

Надежность и эффективность в технике

Справочник в десяти томах

Надежность и эффективность в технике

Справочник в десяти томах

Редакционный совет:

*В.С. АВДУЕВСКИЙ, акад. АН СССР
(председатель);
В.И. КУЗНЕЦОВ, акад. АН СССР;
Н.Д. КУЗНЕЦОВ, акад. АН СССР;
В.А. МЕЛЬНИКОВ, акад. АН СССР;
В.П. МИШИН, акад. АН СССР;
В.Ф. УТКИН, акад. АН СССР;
К.В. ФРОЛОВ, акад. АН СССР;
Б.В. ГНЕДЕНКО, акад. АН УССР;
И.Н. КОВАЛЕНКО, акад. АН УССР;
Б.Ф. ЛОМОВ, чл.-корр. АН СССР*

Том
2

*Математические
методы в теории
надежности
и эффективности*

*Под редакцией акад. АН УССР
Б. В. ГНЕДЕНКО*



Москва
«МАШИНОСТРОЕНИЕ»
1987

ББК 30.14
Н17
УДК 621-192 (035)

Авторы тома: В. В. Белов, Ю. К. Беляев, А. Г. Давтян, В. А. Каштанов, И. Н. Коваленко, В. Г. Кривуца, Н. Ю. Кузнецов, Г. В. Мартынов, В. И. Питербарг, А. Д. Соловьев, В. М. Хаметов, Е. В. Чепурин, Н. А. Шимко.

Рецензенты: кандидаты физ.-мат. наук О. П. Виноградов, С. В. Жуленев

Надежность и эффективность в технике: Справочник.
Н17 В 10 т./Ред. совет: В. С. Авдуевский (пред.) и др. — М.: Машиностроение, 1987. — (В пер.),

Т. 2: Математические методы в теории надежности и эффективности/Под ред. Б. В. Гнеденко. — 280 с.: ил. — 1 р. 60 к.

Даны математические методы детерминированного и стохастического анализа, используемые при исследовании, оценке и контроле надежности и эффективности.

Предназначен для инженерно-технических работников, занятых проектированием, изготовлением, испытаниями и эксплуатацией техники. Будет полезен студентам и преподавателям высших технических учебных заведений.

Н 2701000000-612 Подписное
038 (01)-87

ББК 30.14

СПРАВОЧНИК СПЕЦИАЛИСТА

Владимир Владимирович Белов, Юрий Константинович Беляев, Александр Георгиевич Давтян и др.

НАДЕЖНОСТЬ И ЭФФЕКТИВНОСТЬ В ТЕХНИКЕ

ТОМ 2

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ТЕОРИИ НАДЕЖНОСТИ И ЭФФЕКТИВНОСТИ

Редактор Т. С. Грачева. Художественный редактор С. С. Водчиц.
Переплет художника А. Я. Михайлова. Технический редактор Е. П. Смирнова.
Корректоры: Н. Г. Богомолова, Л. Л. Георгиевская

ИБ № 4673

Сдано в набор 28.04.86. Подписано в печать 24.09.86. Т-15429. Формат 60×90^{1/16}.
Бумага офсетная № 2. Гарнитура литературная. Печать офсетная. Усл. печ. л. 17,5.
Усл. кр.-отт. 17,5. Уч.-изд. л. 23, 84. Тираж 12712 экз. Заказ 92. Цена 1 р. 60 к.

Ордена Трудового Красного Знамени издательство «Машиностроение», 107076, Москва, Стромынский пер., 4.

Ленинградская типография № 6 ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 193144, г. Ленинград, ул. Моисеенко, 10.

Оглавление

Предисловие	8	Глава 5. ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ (<i>В. А. Каштанов</i>)	65
часть I			
МАТЕМАТИЧЕСКИЕ МЕТО- ДЫ АНАЛИЗА	11	1. Основные понятия	65
Глава 1. ЭЛЕМЕНТЫ ТЕО- РИИ МНОЖЕСТВ (<i>В. А. Каштанов</i>)	11	2. Элементарные формулы	69
1. Основные понятия	11	3. Случайные величины	70
2. Операции над множествами	13	4. Системы случайных величин	74
3. Классы (системы, множест- ва) множеств	16	5. Функции от случайных ве- личин	76
4. Упорядоченные множества	17	6. Числовые характеристики случайных величин	78
5. Функции	18	7. Неравенства Чебышева и Колмогорова	81
Глава 2. ЭЛЕМЕНТЫ МАТЕ- МАТИЧЕСКОЙ ЛО- ГИКИ (<i>Н. А. Шимко</i>)	21	8. Сходимость последователь- ностей случайных величин	82
1. Булевы функции	21	9. Законы больших чисел	82
2. Исчисление высказываний	25	10. Предельные теоремы	84
3. Исчисление предикатов	28	Глава 6. ЭЛЕМЕНТЫ ТЕО- РИИ СЛУЧАЙНЫХ ПРОЦЕССОВ (<i>Ю. К. Беляев,</i> <i>В. И. Питербарг</i>)	86
Глава 3. ЭЛЕМЕНТЫ ТЕО- РИИ ГРАФОВ (<i>В. В. Белов</i>)	32	1. Основные понятия и класси- фикация случайных процес- сов	86
1. Основные определения и по- нятия	32	2. Марковские цепи	87
2. Операции над графами	37	3. Процессы восстановления. Случайные точечные про- цессы	91
3. Специальные типы графов	41	4. Мартингалы	93
4. Матрицы графа. Бинарные отношения и графы	45	5. Процессы с ортогональными приращениями. Стохастиче- ские интегралы. Винеров- ский процесс	95
Список литературы	47	6. Стационарные последова- тельности и процессы	96
Глава 4. ЭЛЕМЕНТЫ ТЕО- РИИ ДИФФЕРЕН- ЦИАЛЬНЫХ УРАВ- НЕНИЙ (<i>А. Г. Давтян</i>)	48	7. Гауссовские случайные про- цессы	99
1. Обыкновенные дифференци- альные уравнения	48	Список литературы	102
2. Уравнения в частных произ- водных первого порядка . .	62		

Глава 7. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ (Ю. К. Беляев)	103		
1. Функция правдоподобия	103		
2. Статистики	104		
3. Оценки	106		
4. Показатели качества оценок	108		
5. Доверительное оценивание	111		
6. Проверка статистических гипотез	112		
7. Инвариантность	114		
Список литературы	115		
Глава 8. МАТЕМАТИЧЕСКОЕ ПРОГРАММИРОВАНИЕ (В. М. Хаметов)	116		
1. Элементы выпуклого анализа	116		
2. Выпуклое программирование	118		
3. Линейное программирование (ЛП)	120		
4. Динамическое программирование	124		
Список литературы	128		
ЧАСТЬ II			
МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ И ЭФФЕКТИВНОСТИ	129		
Глава 9. НАДЕЖНОСТЬ ЭЛЕМЕНТА (А. Д. Соловьев)	129		
1. Надежность невозстанавливаемого элемента	129		
2. Распределения, используемые в теории надежности	130		
3. Надежность восстанавливаемого элемента	131		
Глава 10. НАДЕЖНОСТЬ НЕВОССТАНАВЛИВАЕМЫХ СИСТЕМ (А. Д. Соловьев)	136		
1. Надежность системы с независимо отказывающими элементами	136		
2. Метод путей и сечений. Рекуррентный метод	138		
3. Модели зависимости отказов элементов в системе	141		
4. Марковские однородные процессы с конечным или счетным числом состояний. Процессы чистой гибели. Процесс рождения и смерти	142		
Список литературы	148		
Глава 11. НАДЕЖНОСТЬ ВОССТАНАВЛИВАЕМЫХ СИСТЕМ (А. Д. Соловьев)	149		
1. Регенерирующие процессы. Предельные теоремы	149		
2. Регенерирующий процесс специального типа. Марковские вверх процессы	152		
3. Сходимость случайных величин с рациональным преобразованием Лапласа к показательному распределению	154		
4. Общая модель резервирования с восстановлением	157		
5. Модель сложной восстанавливаемой системы	160		
Глава 12. МОДЕЛИ ОПТИМАЛЬНОГО ОБСЛУЖИВАНИЯ (В. А. Каштанов)	163		
1. Физическое описание задачи технического обслуживания	163		
2. Математическая постановка задачи	164		
3. Управляемые полумарковские процессы	168		
4. Частные случаи	172		
5. Управление регенерирующим случайным процессом	173		
6. Примеры моделей технического обслуживания	175		
Глава 13. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ О НАДЕЖНОСТИ ОБОРУДОВАНИЯ В ЭКСПЛУАТАЦИИ (Е. В. Чепурин)	181		
1. Анализ надежности показательных последовательных систем с независимыми компонентами (случай двух компонент)	181		

2. Анализ надежности показательных последовательных систем с зависимыми компонентами	183	3. Моделирование цепи Маркова с непрерывным временем	213
3. Анализ систем, описываемых марковским процессом с дискретным временем	187	4. Моделирование полумарковских процессов	214
4. Анализ систем, описываемых марковским процессом с непрерывным временем	189	5. Точность оценок и количество реализаций	215
5. Анализ систем, описываемых полумарковским процессом	191	6. Некоторые общие методы уменьшения дисперсии	219
6. Анализ систем, описываемых процессом восстановления	192	7. Аналитико-статистические методы	223
Глава 14. СТАТИСТИЧЕСКИЙ КОНТРОЛЬ КАЧЕСТВА (Ю. К. Беляев)	195	8. Оценка распределения времени безотказной работы систем, описываемых регенерирующими процессами	225
1. Планы контроля	195	9. Оценка нестационарных характеристик надежности систем аналитико-статистическим методом	227
2. Числовые показатели планов контроля	198	10. Кусочно-линейный агрегат	237
3. Априорные распределения числа дефектных изделий	200	11. Моделирование кусочно-линейных агрегатов	238
4. Системы планов контроля	201	12. Основные системы моделирования	239
5. Последующие оценки	202	Список литературы	249
6. Контроль непрерывного потока штучной продукции	204	Глава 16. АЛГОРИТМЫ ВЫЧИСЛЕНИЯ ВЕРОЯТНОСТНЫХ РАСПРЕДЕЛЕНИЙ (Г. В. Мартынов)	251
Список литературы	207	1. Непрерывные распределения	251
Глава 15. МЕТОД СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ (МЕТОД МОНТЕКАРЛО) (И. Н. Коваленко, Н. Ю. Кузнецов, В. Г. Кривуца)	208	2. Нормальное распределение	255
1. Моделирование случайных величин	208	3. Распределение χ^2 и его обобщения	267
2. Моделирование пуассоновского потока с параметром $\lambda(t)$, $\int_0^{\infty} \lambda(t) dt = \infty$	213	4. t -распределение Стьюдента	269
		5. F -распределение	271
		6. Бета-распределение (Е. В. Чепурин)	273
		7. Дискретные распределения (Е. В. Чепурин)	274
		Список литературы	
		Предметный указатель	277

Предисловие

Математика неотделима от проблем теории надежности. Точные количественные методы необходимы при расчете надежности сложных систем, при испытаниях на надежность, при проверке гипотез о характере потери изделиями их надежности. Без глубокого математического анализа невозможно прогнозировать надежность изделия при длительной эксплуатации. Более того, само определение надежности и выделение показателей надежности нуждаются в математических средствах. Одними описательными и экспериментальными средствами нельзя обойтись при постановке и решении многочисленных и разнообразных по своему характеру задач теории и практики надежности. При таком многообразии назначений математики в теории надежности естественно разобратся, какие же математические дисциплины представляют основной интерес для теории надежности?

Представление о том, что каждое изделие в период существования проходит ряд последовательных изменений своих состояний, приводит к понятию множества состояний технической системы и изменениях их со временем. Выделение подмножества состояний, принадлежность к которому считается отказом системы, подчеркивает важность представлений теории множеств для описания простейших и одновременно центральных понятий теории надежности. Именно по этой причине в настоящем томе некоторое внимание уделено понятию множества и действиям с множествами.

Работы Дж. Неймана и его последователей над проблемой создания высоконадежных систем из малонадежных элементов привели к широкому использованию средств математической

логики в задачах теории надежности. Позднее выяснилось, что методы математической логики находят многочисленные и фундаментальные применения во многих задачах теории надежности. Естественно, что мы не могли в нашем томе обойти математическую логику с ее логическими операциями.

Для всего процесса математизации наших знаний огромную роль играет математический анализ. Собственно, добрая половина всех математических моделей реальных явлений непосредственно связана с математическим анализом и такими его частями, как дифференциальные, интегральные и интегродифференциальные уравнения. Этот исключительно мощный и гибкий математический аппарат позволяет во многих случаях дать достаточно адекватное описание действительно происходящих процессов изменения состояния вещества технического изделия, а также сопряжения его узлов и деталей. Заслуживает внимания и тот факт, что когда мы используем представления о стохастическом процессе развития, описание вероятностного характера изменений удается дать также с помощью средств классического математического анализа.

Большой круг проблем теории надежности связан с оптимизацией тех или иных решений. Для примера при отыскании неисправностей в сложной технической системе нельзя действовать наобум или же путем перебора всех элементов, для чего требуется очень большое время для проверки, к тому же этот прием малопродуктивен. Попытки найти более эффективные правила проверки приводят к постановке задачи о наиболее рациональных процедурах поиска неисправ-

ности. Такие процедуры в ряде задач найдены; они позволили принципиально сократить время поиска. Задачи оптимизации составляют важный раздел нашего тома. В нем изложены не только методы разыскания максимума и минимума, но и многочисленные современные приемы — линейное и нелинейное программирование, методы оптимального управления и др. Это большой и весьма важный для теории надежности набор подходов к задачам оптимизации.

Элементы теории графов позволяют часто наглядно и просто представлять процесс изменений, происходящих в технической системе.

Хорошо известно, что реальные устройства несут в себе элементы случайности. Это связано с существом дела, поскольку локальные неоднородности вещества, из которого создаются технические системы, приводят к неравномерности снашивания, старения, а вместе с тем и к разбросу сроков службы изделий, изготовленных, казалось бы, в тождественных условиях. Но это не все, поскольку в процессе изготовления допускаются, пусть каждый раз и небольшие, но отклонения от заданных размеров, небольшие нарушения в температурной или иной обработке. В результате качество изделий оказывается неоднородным, разбросанным по случаю. Теория надежности во всей ее широте и глубине основана на теории вероятностей как для расчетов, так и для определения самих понятий.

Теория случайных процессов по своему существу тесно связана с процессами потери надежности техническими системами. Дело в том, что как процессы изнашивания, так и процессы старения должны рассматриваться как случайные процессы, к каждому из которых необходимо найти свой подход. Например, процесс потери надежности изоляции электрического кабеля существенно отличается от процесса потери надежности электрического генератора или же двигателя самолета. Вот почему для тех, кто занимается проблемами теории надежности, важно и необходимо не просто познакомиться с методами уже существующей теории случай-

ных процессов, но научиться создавать новые представления и новые методы в уже созданном круге идей. Иными словами, требуется творческое освоение математической теории.

Ряд задач теории надежности — вопросы резервирования, определения необходимых для поддержания надежности парка машин запасных узлов и деталей, проблемы профилактического обслуживания и другие — оказался неразрывно связанным с постановками проблем и методами, разработанными в теории массового обслуживания. Про теорию массового обслуживания можно сказать, что она стала как бы главной теории надежности.

Математическая статистика, ставшая в наши дни одной из центральных частей прикладной математики, в теории надежности заняла особую роль. Прежде всего с ней связана обработка первичного экспериментального материала, служащего фундаментом для теоретических обобщений. Мало провести экспериментальные исследования или испытания, необходимо их результаты проанализировать и сделать выводы, соответствующие реальному течению процессов. Без математической статистики и ее методов этого сделать невозможно. Далее следует разработать рациональные правила проведения испытаний и последующей обработки полученных данных. В частности, необходимо разработать правила испытаний на надежность, которые проводятся на разных этапах создания технических систем: 1) на этапе разработки, 2) изготовления, 3) приемки, 4) проверки сохранности надежности после транспортирования или хранения. Математическая статистика необходима во всех инженерных дисциплинах и для целей проверки качества моделей и разработанных теорий, иными словами, для проверки соответствия теории реальному протеканию изучаемых явлений.

Еще одно направление использования методов математической статистики в практике надежности выделено особо — речь идет о статистических методах контроля качества массовой продукции и управлении качеством в процессе производства. Эти методы

уже нашли широкое применение, но их будущее должно быть несравненно большим. Особенно важно подчеркнуть значение методов управления качеством, поскольку их цель состоит в том, чтобы не допустить саму возможность изготовления некачественных изделий.

Таким образом, практически все разделы современной математики находят применение. Возможно только теория чисел и теория функций действительного переменного в обозримом будущем будут принимать в развитии теории надежностей лишь скромное участие.

Теория надежности — молодая наука, она не насчитывает и пятидесяти лет. Развиваться же ей предстоит столько времени, сколько будет существовать промышленность и массовое производство. При этом будут

совершенствоваться и математические методы, которые используются в теории надежности. Более того, несомненно, теория надежности выдвинет такие проблемы, которые заставят прогрессировать саму математику. Для того чтобы это случилось, необходимо на каждом этапе развития науки подводить итоги того, что в ней уже достигнуто. Такое понимание, что уже сделано и на что наука уже способна, как раз и является предпосылкой для дальнейшего ее процветания и совершенствования. Надеюсь, что настоящий том внесет свою долю не только в ознакомление с тем, что уже достигнуто в математической теории надежности, но и в понимание того, что предстоит срочно дополнительно сделать.

Б. В. ГНЕДЕНКО,
акад. АН УССР

Часть I. МАТЕМАТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА

Глава 1. Элементы теории множеств

1. ОСНОВНЫЕ ПОНЯТИЯ

1.1. МНОЖЕСТВА

Множеством будем называть совокупность, набор некоторых элементов. При таком интуитивном определении понятие множества в строгом математическом смысле следует считать неопределяемым и первичным. В рамках неформального подхода будем называть элементами множества те объекты, из которых оно составлено;

Множества обозначают большими буквами $A, B, \dots, X, Y, \dots, \mathfrak{A}, \Omega, \dots$

Элементы множества обозначают малыми буквами $a, b, \dots, x, y, \dots, \alpha, \beta, \dots, \omega, \dots$,

Способы задания множеств. Множество может быть задано перечислением всех его элементов:

$$\stackrel{\text{def}}{A} = \{a_1, a_2, \dots, a_n\};$$

$$\stackrel{\text{def}}{B} = \{b_1, b_2, \dots, b_m, \dots\}.$$

Множество может быть задано перечислением тех свойств (условий), которыми обладают все элементы определяемого множества:

$$\stackrel{\text{def}}{A} = \{a : U(a)\}.$$

Читается: A есть множество элементов, для которых выполняются условия $U(a)$.

Замечание. Символ $\stackrel{\text{def}}{=}$ обозначает символ «определения», т. е. равенство левой и правой части выполняется по определению (иногда «def» опускается).

Символ включения. Если некоторый элемент a принадлежит множеству A , то записывают $a \in A$. Если некоторый элемент b не принадлежит множеству B , то записывают $b \notin B$ или $b \bar{\in} B$.

Пример множеств:
натуральных чисел $N = \{0, 1, 2, \dots\}$;
целых чисел $Z = \{0, \pm 1, \pm 2, \dots\}$;
действительных чисел $R^{(1)} = \{x : -\infty < x < +\infty\}$;
положительных чисел $R^+ = \{x : x > 0\}$;

четных чисел $\left\{x : \frac{x}{2} \in N\right\}$;

корней алгебраического уравнения $\sum_{i=0}^n a_i x^i = 0$, т. е. $\left\{x : \sum_{i=0}^n a_i x^i = 0\right\}$;

n -мерных векторов с действительными координатами

$$R^{(n)} = \{x = (x_1, \dots, x_n) :$$

$$x_k \in R^{(1)}, k = 1, 2, \dots, n\}.$$

Подмножеством B множества A называют любую совокупность (множество) элементов, каждый из которых принадлежит множеству A . Если множество B есть подмножество множества A , то записывают $B \subset A$. Соотношение $B \subset A$ читается: множество B включается в множество A . Истинным подмножеством множества A называют подмножество B , для которого $A \setminus B \neq \emptyset$. (Операция разности множеств и пустое множество определены ниже).

Примеры подмножеств:

само множество A является подмножеством A , т. е. $A \subset A$;

множество $\{a\}$, состоящее из одного элемента a , принадлежащего A ($a \in A$), является подмножеством A , т. е. $\{a\} \subset A$;

множество четных чисел есть подмножество множества натуральных чисел, т. е. $\left\{x : \frac{x}{2} \in N\right\} \subset N$.

Множество A равно множеству B (записывается $A = B$), если одновременно выполняются два условия:

A есть подмножество B , т. е. $A \subset B$;

B есть подмножество A , т. е. $B \subset A$.

Замечание. Если учесть определение подмножества, то равенство двух множеств можно определить следующим образом: множество A равно множеству B , если любой элемент множества A принадлежит множеству B , и, наоборот, любой элемент множества B принадлежит множеству A .

Пустым множеством называют множество, не содержащее ни одного элемента. Пустое множество обозначают символом \emptyset .

Замечания: 1. Определение пустого множества не подпадает под определение 1.1. Поэтому то, что этот объект есть множество, постулируется.

2. По определению считаем, что пустое множество \emptyset есть подмножество любого множества A , т. е. имеет место включение $\emptyset \subset A$.

Примеры: множество действительных корней уравнения $x^2 + 1 = 0$;

$$\{x : e^x < 0\} = \emptyset;$$

$$\{(x, y) : x^2 + y^2 < 0\} = \emptyset.$$

множество пар действительных чисел, для которых выполняется неравенство $x^2 + y^2 < 0$.

1.2. МОЩНОСТЬ МНОЖЕСТВА

Конечные и бесконечные множества. Множество, содержащее конечное число элементов, называют конечным. Множество A называют бесконечным, если для любого $N > 0$ существует такое конечное истинное подмножество B множества A , для которого число элементов равно N .

Эквивалентность множеств. Два конечных множества называют эквивалентными, если они содержат равное число элементов. Два бесконечных множества называют эквивалентными, если между элементами этих множеств можно установить взаимно однозначное соответствие, т. е. каждому элементу одного множества поставить в соответствие один и только один элемент другого множества и наоборот.

Счетные множества. Множество A называется счетным, если можно установить взаимно однозначное соответствие между элементами множества натуральных чисел $N = \{1, 2, 3, \dots, n, \dots\}$ и элементами множества A . Это соответствие записывается в виде $(a_1, a_2, \dots, a_n, \dots)$.

С в о й с т в а: всякое подмножество счетного множества либо конечно, либо счетно; сумма * конечного или счетного числа счетных множеств есть счетное множество.

Примеры счетных множеств:

четные числа $\left\{x : \frac{x}{2} \in N\right\}$;

целые числа $Z = \{0, \pm 1, \pm 2, \pm k, \dots\}$;

рациональные числа;

точки на плоскости, имеющие рациональные координаты.

Несчетные множества. Бесконечное множество, не являющееся счетным, называют несчетным.

С в о й с т в а: 1) несчетное множество содержит счетное подмножество; 2) бесконечное множество эквивалентно некоторому своему истинному подмножеству.

Примеры несчетных множеств:

действительные числа, принадлежащие отрезку $[0, 1]$, $\{x : x \in [0, 1]\}$;

* Определение суммы дано на с. 13.

действительные числа, $\{x \mid x \in R^{(1)}\}$; точки на плоскости.

Мощность множества есть характеристика количества его элементов. Для конечного множества, когда определено число элементов, мощностью называют число элементов. Для бесконечного множества, когда число его элементов не определено, возможно только сравнение его с другими множествами (см. определение эквивалентности с. 12).

Если A и B — два произвольных множества, $m(A)$ и $m(B)$ — мощности этих множеств, то по определению:

для эквивалентных множеств $m(A) = m(B)$;

если A эквивалентно некоторому подмножеству B и нет ни одного подмножества A , которое было бы эквивалентно B , то $m(A) < m(B)$.

Так как все счетные множества эквивалентны, то говорят, что множество, эквивалентное счетному, имеет мощность счетного множества. Множества, эквивалентные множеству действительных чисел между нулем и единицей, имеют мощность континуума.

Мощность множества действительных чисел между 0 и 1 больше мощности счетного множества.

Если A — произвольное множество и \mathfrak{A} — множество всевозможных подмножеств множества A , то имеет место неравенство $m(A) < m(\mathfrak{A})$.

Счетное множество имеет наименьшую мощность среди всех бесконечных множеств.

2. ОПЕРАЦИИ НАД МНОЖЕСТВАМИ

2.1. ОБЪЕДИНЕНИЕ, ПЕРЕСЕЧЕНИЕ, РАЗНОСТЬ МНОЖЕСТВ. ДОПОЛНЕНИЕ

Объединением (суммой) множеств A_1, A_2, \dots, A_n называют множество B , состоящее из всех тех элементов, каждый из которых принадлежит хотя бы одному из A_1, \dots, A_n . Записывается: $B = A_1 \cup A_2 \cup \dots \cup A_n$ или $B = \bigcup_{k=1}^n A_k$. Из этого определения при $n = 2$ получаем определение объединения двух множеств: объединением двух множеств A_1 и A_2 называют множество B , состоящее из всех тех эле-

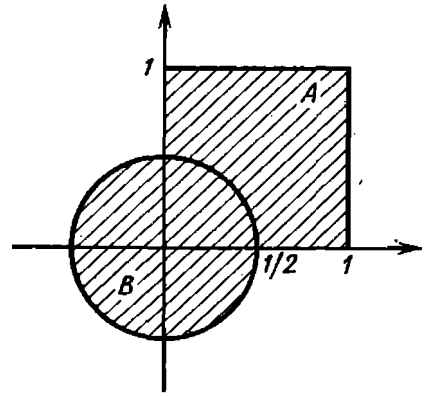


Рис. 1. Пример объединения двух множеств

ментов, которые принадлежат либо A_1 , либо A_2 , либо обоим множествам одновременно $B = A_1 \cup A_2$.

Для бесконечного набора множеств $\{A_k\}, k \in N$ (или $k \in R^{(+)}$) объединением $\bigcup_{k \in N} A_k$ ($\bigcup_{k \in R^+} A_k$) называют

множество B , состоящее из всех тех элементов, каждый из которых принадлежит хотя бы одному множеству заданного набора $\{A_k\}, B = \bigcup_{k \in N} A_k$;

$$B = \bigcup_{k \in R^+} A_k.$$

Замечания: 1. Из равенства $B = \bigcup_k A_k$ следует, что для любого

элемента x множества B найдется такой индекс k_0 , принадлежащий множеству индексов (в приведенных выше определениях $k_0 \in \{1, 2, \dots, n\}; k_0 \in N; k_0 \in R^+$), что $x \in A_{k_0}$.

2. Из равенства $B = \bigcup_k A_k$ следует, что для любого k_0 , принадлежащего набору индексов, элемент x , принадлежащий $A_{k_0}, x \in A_{k_0}$, принадлежит множеству B , т. е. $x \in B$.

Свойства:

для любого множества $A \quad A \cup A = A, A \cup \emptyset = A$;

если $B \subset A$, то $A \cup B = A$;

коммутативность операции объединения, т. е. $A \cup B = B \cup A$;

ассоциативность операции объединения, т. е. $(A \cup B) \cup C = A \cup (B \cup C)$.

Все приведенные свойства следуют из определения объединения множеств и равенства множеств.

Пример. На рис. 1 приведен пример объединения двух множеств: мно-

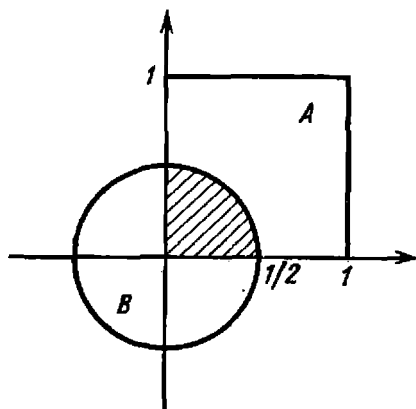


Рис. 2. Пример пересечения двух множеств

жества A — квадрата со стороной единица и множества B — круга радиуса $r = 1/2$. Множества $C = A \cup B$ заштриховано.

Пересечением множеств A_1, A_2, \dots, A_n называют множество B , состоящее из всех тех элементов, которые принадлежат всем множествам набора $\{A_k\}$. Записывается: $B = A_1 \cap A_2 \cap \dots \cap A_n$, $B = A_1, A_2, \dots, A_n$ или $B = \bigcap_{k=1}^n A_k$.

При $n = 2$ из этого определения можно получить определение пересечения двух множеств: пересечением двух множеств A_1 и A_2 называют множество B , состоящее из всех точек, принадлежащих одновременно и A_1 и A_2 , $B = A_1 \cap A_2$.

Для бесконечного набора множеств $\{A_k\}$, $k \in N$ (или $k \in R^+$) пересечением $\bigcap_{k \in N} A_k$ ($\bigcap_{k \in R^+} A_k$) называют множество B , состоящее из всех тех элементов, которые принадлежат всем множествам набора $\{A_k\}$; $B = \bigcap_{k \in N} A_k$

или $B = A_1, A_2, \dots, A_n, \dots$

Замечания: 1. Из равенства $B = \bigcap_{k=1}^n A_k$ следует, что для любого элемента $x \in B$ имеет место включение $x \in A_1, x \in A_2, \dots, x \in A_n$ (аналогично при $k \in N$ или $k \in R^+$).

2. Из равенства $B = \bigcap_{k=1}^n A_k$ и включений $x \in A_1, x \in A_2, \dots, x \in A_n$ следует $x \in B$.

Свойства:

для любого множества A : $A \cap A = A$, $A \cap \emptyset = \emptyset$;

если $B \subset A$, то $A \cap B = B$;

коммутативность операции пересечения, т. е. $A \cap B = B \cap A$;

ассоциативность операции пересечения, т. е. $(A \cap B) \cap C = A \cap (B \cap C)$;

если $A_k \subset B_k$, $k \in N$, тогда $\bigcap_{k \in N} A_k \subset \bigcap_{k \in N} B_k$.

Все перечисленные свойства следуют из определения пересечения множества (см. с. 14) и равенства множества.

Пример. На рис. 2 приведен пример пересечения двух множеств A и B (определения A и B см. на с. 13). Множество $C = A \cap B$ заштриховано.

Непересекающиеся множества. Два множества A и B называют непересекающимися, если их пересечение есть пустое множество, т. е. $A \cap B = \emptyset$.

Разностью множеств A и B называют множество C , состоящее из всех элементов множества A , которые не принадлежат множеству B . Записывается: $C = A \setminus B$.

Пример. На рис. 3 приведен пример разности двух множеств A и B (множества A и B определены на с. 13). Множество $C = A \setminus B$ заштриховано.

Свойства:

если $A \cap B = \emptyset$, то $A \setminus B = A$ и $B \setminus A = B$;

если $A \subset B$ ($A \cap B = A$), то $A \setminus B = \emptyset$;

для любых множеств A и B справедливо равенство $A \cap (B \setminus A) = \emptyset$, т. е. множества A и $(B \setminus A)$ непересекающиеся.

Указанные свойства вытекают из определения разности множеств.

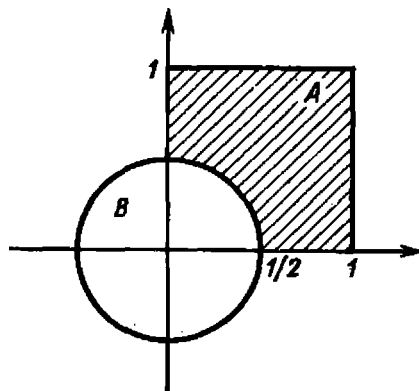


Рис. 3. Пример разности двух множеств

Дополнение. Пусть рассматриваются только подмножества некоторого множества E . Дополнением множества A в E (чаще *дополнением* множества A) называют множество элементов E , не принадлежащих множеству A , т. е. дополнение есть разность $E \setminus A$. Обозначается $\bar{A} \stackrel{\text{def}}{=} E \setminus A$.

Свойства. Для любого множества A :

множества A и \bar{A} непересекающиеся,
 $A \cap \bar{A} = \emptyset$;

выполняется равенство $A \cup \bar{A} = E$;

если $A \subset B$, то $\bar{A} \supset \bar{B}$.

При $A = E$ имеем:

дополнение множества E есть пустое множество \emptyset , т. е. $\bar{E} = \emptyset$;

дополнение пустого множества \emptyset есть множество E , т. е. $\bar{\emptyset} = E$.

Для любого набора $\{A_k\}$ справедливы равенства:

дополнение объединения некоторых множеств равно пересечению дополнений этих множеств $\overline{\bigcup_k A_k} = \bigcap_k \bar{A}_k$;

дополнение пересечения некоторых множеств равно объединению дополнений этих множеств $\overline{\bigcap_k A_k} = \bigcup_k \bar{A}_k$.

Принцип двойственности. Если справедливо некоторое соотношение между множествами, имеющее вид равенства или включения и выраженное в терминах объединений, пересечений и дополнений, то справедливо и соотношение такого же рода, которое получается из исходного, если в нем операции объединения \cup , пересечения \cap , включения \subset и \supset заменить соответственно на операции пересечения \cap , объединения \cup , включения \supset и \subset , равенства сохранить, а каждое множество заменить на его дополнение.

2.2. ПРЕДЕЛЫ МНОЖЕСТВ

Верхний предел последовательности множеств. Для последовательности множеств $\{A_k, k \in N\}$ верхним пределом A^* называют множество всех элементов, каждый из которых принадлежит бесконечно многим множествам A_k . Обозначается: $A^* = \limsup_k A_k$.

Для множества A^* справедливо равенство

$$A^* = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

Нижний предел последовательности множеств. Для последовательности множеств $\{A_k, k \in N\}$ нижним пределом A_* называют множество всех элементов, каждый из которых принадлежит всем A_k , за исключением конечного числа множеств. Обозначается: $A_* = \liminf_k A_k$. Это определение означает, что для каждого элемента x множества A_* найдется такой номер m , что x принадлежит всем $A_k, k \geq m$,

т. е. $x \in \bigcap_{k=m}^{\infty} A_k$.

Для множества A_* справедливо равенство

$$A_* = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Для множеств A^* и A_* справедливо включение $A^* \supset A_*$.

Предел последовательности множеств. Если последовательность множеств $\{A_k, k \in N\}$ такова, что $A^* = A_*$, то множество A^* (или A_*) называют пределом последовательности множеств $\{A_k\}$.

Если $A^* = A_*$, то говорят, что предел последовательности множеств $\{A_k\}$ существует. Если $A^* \neq A_*$, то говорят, что предела последовательности множеств $\{A_k\}$ не существует.

Монотонные последовательности множеств. Последовательность множеств $\{A_k, k \in N\}$ называют *возрастающей*, если при любом k имеет место включение: $A_k \subset A_{k+1}$. Предел *возрастающей* последовательности множеств существует и равен $\bigcup_{k=1}^{\infty} A_k$.

Последовательность множеств $\{A_k, k \in N\}$ называют *убывающей*, если при любом k имеет место включение $A_k \supset A_{k+1}$. Предел *убывающей* последовательности множеств существует и равен $\bigcap_{k=1}^{\infty} A_k$.

3. КЛАССЫ (СИСТЕМЫ, МНОЖЕСТВА) МНОЖЕСТВ

3.1. КОЛЬЦО И АЛГЕБРА МНОЖЕСТВ

Непустое множество (система, класс) R подмножеств некоторого множества E называют кольцом, если из условий $A \in R$ и $B \in R$ следуют включения:

1. $A \cup B \in R$; 2. $A \setminus B \in R$.

Иногда говорят, что *кольцо* есть *непустой класс подмножеств, замкнутый относительно операции объединения двух множеств и вычитания*.

Свойства:

всякое кольцо R содержит пустое множество \emptyset , так как если $A \in R$, то $A \setminus A = \emptyset \in R$;

кольцо R есть *непустой класс подмножеств, замкнутый относительно операции пересечения двух элементов*, т. е. из условий $A \in R$ и $B \in R$ следует $A \cap B \in R$, так как $A \cap B = (A \cup B) \setminus ((A \setminus B) \cup (B \setminus A))$;

если $A_i \in R$ $i = 1, 2, \dots, n$, то $\bigcup_{i=1}^n A_i \in R$, $\bigcap_{i=1}^n A_i \in R$, т. е. *кольцо есть непустой класс подмножеств, замкнутый относительно операции объединений и пересечения в конечном числе*;

пересечение любого множества колец есть кольцо.

Примеры колец:

пара множеств $\{\emptyset, E\}$;

система множеств $\{A, \bar{A}, \emptyset, E\}$ $A \subset E$;

множество всех подмножеств множества E ;

класс всевозможных множеств вида

$$\bigcup_{k=1}^n \{x: -\infty < a_k \leq x < b_k < +\infty\},$$

т. е. класс всевозможных конечных объединений ограниченных интервалов, замкнутых слева и открытых справа;

класс всех конечных или счетных подмножеств произвольного множества E .

Определение 1. *Непустой класс R подмножеств множества E называют алгеброй, если:*

1) из условий $A \in R$, $B \in R$ следует $A \cup B \in R$; 2) из условия $A \in R$ следует $\bar{A} \in R$.

Свойства:

всякая алгебра R содержит множество E , $E \in R$, так как $E = A \cup \bar{A}$; из условий $A \in R$, $B \in R$ следует

$A \cap B \in R$, так как $A \cap B = \overline{\bar{A} \cup \bar{B}}$ (см. принцип двойственности с. 15);

если $A_k \in R$, $k = 1, 2, \dots, n$, то $\bigcup_{k=1}^n A_k \in R$, $\bigcap_{k=1}^n A_k \in R$;

всякая алгебра есть кольцо (обратное неверно).

Множество E называют *единицей* некоторой системы множеств R , если $E \in R$, и для любого $A \in R$ имеет место равенство $A \cap E = A$.

Определение 2. Кольцо множеств с единицей называют алгеброй.

3.2. σ -КОЛЬЦО И σ -АЛГЕБРА

Кольцо R подмножеств некоторого множества E называют σ -кольцом, если из условий $A_i \in R$, $i \in N$ следует

$\bigcup_{i=1}^{\infty} A_i \in R$; т. е. σ -кольцо есть *кольцо, замкнутое относительно образования счетных объединений*.

Свойства:

σ -кольцо замкнуто относительно образования счетного пересечения, т. е.

если $A_k \in R$, $k \in N$, то $\bigcap_{k=1}^{\infty} A_k \in R$;

если $A_k \in R$, $k \in N$, то $A^* \in R$; и $A_* \in R$;

пересечение любого числа σ -колец есть σ -кольцо.

σ -алгеброй называют σ -кольцо с единицей.

Свойства:

σ -алгебра подмножеств множества E замкнута относительно образования счетного пересечения и объединения; пересечение любого числа σ -алгебр есть σ -алгебра.

Пусть $E = \{E\}$ — некоторая система множеств, $R(E)$ — все σ -алгебры, для которых $X = \bigcup_{E \in E} E$ есть единица

и любое $E \in E$ принадлежит R , т. е. $E \in R(E)$. Тогда для системы E существует единственная σ -алгебра R_0 , содержащая E и содержащаяся в лю-

бой σ -алгебре, содержащей E , т. е. $R_0 \subset R(E)$, $R_0 = \bigcap_{R \in \mathcal{R}(E)} R$. σ -алгебра R_0 называется минимальной σ -алгеброй над системой E .

Замечания: 1. Говорят, что система E порождает минимальную σ -алгебру R_0 , причем единственную.

2. Элементы минимальной σ -алгебры R_0 (борелевской алгебры множеств), порожденной некоторой системой множеств E , называют борелевскими множествами.

Борелевские множества на прямой. Борелевской алгеброй на прямой называют минимальную σ -алгебру над системой всех сегментов $[a, b]$. Элементы борелевской σ -алгебры на прямой называют борелевскими множествами на прямой.

Определение не изменится, если вместо указанной системы множеств всех сегментов $[a, b]$ взять систему бесконечных интервалов $(-\infty, x)$.

Борелевские множества в $R^{(n)}$. Борелевской алгеброй $\mathfrak{A}^{(n)}$ в $R^{(n)}$ называют минимальную алгебру подмножеств пространства $R^{(n)}$, содержащую множество всевозможных параллелепипедов вида

$$\{(x_1, \dots, x_n) : a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\}, a_i \leq b_i, 1 \leq i \leq n.$$

При определении борелевской алгебры в $R^{(n)}$ можно брать систему множеств вида

$$\{(x_1, \dots, x_n) : -\infty < x_1 \leq b_1, \dots, -\infty < x_n \leq b_n\}.$$

3.3. ИЗМЕРИМОЕ ПРОСТРАНСТВО

Множество E и некоторая σ -алгебра подмножеств \mathfrak{A} множества E образует измеримое пространство (E, \mathfrak{A}) .

Элементы $A \in \mathfrak{A}$ называют измеримыми множествами.

Примеры:

$E = [0, 1]$, \mathfrak{A} — σ -алгебра борелевских подмножеств, принадлежащих $[0, 1]$;

$E = \{e_0, e_1, \dots, e_N\}$ — конечное множество, \mathfrak{A} — множество всевозможных подмножеств множества E .

4. УПОРЯДОЧЕННЫЕ МНОЖЕСТВА.

4.1. ОТНОШЕНИЯ ПОРЯДКА И ЭКВИВАЛЕНТНОСТИ

Говорят, что между a_1 и a_2 ($a_1, a_2 \in A$) установлено отношение порядка O (\succ или \geq и т. п.), если высказывание « a_1 находится в отношении O к a_2 » истинно. Записывается: $a_1 O a_2$ ($a_1 \succ a_2$, $a_1 \geq a_2$ и т. п.).

Примеры отношений порядка:

а) $a_1, a_2 \in R^{(1)}$, $R^{(1)}$ — множество действительных чисел; тогда $a_1 \geq a_2$ есть отношение порядка (a_2 предшествует a_1);

б) $\{a_1, a_2, \dots, a_n\}$ множество сотрудников института, отношение порядка: « a_k есть сотрудник той же лаборатории, что и a_1 »;

с) $\{a_1, a_2, \dots, a_n\}$ — некоторое множество людей, отношение порядка: « a_k знаком с a_m »;

д) отношение порядка « a есть ученик b ».

Отношение порядка O называют рефлексивным, если $x O x$ выполняется для всех $x \in X$.

Примеры а), б), с) есть примеры рефлексивных отношений; пример д) есть пример нереклексивного отношения.

Отношение порядка O называют симметричным, если для любых $x_1, x_2 \in X$ из выполнения $x_1 O x_2$ следует $x_2 O x_1$.

Примеры а), д) не обладают свойством симметричности; примеры б), с) — примеры симметричных отношений.

Отношение O называют транзитивным, если для любых $x_1, x_2, x_3 \in X$ из $x_1 O x_2$ и $x_2 O x_3$ следует $x_1 O x_3$.

Примеры а) и б) есть примеры транзитивных отношений, в примерах с) и д) отношения не обладают свойством транзитивности.

Всякое рефлексивное, симметричное и транзитивное отношение называют отношением эквивалентности.

Отношение эквивалентности разбивает множество X , на котором оно определено, на непересекающиеся подмножества — классы эквивалентности [в примере б) — разбиение сотрудников института на лаборатории].

Всякое рефлексивное транзитивное отношение называют отношением слабого упорядочения.

Для отношения слабого упорядочения O_c , определенного на некотором множестве X , для любой пары $x_1, x_2 \in X$ возможны следующие соотношения:

1) справедливы отношения $x_1 O_c x_2$ и $x_2 O_c x_1$ (симметрия); тогда элементы «подобны», отношение O_c относит их к одному классу эквивалентности;

2) справедливо $x_1 O_c x_2$, но несправедливо $x_2 O_c x_1$; тогда x_1 предшествует x_2 ;

3) справедливо $x_2 O_c x_1$, но несправедливо $x_1 O_c x_2$; тогда « x_2 предшествует x_1 »;

4) несправедливо $x_1 O_c x_2$ и несправедливо $x_2 O_c x_1$; тогда оба элемента несравнимы.

4.2. ЧАСТИЧНО УПОРЯДОЧЕННЫЕ МНОЖЕСТВА

Множество X называют частично упорядоченным, если на нем определено отношение слабого упорядочения O_c и для любых $x_1, x_2 \in X$ отношения $x_1 O_c x_2, x_2 O_c x_1$ одновременно справедливы только при $x_1 = x_2$.

Так как при определении слабого упорядочения в множестве X устанавливаются отношения порядка не обязательно между любой парой элементов множества X (могут быть несравнимые элементы), то определяемый таким образом объект называют частично упорядоченным.

4.3. УПОРЯДОЧЕННЫЕ МНОЖЕСТВА

Множество X называют упорядоченным, если на нем определено отношение слабого упорядочения, причем нет несравнимых элементов и для любых $x_1, x_2 \in X$ одновременное выполнение отношений $x_1 O_c x_2$ и $x_2 O_c x_1$ справедливо только при $x_1 = x_2$.

Примеры: 1. Множество действительных чисел $R^{(1)}$ с отношением порядка $x_1 \geq x_2, x_1, x_2 \in R^{(1)}$ есть упорядоченное множество.

2. Множество подмножеств счетного множества N целых чисел ($N = \{0, 1, 2, \dots\}$) с отношением порядка O по включению ($A_1 O A_2$ означает, что $A_1 \supset$

$\supset A_2, A_1, A_2 \subset N$) есть пример частично упорядоченного множества (есть несравнимые элементы).

5. ФУНКЦИИ

5.1. ОТОБРАЖЕНИЕ

Функцией (отображением) F , определенной на некотором множестве X и принимающей значения из некоторого множества Y , называют однозначное соответствие, при котором каждому элементу x множества X ($x \in X$) сопоставляется (ставится в соответствие) один и только один элемент y множества Y ($y \in Y$).

Это соответствие записывают так: $y = F(x)$.

Множество X называют областью определения функции F ; множество Y называют областью или пространством возможных значений.

В приведенном определении соответствие однозначно в одну сторону, т. е. если $y_1 = F(x_1)$ и $y_2 = F(x_2)$, то и из неравенства $x_1 \neq x_2$ не следует $y_1 \neq y_2$ (может быть $y_1 = y_2$ при $x_1 \neq x_2$).

Отображение F называют взаимно однозначным, если из равенства $F(x_1) = F(x_2)$ следует равенство $x_1 = x_2$ для любых $x_1, x_2 \in X$.

Образом множества A ($A \subset X$) в пространстве Y называют множество всех тех элементов y ($y \in Y$), для каждого из которых существует хотя бы один $x \in X$, такой, что выполняется условие $y = F(x)$. Обозначается $F(A)$. Записывается:

$$F(A) = \bigcup_{x \in A} \{y : y = F(x)\}.$$

Замечания:

образ $F(A)$ есть подмножество множества Y , $F(A) \subseteq Y$;

образ $F(X)$ называют пространством значений функции F (может быть $F(X) \subseteq Y$);

определение образа позволяет любому подмножеству A множества X поставить в соответствие одно и только одно подмножество $F(A)$ множества Y

Функцию F , которая каждому подмножеству E из некоторого класса E подмножеств множества X ставит в соответствие некоторое подмножество

R множества Y , называют **функцией множеств** с областью определения E и значениями из Y

Записывается: $R = F(E)$, $E \in E$, $R \subseteq Y$

Говорят, что функция F определена на измеримом пространстве (X, E) .

В частном случае, если $\{x\} \in E$, то $R = F(x)$, только здесь не требуется, чтобы R состояло из одного элемента y множества Y .

Прообразом множества B ($B \subset Y$) в пространстве X называют множество всех тех элементов x ($x \in X$), для каждого из которых найдется хотя бы один $y \in B$, такой, что имеет место соответствие $y = F(x)$. Если для любого $x \in X$ не найдется указанного в определении $y \in Y$, то считаем прообраз \emptyset . Обозначение прообраза: $F^{-1}(B)$. Записывается:

$$F^{-1}(B) = \bigcup_{y \in B} \{x : F(x) = y\} = \{x : F(x) \in B\}.$$

Замечания:

прообраз $F^{-1}(B)$ есть подмножество множества X ;

очевидно соотношение $F(F^{-1}(B)) = B$;

если $B \cap F(X) \neq \emptyset$, то $F^{-1}(B) \neq \emptyset$;

определение прообраза позволяет подмножеству множества Y поставить в соответствие подмножество множества X .

Функцию множеств $F^{-1}(B)$, определенную на некотором классе подмножеств множества Y ($B \subset Y$) со значениями в пространстве X , называют **обратным отображением** [порожденным функцией $F(A)$], если для любого $B \subset Y$

$$F^{-1}(B) = \{x : F(x) \in B\}.$$

Свойства:

$$F^{-1}(B_1) F^{-1}(B_2) = F^{-1}(B_1 \cap B_2);$$

$$F^{-1}(B_1 \cup B_2) = F^{-1}(B_1) \cup F^{-1}(B_2);$$

если $B_1 \subset B_2$, то $F^{-1}(B_1) \subseteq F^{-1}(B_2)$;

$$F^{-1}\left(\bigcap_{k=1}^{\infty} B_k\right) = \bigcap_{k=1}^{\infty} F^{-1}(B_k);$$

$$F^{-1}\left(\bigcup_{k=1}^{\infty} B_k\right) = \bigcup_{k=1}^{\infty} F^{-1}(B_k);$$

если областью определения обратного отображения является σ -алгебра подмножеств множества Y , то прообразы образуют σ -алгебру подмножеств X .

Функцию F , определенную на элементах множества X с некоторой σ -алгеброй R подмножеств множества X и принимающую значения из $R^{(1)}$, называют **измеримой**, если прообразы всех борелевских множеств являются элементами σ -алгебры R , т. е. для любого борелевского множества B

$$F^{-1}(B) = \{x : F(x) \in B\} \in R.$$

Говорят, что измеримая функция F определена на измеримом пространстве (X, R) , а принимает значения в борелевском измеримом пространстве $(R^{(1)}, \mathfrak{A})$ где \mathfrak{A} — борелевская σ -алгебра.

Свойства:

для того чтобы функция $F(x)$ была измеримой, необходимо и достаточно, чтобы при любом действительном c

$$\{x : F(x) < c\} \in R;$$

предел сходящейся последовательности измеримых функций есть функция измеримая;

сумма, разность и произведение измеримых функций есть функция измеримая;

частное двух измеримых функций при условии, что знаменатель не обращается в ноль, измеримая функция;

измеримая функция от измеримых функций есть функция измеримая.

Пример неизмеримой функции. Функция $F(x, y)$ определена в области $X = \{0 \leq x \leq 1, 0 \leq y \leq 1\}$, принимает значения в области $x \leq y$ $F(x, y) = 1$, в области $x > y$ $F(x, y) = 0$. Если в пространстве X определить σ -алгебру множеств, состоящую из четырех элементов

$$(X, \emptyset, \{(x, y) : 0 \leq x \leq \frac{1}{2},$$

$$0 \leq y \leq 1\}, \{(x, y) : \frac{1}{2} < x \leq 1,$$

$$0 \leq y \leq 1\}),$$

то функция $F(x, y)$ не будет σ -измеримой.

5.2. ПРИМЕРЫ ФУНКЦИЙ МНОЖЕСТВ

Характеристическая функция множества. Функцию $\chi_A(x)$, область определения которой есть множество X , называют **характеристической функцией** множества A или **индикатором** множества A , если

$$\chi_A(x) = \begin{cases} 0, & \text{если } x \in \bar{A} = X \setminus A; \\ 1, & \text{если } x \in A; \quad A \subset X. \end{cases}$$

Свойства:

справедливы тождества $\chi_{\emptyset}(x) \equiv 0$,
 $\chi_X(x) \equiv 1$;

$\chi_A(x) \leq \chi_B(x)$, если $A \subset B$;

$\chi_{A \cap B}(x) = \chi_A(x) \chi_B(x)$;

$\chi_{\bigcap_k A_k}(x) = \prod_k \chi_{A_k}(x)$;

$\chi_{A \cup B}(x) = \chi_A(x) + \chi_B(x) -$

$-\chi_{A \cap B}(x)$;

если $A_i A_j = \emptyset$ при $i \neq j$, то

$\chi_{\bigcup_k A_k}(x) = \sum_k \chi_{A_k}(x)$.

Мера σ -аддитивной мерой называют функцию множества $\mu(A)$, для которой:

1) область определения есть борелевская алгебра R ;

2) значения функции $\mu(A)$ для любого элемента $A \in R$ действительны и неотрицательны;

3) выполняется свойство счетной аддитивности, т. е. для любых множеств $A_1, A_2, \dots, A_n, A_i \in R, i = 1, 2, \dots$, таких, что $A_i A_j = \emptyset$ при $i \neq j$, выполняется равенство

$$\mu\left(\bigcup_1^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

Вероятностной мерой называют σ -аддитивную меру $P(A)$, для которой $P(E) = 1$, где E — единица борелевской алгебры R .

5.3. ФУНКЦИОНАЛЫ

Функционал Φ есть отображение произвольного пространства X в некоторое числовое множество Y

Часто функционалом называют такое отображение, когда элементы множества X есть функции, т. е. функционал ставит в соответствие некоторой функции из множества X число, принадлежащее множеству Y .

Пространство (множество) называют **линейным**, если:

1) для любых двух элементов $x_1 \in X$ и $x_2 \in X$ однозначно определен элемент $x_3 \in X$, который называют суммой $x_3 = x_1 + x_2$, причем $x_1 + x_2 = x_2 + x_1$; $x_0 + (x_1 + x_2) = (x_0 + x_1) + x_2$; существует элемент $0 \in X$, такой, что для любого $x \in X$ справедливо $x + 0 = x$; для каждого $x \in X$ существует элемент $-x \in X$, такой, что $x + (-x) = 0$;

2) для любого числа α и любого $x \in X$ определен элемент $\alpha x \in X$, причем

$$\alpha(\beta x) = (\alpha\beta)x, \quad 1x = x;$$

3) операции сложения и умножения обладают следующими свойствами:

$$(\alpha + \beta)x = \alpha x + \beta x, \quad \alpha(x + y) = \alpha x + \alpha y.$$

Функционал Φ , определенный на некотором линейном пространстве X , называют **линейным**, если для любых $x_1, x_2 \in X$ и любых $\alpha_1, \alpha_2 \in R^{(1)}$ имеет место равенство

$$\Phi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \Phi(x_1) + \alpha_2 \Phi(x_2).$$

Функционал $\Phi(x)$, равный отношению двух линейных функционалов $\Phi_1(x), \Phi_2(x)$ ($\Phi_2(x) \neq 0$), называют **дробно-линейным**.

Функционал $\Phi(x)$, $x \in X$ называют **выпуклым**, если для любых x_1 и $x_2 \in X$ и $0 \leq \alpha \leq 1$ справедливо неравенство

$$\Phi(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha \Phi(x_1) + (1 - \alpha)\Phi(x_2).$$

1. БУЛЕВЫ ФУНКЦИИ

1.1. СПОСОБЫ ЗАДАНИЯ
БУЛЕВЫХ ФУНКЦИЙ

Определение булевой функции. Развитию аппарата булевых функций способствовало широкое применение его в решении прикладных задач. Одной из первых и наиболее простой из таких задач является построение релейно-контактных схем с заданными свойствами. Как известно, реле имеет два устойчивых состояния, поэтому поведение релейных схем хорошо описывается на языке булевых функций. В дальнейшем аппарат булевых функций стал широко применяться при изучении поведения широкого класса дискретных управляющих систем.

Функцию $f(x_1, \dots, x_n)$ называют **булевой (переключательной, функцией алгебры логики)**, если она и ее переменные принимают только два значения. В качестве этих двух значений в математической логике рассматривают значения «истина» и «ложь» или 1 и 0.

Из определения булевой функции от n переменных следует, что областью определения булевой функции служит множество всех наборов из 0 и 1 длины n . Для задания функции достаточно указать значения функции на каждом наборе, т. е. выписать таблицу (табл. 1).

Известно, что число различных наборов из 0 и 1 длины n равно 2^n . Следовательно, число булевых функций, зависящих от n переменных, равно 2^{2^n} .

Булевы функции одной и двух переменных. Все 16 функций двух переменных приведены в табл. 2.

Функции $f_0(x, y)$ и $f_{15}(x, y)$ не зависят ни от одной переменной; их

называют константами и обозначают соответственно 0 и 1.

Функции $f_3(x, y)$, $f_5(x, y)$, $f_{10}(x, y)$ и $f_{12}(x, y)$ зависят от одной переменной ($f_3(x, y)$ и $f_{12}(x, y)$ — от x , а $f_5(x, y)$ и $f_{10}(x, y)$ — от y). Значения функций $f_3(x, y)$ и $f_5(x, y)$ совпадают со значениями переменных, от которых они зависят. Их обозначают соответственно x и y . Функции $f_{10}(x, y)$ и $f_{12}(x, y)$ принимают значения, противоположные значениям переменных, от которых они зависят. Их называют *отрицанием* или *инверсией* и обозначают соответственно \bar{y} и \bar{x} . Остальные функции зависят от двух переменных; для них приняты следующие названия и обозначения:

$f_2(x, y)$ — *конъюнкция* (или логическое умножение), $x \wedge y$. Часто знак « \wedge » опускают или вместо него употребляют знак « \cdot »;

$f_7(x, y)$ — *дизъюнкция*, $x \vee y$;

$f_9(x, y)$ — *эквивалентность*, $x \sim y$;

$f_{13}(x, y)$ и $f_{11}(x, y)$ — *импликация*, соответственно $x \rightarrow y$ и $y \rightarrow x$;

$f_8(x, y)$ — *сумма по модулю 2*, $x \oplus y$ или $x \dot{+} y \pmod{2}$;

$f_{14}(x, y)$ — *штрих Шеффера*, $x | y$;

$f_6(x, y)$ — *стрелка Пирса*, $x \downarrow y$;

1. Таблица значений
булевой функции

x_1	x_2	\dots	x_{n-1}	x_n	$f(x_1, x_2, \dots, x_{n-1}, x_n)$
0	0	\dots	0	0	$f(0, 0, \dots, 0, 0)$
0	0	\dots	0	1	$f(0, 0, \dots, 0, 1)$
1	1	\dots	1	0	$f(1, 1, \dots, 1, 0)$
1	1	\dots	1	1	$f(1, 1, \dots, 1, 1)$

2. Значения булевых функций двух переменных

x	y	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

$f_2(x, y)$ и $f_4(x, y)$ — функции запрета соответственно по y и x , $x \Delta y$ и $y \Delta x$.

Формулы. До сих пор булевы функции задавались таблицами значений, рассмотрим теперь задание булевых функций формулами.

Если дано некоторое множество B булевых функций, называемое *базисом*, то формулой (над базисом B) является:

а) обозначение произвольной функции из B ;

б) выражение $g(F_1, \dots, F_m)$, где F_1, \dots, F_m — формулы или переменные, а $g(x_1, \dots, x_m)$ — обозначение функции из B .

Например, выражение $\overline{(x \rightarrow y)} \wedge (\bar{y} \vee z)$ является формулой над базисом $\{x \rightarrow y, x \wedge y, x \vee y, \bar{x}\}$.

Все формулы, которые порождаются в процессе построения данной формулы, называют ее *подформулами*. В рассмотренном выше примере подформулами являются выражения

$$x \rightarrow y, \overline{x \rightarrow y}, \bar{y}, \bar{y} \vee z,$$

$$\overline{(x \rightarrow y)} \wedge (\bar{y} \vee z).$$

3. Значения формулы $\overline{(x \rightarrow y)} \wedge (\bar{y} \vee z)$

x	y	z	$\overline{(x \rightarrow y)} \wedge (\bar{y} \vee z)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

Каждой формуле соответствует функция, про которую говорят, что она реализуется данной формулой. Зная таблицы функций из B , по формуле можно построить таблицу функции, реализуемой ею.

Пример 1. Построим таблицу функции, реализуемой формулой $\overline{(x \rightarrow y)} \wedge (\bar{y} \vee z)$. Используя табл. 2, вычислим значение формулы на каждом наборе значений переменных x, y, z . Прделаем это для набора (110):

$$\begin{aligned} \overline{(1 \rightarrow 1)} \wedge (\bar{1} \vee 0) &= \bar{1} \wedge (0 \vee 0) = \\ &= 0 \wedge 0 = 0. \end{aligned}$$

Поступая аналогично для остальных наборов, получим табл. 3 этой функции.

Равносильные формулы. Формулы F_1 и F_2 называют *равносильными* ($F_1 = F_2$), если они на любом наборе значений переменных принимают одинаковые значения.

Пример 2.
Законны де Моргана

$$\overline{A \vee B} = \bar{A} \wedge \bar{B}, \quad \overline{A \wedge B} = \bar{A} \vee \bar{B};$$

законны поглощения

$$\begin{aligned} A \vee (A \wedge B) &= \\ &= A \wedge (A \vee B) = A; \end{aligned}$$

законны идемпотентности

$$A \wedge A = A \vee A = A;$$

закон двойного отрицания

$$\overline{\bar{A}} = A.$$

Здесь A и B — произвольные формулы.

Замена в формуле подформулы на равносильную последней приводит к формуле равносильной исходной.

Булева алгебра. Множество всех булевых функций вместе с операциями конъюнкция, дизъюнкция и отрицание образуют *булеву алгебру*.

Из определения конъюнкции, дизъюнкции и отрицания (см. табл. 2), следует справедливость аксиом 1—5 булевой алгебры:

1) *законы коммутативности:*

$$а) x \wedge y = y \wedge x; б) x \vee y = y \vee x;$$

2) *законы ассоциативности:*

$$а) x \wedge (y \wedge z) = (x \wedge y) \wedge z;$$

$$б) x \vee (y \vee z) = (x \vee y) \vee z;$$

3) *законы дистрибутивности:*

$$а) x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z);$$

$$б) x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z);$$

$$4) а) x \vee 0 = x; б) x \wedge 1 = x;$$

$$5) а) x \vee \bar{x} = 1; б) x \wedge \bar{x} = 0.$$

Любое соотношение равносильности можно получить с помощью этих аксиом. Например, закон идемпотентности можно доказать следующими преобразованиями:

$$\begin{aligned} x \vee x &= (4.б) = (x \vee x) \wedge 1 = (5.а) = \\ &= (x \vee x) \wedge (x \vee \bar{x}) = (3.б) = \\ &= x \vee (x \wedge \bar{x}) = (5.б) = x \vee 0 = \\ &= (4.а) = x. \end{aligned}$$

1.2. НОРМАЛЬНЫЕ ФОРМЫ

Далее будем рассматривать формулы в базисе $\{x \vee y, x \wedge y, \bar{x}\}$. Этот базис представляет интерес, так как он является *функционально полным*, т. е. любая булева функция может быть реализована формулой над этим базисом.

Элементарной конъюнкцией (дизъюнкцией) называют конъюнкцию (дизъюнкцию) переменных или их отрицаний, причем каждая переменная входит не более одного раза.

Дизъюнктивной (конъюнктивной) нормальной формой называют дизъюнкцию (конъюнкцию) различных элементарных конъюнкций (дизъюнкций).

Пример 3. Формулы $\bar{x}\bar{y} \vee \bar{x}y \vee xy$; $\bar{x} \vee xy$; $\bar{x}\bar{y} \vee y$ и $\bar{x} \vee y$ являются дизъюнктивными нормальными фор-

мами и реализуют функцию $x \rightarrow y$.

Совершенные нормальные формы. Если в каждом члене нормальной формы представлены все переменные, то ее называют *совершенной нормальной формой*.

Если функция $f(x_1, \dots, x_n)$ не является константой 0 (1), то по ее табличному заданию можно построить совершенную дизъюнктивную (конъюнктивную) нормальную форму, следуя правилам:

1) найти все наборы, на которых функция равна 1 (0);

2) для каждого найденного набора написать конъюнкцию (дизъюнкцию) переменных или их отрицаний следующим образом. Если i -я компонента набора равна нулю, то в конъюнкцию (дизъюнкцию) входит \bar{x}_i (x_i). Если i -я компонента набора равна единице, то в конъюнкцию (дизъюнкцию) входит x_i (\bar{x}_i);

3) полученные конъюнкции (дизъюнкции) объединить операцией дизъюнкции (конъюнкции).

Пример 4. Функции, заданной табл. 3, соответствуют совершенные нормальные формы:

$$\begin{aligned} &(x\bar{y}\bar{z}) \vee (x\bar{y}z); \\ &(x \vee y \vee z) \wedge (x \vee y \vee \bar{z}) \wedge \\ &\wedge (x \vee \bar{y} \vee z) \wedge (x \vee \bar{y} \vee \bar{z}) \wedge \\ &\wedge (\bar{z} \vee \bar{y} \vee z) \wedge (\bar{x} \vee \bar{y} \vee \bar{z}). \end{aligned}$$

Совершенные нормальные формы реализуют функцию, для которой они построены.

Минимальные нормальные формы. Как видно из примера 3, булеву функцию можно реализовать различными формулами над базисом $\{x \vee y, x \wedge y, \bar{x}\}$. Эти формулы отличаются друг от друга числом вхождений переменных и знаков операций. Поэтому естественно ставить задачу о реализации функции минимальной в некотором смысле формулой. Ограничимся рассмотрением реализации функции дизъюнктивными нормальными формами.

Минимальной дизъюнктивной нормальной формой функции называют дизъюнктивную нормальную форму (ДНФ), содержащую наименьшее число вхождений переменных среди всех ДНФ, реализующих эту функцию.

Существует много алгоритмов получения минимальной ДНФ. Рассмотрим алгоритм, который в качестве исходных данных использует совершенную ДНФ.

Элементарную конъюнкцию U называют простым импликантом функции $f(x_1, \dots, x_n)$, если $U \rightarrow f(x_1, \dots, x_n) = 1$, а при удалении из U любой переменной это условие для полученной таким образом конъюнкции не выполняется.

Все простые импликанты функции можно получить из совершенной ДНФ применением двух тождественных преобразований:

операции *неполного склеивания*

$$Ax \vee A\bar{x} = Ax \vee A\bar{x} \vee A,$$

операции *поглощения*

$$A \vee AB = A.$$

Здесь A и B — произвольные конъюнкции. Операции применяются последовательно, сначала — первая столько раз, сколько это возможно, затем — вторая.

Пример 5. Получим все простые импликанты функции $x \rightarrow y$ заданной совершенной ДНФ:

$$\begin{aligned} x \rightarrow y &= \bar{x}\bar{y} \vee \bar{x}y \vee xy = \\ &= \bar{x}\bar{y} \vee \bar{x}y \vee \bar{x} \vee xy = \bar{x}\bar{y} \vee \bar{x}y \vee \bar{x} \vee \\ &\vee xy \vee y = \bar{x}y \vee \bar{x} \vee xy \vee y = \\ &= \bar{x} \vee xy \vee y = \bar{x} \vee y. \end{aligned}$$

ДНФ, состоящую из всех простых импликантов функции, называют *сокращенной ДНФ* этой функции.

Тупиковой ДНФ функции называют ДНФ, для которой выполнены следующие условия: она реализует функцию и состоит из простых импликантов функции; если из нее удалить произвольный простой импликант, то она перестает реализовать исходную функцию.

Всякая минимальная ДНФ является тупиковой, но не любая тупиковая ДНФ является минимальной.

Алгоритм получения всех тупиковых ДНФ функции из ее сокращенной ДНФ состоит в следующем:

каждому простому импликанту U_i сопоставляют булеву переменную u_i ; для каждого набора $\bar{a}^j = (\alpha_1^j, \dots, \alpha_n^j)$, такого, что $f(\bar{a}^j) = 1$, находят все простые импликанты U_{j_l} , такие, что $U_{j_l}(\bar{a}^j) = 1$, $1 \leq l \leq k$, и образуют дизъюнкцию соответствующих им переменных $u_{j_1} \vee \dots \vee u_{j_k}$;

образуют конъюнкцию элементарных дизъюнкций, построенных на предыдущем шаге;

пользуясь законом дистрибутивности и идемпотентности, получают из конъюнктивной нормальной формы предыдущего шага ДНФ;

находят сокращенную ДНФ функции, реализуемой ДНФ, полученной на предыдущем шаге;

каждой конъюнкции $u_{i_1} \dots u_{i_r}$ сокращенной ДНФ, полученной на предыдущем шаге, сопоставляют ДНФ $U_{i_1} \vee \dots \vee U_{i_r}$. Полученные ДНФ являются тупиковыми ДНФ исходной функции, и других тупиковых ДНФ у нее нет.

Пример 6. Используя алгоритм, можно проверить, что тупиковыми ДНФ функции, реализуемой сокращенной ДНФ

$$x_1x_2 \vee x_1\bar{x}_4 \vee \bar{x}_2\bar{x}_4 \vee x_2x_3x_4 \vee \bar{x}_1x_3x_4 \vee \bar{x}_1\bar{x}_2x_3,$$

являются две ДНФ:

$$\bar{x}_2\bar{x}_4 \vee x_1x_2 \vee \bar{x}_1x_3x_4;$$

$$\bar{x}_2\bar{x}_4 \vee x_1x_2 \vee x_2x_3x_4 \vee \bar{x}_1\bar{x}_2x_3.$$

Сравнивая тупиковые ДНФ функции, получим минимальные ДНФ. Так, для функции из примера 6 существует единственная минимальная ДНФ. Она совпадает с первой тупиковой ДНФ.

С ростом числа переменных число тупиковых ДНФ почти для всех функций очень быстро растет, поэтому алгоритмы получения минимальных ДНФ, основанные на сравнении всех тупиковых ДНФ, применяют только для функций от небольшого числа переменных. При большом числе переменных применяют разнообразные приближенные методы, которые не гарантируют минимальности ДНФ, но дают хорошие практические результаты.

4. Логические связи

Название связи	Обозначение	Чтение
Конъюнкция Дизъюнкция	\wedge (&, \cdot) \vee	«и» «или» «...или или оба»
Импликация	\rightarrow (\supset)	«если то ...», «влечет». «следует»
Эквивалентность	\sim ($=$, \leftrightarrow , \equiv , \Leftrightarrow)	«...тогда и только тогда, когда», «эквивалентно»
Отрицание	\neg ($\bar{}$, \sim)	«не», «неверно, что...»

2. ИСЧИСЛЕНИЕ
ВЫСКАЗЫВАНИЙ

2.1. ЛОГИКА ВЫСКАЗЫВАНИЙ

Высказывания. Первичным понятием исчисления высказываний является высказывание, т. е. утверждение, о котором можно сказать, что оно истинно или ложно, но не то и другое вместе. Например, «Два больше трех». « $\sin 90^\circ = 1$ ». Первое высказывание ложно, а второе истинно. Высказывания могут иметь различный смысл, но в исчислении высказываний не интересуются вопросом о смысле высказываний, а интересуются только их свойством быть истинными или ложными. «Истину» или «ложь», приписанную высказыванию, называют истинностным значением этого высказывания и обозначают соответственно 1 или 0 (И или Л, Т или F, TRUE или FALSE).

Логические связи. В естественном языке сложные предложения образуются из простых с помощью связок (союзов, частиц и т. п.). Например, если обозначить через P высказывание «8 делится на 2», а через Q высказывание «8 делится на 4», то следующие высказывания можно рассматривать как составные. P и $Q =$ «8 делится на 2 и 8 делится на 4». Если Q , то $P =$ «Если 8 делится на 4, то 8 делится на 2». Не $P =$ «8 не делится на 2». Эти связи можно рассматривать как операции над высказываниями. В исчислении высказы-

ваний обычно используют пять связок, которые называют *логическими* (или *пропозициональными*) связками. Они приведены в табл. 4.

Формула исчисления высказываний. Высказывания обозначают большими буквами латинского алфавита (иногда с индексами): $P, Q, R, \dots, P_1, P_2, \dots$. Символы, которые используют для обозначения высказываний, называют *элементарными* (атомарными) *формулами* или *атомами*.

Формулой исчисления высказываний является:

- элементарная формула;
- выражения $(A \wedge B)$, $(A \vee B)$, $(A \rightarrow B)$, $(A \sim B)$ и \bar{A} , где A и B — формулы исчисления высказываний.

Можно обойтись без некоторых скобок, если считать, что при отсутствии скобок сначала выполняется операция конъюнкции, потом дизъюнкции, затем импликации и, наконец, эквивалентности.

Интерпретация. Интерпретацией I формулы A называют приписывание истинностных значений всем атомам, входящим в формулу A .

Чтобы получить истинностное значение формулы A в данной интерпретации, надо указать правила изменения истинностных значений после применения логических связок. Эти правила приведены в табл. 2.

Пример 7. Рассмотрим все интерпретации (табл. 5) формулы $((P \vee Q) \wedge \bar{R}) \rightarrow (\bar{P} \rightarrow Q)$.

5. Интерпретации формулы $((P \vee Q) \wedge \bar{R}) \rightarrow (\bar{P} \rightarrow Q)$

P	Q	R	$P \vee Q$	\bar{R}	$(P \vee Q) \wedge \bar{R}$	$\bar{P} \rightarrow Q$	$(P \vee Q) \wedge \bar{R} \rightarrow (\bar{P} \rightarrow Q)$
0	0	0	0	1	0	0	1
0	0	1	0	0	0	0	1
0	1	0	1	1	1	0	0
0	1	1	1	0	0	0	1
1	0	0	1	1	1	1	1
1	0	1	1	0	0	1	1
1	1	0	1	1	1	0	0
1	1	1	1	0	0	0	1

Построенную таблицу называют *истинностной таблицей формулы*. Говорят, что *формула истинна в данной интерпретации* (или интерпретация является *моделью формулы*), если ее истинностное значение равно единице в этой интерпретации, в противном случае говорят, что *формула ложна в данной интерпретации*.

Общезначимость и противоречивость. В зависимости от истинностных значений формул в интерпретациях выделяют следующие классы формул.

Формулу называют *общезначимой* (или *тавтологией*) тогда и только тогда, когда она истинна во всех интерпретациях, в противном случае ее называют *необщезначимой*.

Формулу называют *противоречивой* (или *невыполнимой*) тогда и только тогда, когда она ложна во всех интерпретациях, в противном случае ее называют *непротиворечивой* (или *выполнимой*).

Из истинностной табл. 5 следует, что формула $((P \vee Q) \wedge \bar{R}) \rightarrow (\bar{P} \rightarrow Q)$ выполнима. Построив соответствующие истинностные таблицы, мож-

но убедиться в том, что формула $P \vee \bar{P}$ общезначима, а формула $P \wedge \bar{P}$ противоречива.

Аналогично тому, как это было сделано для формул над базисом, в исчислении высказываний можно ввести понятие равносильных формул.

Все результаты, полученные для формул над базисом теории булевых функций, имеют место и в исчислении высказываний.

Логические следствия. В приложениях логики важную роль играет понятие логического следствия.

Если A_1, \dots, A_n, B — произвольные формулы, то говорят, что формула B *логически следует из формул* A_1, \dots, A_n ($A_1, \dots, A_n \models B$), если в любой интерпретации, в которой истинны все формулы A_1, \dots, A_n , истинна также формула B .

Понятие логического следствия из конечного числа формул сводится к понятию общезначимости. А именно: для любых формул A_1, \dots, A_n, B формула B логически следует из A_1, \dots, A_n тогда и только тогда, когда формула $A_1 \wedge \dots \wedge A_n \rightarrow B$ общезначима.

Пример 8. Покажем, что $P \rightarrow Q, \bar{Q} \models \bar{P}$ (табл. 6).

2.2. АКСИОМАТИЧЕСКАЯ ТЕОРИЯ ИСЧИСЛЕНИЯ ВЫСКАЗЫВАНИЙ

Рассмотрим построение исчисления высказываний как аксиоматической теории. Начальным шагом построения любой аксиоматической теории является описание языка этой теории.

6. Интерпретации формул $P \rightarrow Q, \bar{Q}, \bar{P}$

P	Q	$P \rightarrow Q$	\bar{Q}	\bar{P}
0	0	1	1	1
0	1	1	0	1
1	0	0	1	0
1	1	1	0	0

Язык теории включает в себя алфавит и понятие слова, осмысленного в этом языке. В исчислении высказываний алфавит состоит из символов, обозначающих атомы, логических связей и скобок, а осмысленное слово — из формул исчисления высказываний. Следующим шагом в описании аксиоматической теории является выделение класса формул, выводимых в этой теории. Определение выводимых формул так же, как и определение формулы, носит рекурсивный характер. Сначала определяют исходные выводимые формулы — *аксиомы*, а затем определяют *правила вывода*, позволяющие из имеющихся выводимых формул образовывать новые выводимые формулы.

Аксиомы исчисления высказываний:

1. $P \rightarrow (Q \rightarrow P)$;
2. $(P \rightarrow (Q \rightarrow R)) \rightarrow ((P \rightarrow Q) \rightarrow (P \rightarrow R))$;
3. $P \wedge Q \rightarrow P$;
4. $P \wedge Q \rightarrow Q$;
5. $(P \rightarrow Q) \rightarrow ((P \rightarrow R) \rightarrow (P \rightarrow Q \wedge R))$;
6. $P \rightarrow P \vee Q$;
7. $Q \rightarrow P \vee Q$;
8. $(P \rightarrow R) \rightarrow ((Q \rightarrow R) \rightarrow (P \vee Q \rightarrow R))$;
9. $(P \sim Q) \rightarrow (P \rightarrow Q)$;
10. $(P \sim Q) \rightarrow (Q \rightarrow P)$;
11. $(P \rightarrow Q) \rightarrow ((Q \rightarrow P) \rightarrow (P \sim Q))$;
12. $(P \rightarrow Q) \rightarrow (\bar{Q} \rightarrow \bar{P})$;
13. $\bar{\bar{P}} \rightarrow P$;
14. $P \rightarrow \bar{\bar{P}}$.

Правила вывода:

1. *Правило подстановки*. Если A — формула исчисления высказываний, содержащая атом P , то говорят, что любая формула, полученная из формулы A заменой атома P всюду на произвольную формулу B , непосредственно выводима из формулы A .

2. *Правило заключения (модус поненс)*. Если $A \rightarrow B$ и A — формулы исчисления высказываний, то говорят, что формула B непосредственно выводима из формул $A \rightarrow B$ и A . Формулы $A \rightarrow B$ и A называют посылками, а формулу B — заключением.

Доказательством (или выводом) формулы B из формул A_1, \dots, A_n называют последовательность B_1, \dots, B_l , в которой каждая формула B_i есть либо одна из формул A_j , либо аксиома, либо непосредственно выводима из последовательности B_1, \dots, B_{i-1} и B_l есть B . Формулы A_1, \dots, A_n называют гипотезами. Выводимость формулы B из гипотез A_1, \dots, A_n обозначают через $A_1, \dots, A_n \vdash B$. Если в доказательстве формулы B не используют гипотезы (а только аксиомы), то эту формулу называют теоремой и обозначают через $\vdash B$.

Пример 9. Докажем, что $\vdash P \sim \bar{\bar{P}}$.

1. $P \rightarrow P$ — по аксиоме 14;
2. $\bar{\bar{P}} \rightarrow P$ — по аксиоме 13;
3. $(P \rightarrow Q) \rightarrow ((Q \rightarrow P) \rightarrow (P \sim Q))$ — по аксиоме 11;
4. $(P \rightarrow \bar{\bar{P}}) \rightarrow ((\bar{\bar{P}} \rightarrow P) \rightarrow (P \sim \bar{\bar{P}}))$ — из п. 3 по правилу подстановки;
5. $(\bar{\bar{P}} \rightarrow P) \rightarrow (P \sim \bar{\bar{P}})$ — из пп. 4 и 1 по правилу заключения;
6. $P \sim \bar{\bar{P}}$ из пп. 5 и 2 по правилу заключения.

Отметим некоторые свойства построенной аксиоматической теории исчисления высказываний.

Построенная аксиоматическая теория *адекватна (или полна)* построенному исчислению высказываний (см. с. 25) в том смысле, что всякая общезначимая формула является теоремой и, наоборот, всякая теорема является общезначимой формулой.

Построенная теория *просто непротиворечива* (в ней не существует такой формулы, что одновременно она и ее отрицание являются теоремами) и *абсолютно непротиворечива* (в ней суще-

7. Операции квантирования

Название операции	Обозначение	Чтение
Квантор общности (всеобщности)	$\forall ((x), \wedge x, \Pi x)$	Для всех
Квантор существования	$\exists (E, \vee x, \Sigma x)$	Существует

стует хотя бы одна формула, не являющаяся теоремой). В исчислении высказываний понятия абсолютной и простой непротиворечивости совпадают.

Аксиомы 1—14 являются *независимыми*, т. е. никакая из аксиом не может быть доказана из остальных.

Аксиомы 1—14 *полны (по Посту)* в том смысле, что добавление к ним любой формулы, не являющейся теоремой, приводит к нарушению непротиворечивости.

3. ИСЧИСЛЕНИЕ ПРЕДИКАТОВ

3.1. ЛОГИКА ПРЕДИКАТОВ

Предикаты. В исчислении высказываний исходными элементами для построения языка являлись высказывания, т. е. предложения, которые истинны или ложны. Однако многие математические утверждения не могут быть формализованы с использованием таких простых исходных элементов, например, следующее умозаключение: «*Всякое рациональное число является действительным, 4 есть рациональное число. Следовательно, 4 есть действительное число.*»

Здесь последнее высказывание не является следствием из первых двух, если эти высказывания рассматривать в рамках исчисления высказываний, т. е. не рассматривать структуру этих высказываний. Для того чтобы формализовать это рассуждение, нужно ввести два новых понятия (предикат и квантор) по сравнению с исчислением высказываний. Раньше чем это сделать, продолжим анализ приведенного выше рассуждения. Введем следующие обозначения:

$P(x)$ — « x есть рациональное число»,
 $Q(x)$ — « x есть действительное число».

Тогда, используя логические связи и введенные обозначения, можно переписать рассуждение в виде:

(Для всякого x , такого, что $(P(x) \rightarrow Q(x)) \wedge P(4) \rightarrow Q(4)$). Теперь становится ясным правильность умозаключения. В самом деле, если $P(x) \rightarrow Q(x)$ истинно для любых значений x , следовательно, и при $x = 4$ $P(4) \rightarrow Q(4)$ истинно. Далее из определения логического следствия получаем, что $Q(4)$ истинно.

Под *предикатом от n переменных* (или *n -местным предикатом*), $n \geq 0$, понимают предложение $P(x_1, \dots, x_n)$, которое становится высказыванием после подстановки вместо переменных x_1, x_2, \dots, x_n их значений из множеств соответственно M_1, M_2, \dots, M_n .

Таким образом, $P(x)$ и $Q(x)$ являются предикатами от одной переменной.

Полагаем, что при $n = 0$ предикат является высказыванием.

В дальнейшем для упрощения изложения будем считать, что $M_1 = M_2 = \dots = M_n = M$.

Кванторы. К предикатам можно применять логические связи $\wedge, \vee, \rightarrow, \sim, -$ и таким образом получать составные предикаты. Например, если к предикатам $P(x, y) = "x = y"$ и $Q(x, y) = "x < y"$, где x и y принимают значения из множества действительных чисел, применить связи \wedge и $-$, то получим предикат $R(x, y) = \overline{P(x, y)} \wedge \overline{Q(x, y)} = "x > y"$.

Помимо логических связей, в исчислении предикатов применяют две особые операции — операции квантирования (табл. 7).

Теперь рассуждение, умозаключение, рассмотренное выше (см. с. 28), будет иметь вид

$$(\forall x (P(x) \rightarrow Q(x))) \wedge P(4) \rightarrow Q(4).$$

Формула исчисления предикатов. Для построения языка исчисления предикатов будем использовать следующие типы символов:

индивидуальные символы (предметы, константы), т. е. элементы множества M . Будем обозначать их малыми буквами латинского алфавита (иногда с индексами):

$$a, b, c, \dots, a_1, a_2,$$

Символы предметных переменных, т. е. переменных, от которых зависит предикат. Будем обозначать их малыми буквами латинского алфавита (иногда с индексами):

$$x, y, z, \dots, x_1, x_2,$$

Предикатные символы. Будем обозначать их большими буквами латинского алфавита (можно с индексами):

$$P, Q, R, \dots, P_1, P_2, \dots, P_i^n,$$

Верхний индекс обозначает местность предикатного символа.

Если символ P есть n -местный предикатный символ и t_1, t_2, \dots, t_n — предметные переменные или предметы, то выражение $P(t_1, t_2, \dots, t_n)$ является *элементарной (атомарной) формулой* или *атомом* исчисления предикатов.

Формулой исчисления предикатов является:

элементарная формула исчисления предикатов;

выражения $(A \vee B)$, $(A \wedge B)$, $(A \rightarrow B)$, $(A \sim B)$ и \bar{A} , если A и B — формулы исчисления предикатов;

выражения $\forall x A$ и $\exists x A$, если A — формула и x предметная переменная.

Сохраним введенные ранее соглашения о порядке выполнения логических связок при отсутствии скобок.

Свободные и связанные предметные переменные. Из определения формулы исчисления предикатов следует, что предметные переменные, входящие в формулу, не равноправны, так как по отношению к некоторым из них

применяются операции квантирования, а по отношению к другим нет.

Определим *область действия квантора*, входящего в формулу, как ту подформулу, к которой этот квантор применяется.

Вхождение предметной переменной x в формулу A называют *связанным*, если она входит в выражение $\forall x$ или $\exists x$ или попадает в область действия таких выражений. Вхождение называют *свободным*, если оно не является связанным.

Рассмотрим формулу $(\forall x (P(x) \rightarrow \rightarrow Q(x))) \wedge R(x, y)$. В ней первое вхождение переменной x связано, так как входит в выражение $\forall x$, второе и третье вхождения связаны, так как попадают в область действия квантора $\forall x$, четвертое вхождение является свободным. Вхождение переменной y в эту формулу свободно.

Проиллюстрируем различие свободных и связанных вхождений следующим примером. Рассмотрим предикат $P(x, y) = "x = 5y"$, здесь x и y принимают целочисленные значения. Истинность и ложность его зависят от выбора значений обеих переменных x и y . Это отражает тот факт, что x и y входят в $P(x, y)$ свободно. Теперь рассмотрим предикат $\exists y P(x, y)$, истинность или ложность его зависит от x , но не зависит от выбора y . Можно сформулировать это утверждение в форме, в которой переменная y даже не встречается: « x делится на 5». Это отражает тот факт, что x входит свободно, а y — связано.

Интерпретация. В исчислении высказываний интерпретацией являлось приписывание истинностных значений атомам. В исчислении предикатов прежде, чем приписывать значения атомам, надо указать область определения предметных переменных, значения констант и предикатных символов.

Интерпретация I формулы A включает в себя следующее:

непустое множество M (область интерпретации);

соответствие $a \rightarrow I(a)$, которое каждой константе сопоставляет элемент из M ;

8. Истинностные значения формулы

$$\overline{P(x)} \rightarrow \exists y (P(y) \wedge R)$$

P	R	x	$\overline{P(x)} \rightarrow \exists y (P(y) \wedge R)$
f_1	0	a	0
f_1	0	b	0
f_1	1	a	0
f_1	1	b	0
f_2	0	a	0
f_2	0	b	1
f_2	1	a	1
f_2	1	b	1
f_3	0	a	1
f_3	0	b	0
f_3	1	a	1
f_3	1	b	1
f_4	0	a	1
f_4	0	b	1
f_4	1	a	1
f_4	1	b	1

соответствие $P \rightarrow I(P)$, которое каждому n -местному предикатному символу сопоставляет отображение M^n в $\{0, 1\}$;

соответствие $x \rightarrow I(x)$, которое каждой свободной предметной переменной сопоставляет элемент из M .

Чтобы получить истинностное значение формулы A в данной интерпретации, надо указать правила изменения истинностных значений после применения логических связок и операций квантирования. Для логических связок эти правила приведены в табл. 2. Для операций квантирования правила выглядят следующим образом:

$I(\exists x A(x)) = 1$ тогда и только тогда, когда существует такой элемент a из M , что $I(A(a)) = 1$, $I(\forall x A(x)) = 1$ тогда и только тогда, когда для всех элементов a из M $I(A(a)) = 1$.

Пример 10. Рассмотрим интерпретации формулы $\overline{P(x)} \rightarrow \exists y (P(y) \wedge R)$ для фиксированного множества $M = \{a, b\}$ (табл. 8). Здесь f_1, f_2, f_3, f_4 — всевозможные отображения множества $\{a, b\}$ в $\{0, 1\}$, которые являются значениями предикатного символа P . Они перечислены в следующей таблице:

M	P			
	f_1	f_2	f_3	f_4
a	0	0	1	1
b	0	1	0	1

Построенную таблицу называют *истинностной таблицей формулы*. Ясно, что если M — конечное множество, то для любой формулы можно построить истинностную таблицу.

Точно так же, как в исчислении высказываний, можно выделить три класса формул: общезначимые, выполнимые и противоречия.

Из табл. 8 следует, что формула $\overline{P(x)} \rightarrow \exists y (P(y) \wedge R)$ является выполнимой, но не общезначимой.

Формула $\forall x P(x) \sim \exists x \overline{P(x)}$ общезначима. Пусть существует интерпретация I , в которой формула ложна. Тогда, согласно определению связки \sim , выполняется одно из двух условий:

либо $I(\forall x P(x)) = 0$, а $I(\exists x \overline{P(x)}) = 1$, либо $I(\forall x P(x)) = 1$, а $I(\exists x \overline{P(x)}) = 0$. В первом случае $I(\forall x P(x)) = 0$ означает, что существует a , такое, что $I(P(a)) = 0$, т. е. $I(\overline{P(a)}) = 1$. Отсюда $I(\exists x \overline{P(x)}) = 1$, т. е.

$I(\exists x \overline{P(x)}) = 0$. Аналогичными рассуждениями можно показать, что во втором случае также получается противоречие, что и доказывает общезначимость формулы.

Аналогично можно показать, что формула $\forall x P(x) \wedge \exists y \overline{P(y)}$ противоречива.

Точно так же, как в исчислении высказываний, в исчислении предикатов можно ввести понятие равносильных формул и использовать замену формулы на равносильную ей.

Пусть A и B — произвольные формулы исчисления предикатов, тогда справедливы следующие равносильности:

- 1) а) $\overline{\forall x A} = \exists x \overline{A}$; б) $\overline{\forall x \overline{A}} = \exists x A$;
- 2) а) $\exists x (A \vee B) = \exists x A \vee \exists x B$;
- б) $\forall x (A \wedge B) = \forall x A \wedge \forall x B$;

3) если предметная переменная x не входит в формулу B свободно, то $(\Box x A) * B = \Box x (A * B)$. Здесь \Box — произвольный квантор, а $*$ — символ конъюнкции или дизъюнкции;

4) если предметная переменная x входит в формулу A , а переменная y не входит в A , то $(\Box x A(x)) = (\Box y A(y))$.

В исчислении предикатов можно ввести понятие логического следствия так же, как это было сделано в исчислении высказываний.

3.2. АКСИМАТИЧЕСКАЯ ТЕОРИЯ ИСЧИСЛЕНИЯ ПРЕДИКАТОВ

Для построения аксиоматической теории исчисления предикатов будем использовать понятие формулы исчисления предикатов, введенное ранее, аксиомы 1—14 и правила вывода исчисления высказываний, добавив к ним правила вывода, которые предоставляют возможность работать с операциями квантирования, и аксиомы:

$$15. \forall x P(x) \rightarrow P(y),$$

$$16. P(y) \rightarrow \exists x P(x).$$

Здесь переменная y не входит в формулу $P(x)$.

К правилам вывода исчисления предикатов, помимо уже известных правил заключения и подстановки, добавляются следующие:

\forall -правило. Если $B \rightarrow A(x)$ — формула исчисления предикатов, то говорят, что формула $B \rightarrow \forall x A(x)$ непосредственно выводима из формулы $B \rightarrow A(x)$.

\exists -правило. Если $A(x) \rightarrow B$ — формула исчисления предикатов, то говорят, что формула $\exists x A(x) \rightarrow B$ непосредственно выводима из формулы $A(x) \rightarrow B$. В этих правилах переменная x не входит свободно в формулу B .

Замечание. Отметим особенности правила подстановки в исчислении предикатов: а) если формулу B подставляют вместо n -местного предиката $P(x_1, \dots, x_n)$, то она должна содер-

жать n свободных переменных t_1, \dots, t_n , которые в результате подстановки заменяются на переменные соответственно x_1, \dots, x_n ; б) если предикат $P(x_1, \dots, x_n)$ находится в области действия квантора, связывающего какую-либо переменную, то эта переменная не входит в формулу B .

Понятия доказательства и теоремы вводятся так же, как в исчислении высказываний.

Пример 11. Покажем, что

$$\vdash \forall x P(x) \rightarrow \exists x P(x).$$

1. $P \rightarrow (Q \rightarrow P)$ по аксиоме 1.

2. $(P(y) \rightarrow \exists x P(x)) \rightarrow (Q \rightarrow (P(y) \rightarrow \exists x P(x)))$ — из п. 1 по правилу подстановки.

3. $P(y) \rightarrow \exists x P(x)$ по аксиоме 16.

4. $Q \rightarrow (P(y) \rightarrow \exists x P(x))$ из пп. 2, 3 по правилу заключения.

5. $\forall x P(x) \rightarrow (P(y) \rightarrow \exists x P(x))$ из п. 4 по правилу подстановки.

6. $(P \rightarrow (Q \rightarrow R)) \rightarrow ((P \rightarrow Q) \rightarrow (P \rightarrow R))$ по аксиоме 2.

7. $(\forall x P(x) \rightarrow (Q \rightarrow R)) \rightarrow ((\forall x P(x) \rightarrow Q) \rightarrow (\forall x P(x) \rightarrow R))$ — из п. 6 по правилу подстановки.

8. $(\forall x P(x) \rightarrow (P(y) \rightarrow R)) \rightarrow ((\forall x P(x) \rightarrow P(y)) \rightarrow (\forall x P(x) \rightarrow R))$ из п. 7 по правилу подстановки.

9. $(\forall x P(x) \rightarrow (P(y) \rightarrow \exists x P(x))) \rightarrow ((\forall x P(x) \rightarrow P(y)) \rightarrow (\forall x P(x) \rightarrow \exists x P(x)))$ из п. 8 по правилу подстановки.

10. $(\forall x P(x) \rightarrow P(y)) \rightarrow (\forall x P(x) \rightarrow \exists x P(x))$ — из пп. 5, 9 по правилу заключения.

11. $\forall x P(x) \rightarrow P(y)$ — по аксиоме 15.

12. $\forall x P(x) \rightarrow \exists x P(x)$ — из пп. 10, 11 по правилу заключения.

Исчисление предикатов так же, как и исчисление высказываний, является просто и абсолютно непротиворечивым. Исчисление предикатов полно в том смысле, что всякая теорема его общезначима и, наоборот, всякая общезначимая формула является теоремой. Вопрос о полноте (по Посту) для исчисления предикатов решается отрицательно.

1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ
И ПОНЯТИЯ

Графом G называют пару (X, E) , где $X \doteq X(G)$ — непустое конечное множество элементов, называемых *вершинами*, $E \doteq E(G)$ — конечное множество неупорядоченных и упорядоченных пар различных элементов из X , называемых соответственно ребрами и дугами. Вершины графа G обозначают символами $x_1, x_2, \dots, x_i, \dots, x_n, \dots$, а ребра и дуги — e_1, e_2, \dots, e_m ; $|X(G)| = n$ — число вершин графа G , $|E(G)| = m$ — число дуг и ребер графа G . Если $e_k = (x_i, x_j)$ — ребро ($(x_i, x_j) = (x_j, x_i)$), то говорят, что ребро e_k соединяет вершины x_i и x_j , и условно изображают диаграммой (рис. 1, а), в которой вершина изображается точкой или кружком, а ребро — отрезком линии, соединяющей точки или кружки, соответствующие вершинам ребра. Если $e_k = (x_i, x_j)$ — дуга ($(x_i, x_j) \neq (x_j, x_i)$), то говорят, что дуга e_k начинается в вершине x_i и заканчивается в x_j , и условно изображают дугу e_k ориентированным отрезком (рис. 1, б), направленным от начальной вершины x_i к конечной x_j .

Вершины, соединенные ребром или дугой, называют смежными. Ребра или дуги, имеющие общую вершину, также называют *смежными*. Ребро (дугу) и любую из его (ее) двух вершин называют *инцидентными*.

Граф $G = (X, E)$, содержащий только дуги, называют *ориентированным* (рис. 2, а), а содержащий только ребра — *неориентированным* (рис. 2, б).

Если в определении графа вместо множества E рассматривать конечное семейство \tilde{E} неупорядоченных или упорядоченных пар элементов из X (не обязательно различных), то пару

$(X, E) = \tilde{G}$ называют *псевдографом* (рис. 2, в). Ребра (рис. 3, а) или дуги (рис. 3, б) псевдографа с одинаковыми концевыми (начальными и конечными) вершинами называют *кратными* или *параллельными*.

Ребро $e_k = (x_i, x_i)$, соединяющее вершину x_i с ней самой, или дугу $e_k = (x_i, x_i)$, начинающуюся и заканчивающуюся в вершине x_i , называют *петлей* в вершине x_i . Если все ребра и дуги псевдографа \tilde{G} являются кратными, то псевдограф \tilde{G} называют *неориентированным мультиграфом*, если \tilde{G} содержит только ребра (рис. 4, а), и *ориентированным мультиграфом*, если только дуги (рис. 4, б). Если псевдограф \tilde{G} не содержит кратных ребер (дуг) и петель, то его называют *простым*. Простой псевдограф — граф, определенный выше.

Ориентированному графу $G = (X, E)$ можно сопоставить неориентированный граф с тем же числом вершин, заменив каждую дугу графа ребром, при этом пару противоположно направленных дуг $\{(x_i, x_j) \in E, (x_j, x_i) \in E\}$ — одним ребром. Получившийся *неориентированный* граф называют *неориентированным дубликатом* графа G (рис. 5, а, б) и обозначают $\bar{G} = (X, \bar{E})$.

Степени и полустепени вершин графа. Степенью $d(x_i)$ вершины x_i называют число ребер, инцидентных x_i ; *полустепенью захода* $d_-(x_i)$ вершины x_i называют число заходящих в x_i дуг, а *полустепенью исхода* $d_+(x_i)$ — число исходящих из x_i дуг.

Для ориентированного графа $G = (X, E)$, $|X| = n$, $|E| = m$ справедливо равенство

$$\sum_{i=1}^n d_+(x_i) = \sum_{i=1}^n d_-(x_i) = m.$$

Для неориентированного графа $G = (X, E)$, $|X| = n$, $|E| = m$ справедливо равенство (лемма «о рукопожатии»)

$$\sum_{i=1}^n d(x_i) = 2m.$$

Замечание. Согласно предыдущей формуле число вершин нечетной степени в любом неориентированном графе четно.

Вершину x , для которой $d(x) = 0$, называют *изолированной*, $d(x) = 1$ — *висячей*, $d_-(x) = 0$ — *исток*, $d_+(x) = 0$ — *сток*.

На рис. 5, б x_1 — изолированная, x_2, x_3, x_4 — висячие вершины; на рис. 5, а x_2 — исток, x_5 — сток.

Маршруты графа. *Ориентированным* (неориентированным) маршрутом μ графа называют последовательность дуг (ребер), в которой конечная вершина всякой дуги (ребра), отличной от последней, является начальной вершиной следующей.

Маршрут $\mu \cong \{e_1, e_2, \dots, e_k, \dots, e_t\}$, $t \in \mathbf{N}$ — *ориентированный*, если $e_k, k = 1, 2, \dots, t$ — дуга, и *неориентированный*, если $e_k, k = 1, 2, \dots, t$ — ребро, можно задать как последовательность вершин, инцидентных дугам (ребрам) маршрута μ , $\mu \cong \{x_{i_0}, x_{i_1}, x_{i_2}, \dots, x_{i_k}, x_{i_{k+1}}, \dots, x_{i_{t-1}}, x_{i_t}\} \cong \mu(x_{i_0} \rightarrow x_{i_t})$, в которой $(x_{i_{k-1}}, x_{i_k}) = e_k, k = 1, 2, \dots, t$. Вершина x_{i_0} — *начало маршрута* μ , x_{i_t} — *конец маршрута* μ ; число дуг (ребер) маршрута $|\mu| = t$ называют *длиной маршрута*.

Вершина x_j *достижима* из вершины x_i , если существует маршрут $\mu(x_i \rightarrow x_j)$. Вершина x_j *контрдостижима* из x_i , если существует маршрут $\mu(x_j \rightarrow x_i)$.

Множество всех вершин графа G , достижимых из x_i , $D(x_i) \cong \{x \in X(G), \exists \mu(x_i \rightarrow x)\}$. Множество $D^t(x_i) \cong \{x \in X(G), \exists \mu(x \rightarrow x_i)\}$ — множество всех *контрдостижимых* вершин из x_i .

Маршрут называют *открытым*, если его концевые вершины различны; в противном случае — *замкнутым*. Ориентированный (неориентированный) мар-

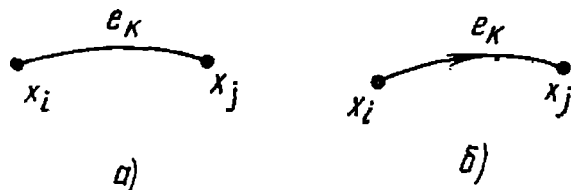


Рис. 1

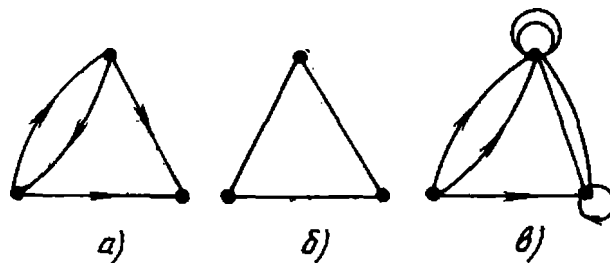


Рис. 2

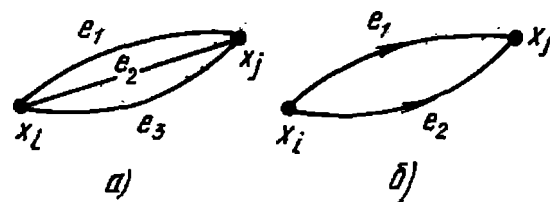


Рис. 3

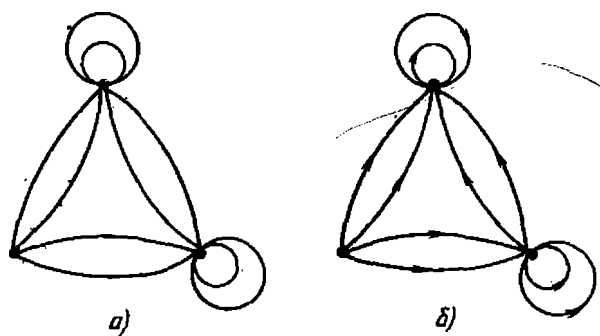


Рис. 4

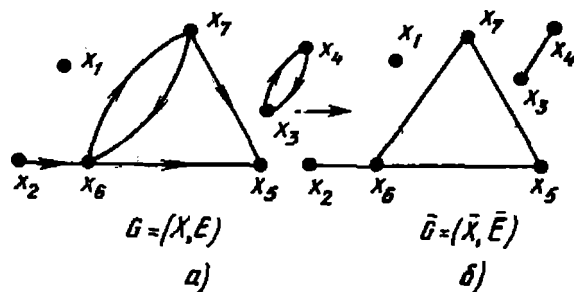


Рис. 5

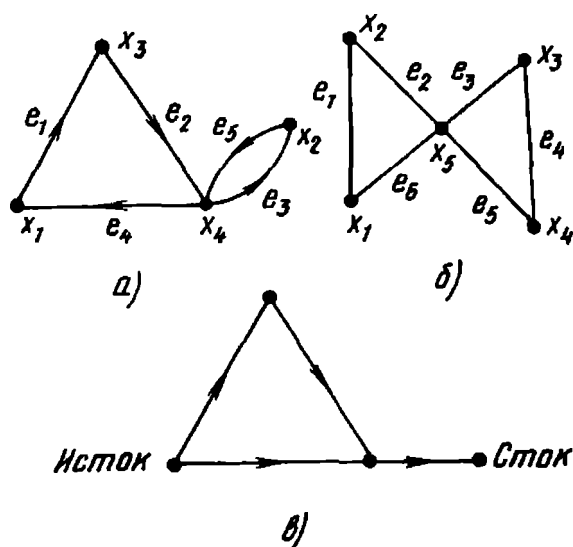


Рис. 6

шрут называют *ориентированной цепью* (*цепью*), если в нем все дуги (ребра) различны; и *путем* (*простой цепью*), если в нем все вершины различны. Замкнутую ориентированную цепь (*цепь*) называют *ориентированным циклом* (*циклом*), а замкнутый путь (простую цепь) — *контуром* (*простым циклом*).

Для ориентированного графа G (рис. 6, а) $\mu_1 = \{e_1, e_2, e_4, e_1, e_2, e_3\}$ — маршрут с началом в вершине x_1 и концом в вершине x_2 , т. е. $\mu_1 = \mu(x_1 \rightarrow x_2)$; $\mu_2 = \{e_1, e_2, e_3, e_5\} = \{x_1, x_3, x_4, x_2, x_4\}$ — ориентированная цепь, не являющаяся путем, а $\mu_3 = \{e_1, e_2, e_3\}$ — путь, $\mu_3 = \mu(x_1 \rightarrow x_2)$; $\{e_1, e_2, e_3, e_5; e_4\} = \mu_4$ — ориентированный цикл, состоящий из двух контуров $\mu_5 = \{e_3, e_5\}$ и $\mu_6 = \{e_1, e_2, e_4\}$.

Для неориентированного графа G (рис. 6, б) $\mu_1 = \{e_2, e_5, e_4, e_3\}$ — цепь, $\mu_2 = \{e_2, e_5, e_4\}$ — простая цепь, $\mu_3 = \{e_2, e_5, e_4, e_3, e_6, e_1\}$ — цикл, состоящий из двух простых циклов $\mu_4 = \{e_2, e_6, e_4\}$ и $\mu_5 = \{e_5, e_4, e_3\}$.

Граф, не содержащий контуров и простых циклов, называют *ациклическим*. В ациклическом ориентированном графе имеется по крайней мере одна вершина (исток) с нулевой степенью захода и одна вершина (сток) с нулевой полустепенью исхода (рис. 6, в).

В неориентированном графе G длину кратчайшей простой цепи, соединяющей вершины x_i и x_j , называют

расстоянием $r(x_i, x_j)$ между вершинами x_i и x_j , т. е. $r(x_i, x_j) = \min_{\mu(x_i \rightarrow x_j)} |\mu(x_i \rightarrow x_j)|$.

В неориентированном графе G , $n = |X(G)|$ *диаметром графа* d_G называют наибольшее расстояние, т. е. $d_G = \max_{1 \leq i, j \leq n} r(x_i, x_j)$; *радиусом графа* r_G называют величину $r_G = \min_{1 \leq i \leq n} \max_{1 \leq j \leq n} r(x_i, x_j)$; *центром графа* — вершину x_{i_0} , для которой $\max_{1 \leq j \leq n} r(x_{i_0}, x_j)$ принимает наименьшее значение.

Маршрут ориентированного графа, содержащий все его вершины, называют *остовным маршрутом*.

Части графа. Подграфом $G' = (X', E')$ графа $G(X, E)$ называют граф с множеством вершин $X' \subseteq X$ и множеством ребер (дуг) $E' \subseteq E$, каждое из которых инцидентно только вершинам из X' .

Подграф $G' = (X', E')$ графа $G = (X, E)$ называют *остовным*, если $X' = X$; *порожденным множеством вершин* $X' \subseteq X$, если множество E' состоит из всех ребер (дуг) графа G , которые соединяют (начинаются и заканчиваются) вершины из X' ; *порожденным множеством ребер (дуг) E'* , если G' не содержит изолированных вершин и множество его вершин X' является наименьшим подмножеством X , содержащим все концевые вершины из E' .

Подграф, порожденный множеством вершин X' , обозначают $\langle X' \rangle$; порожденный множеством ребер (дуг) E' — через $\langle E' \rangle$.

На рис. 7 изображены соответственно ориентированный граф G , его подграф G_1 , подграф G_2 графа G , порожденный множеством вершин $X'_2 = \{x_1, x_2, x_4\}$, подграф G_3 графа G , порожденный множеством дуг $E' = \{e_3, e_4, e_5\}$, и остовный подграф G_4 графа G .

Пусть P — некоторое свойство, которым могут обладать графы. *Максимальным (минимальным) подграфом графа G относительно свойства P* называют порожденный подграф $\langle X_* \rangle$ графа G , обладающий этим свойством и такой, что не существует другого

порожденного подграфа $\langle X' \rangle$, у которого $X' \supset X_*$ ($X' \subset X_*$) и который также обладает свойством P .

Типы связности графов. Неориентированный граф G называют *связным*, если в нем существует цепь между каждой парой вершин. В связном неориентированном графе расстояние $r(x_i, x_j)$ удовлетворяет аксиомам метрики. Ориентированный граф G называют:

а) *сильно связным* (рис. 8, а), если для любых двух вершин x_i и x_j существуют путь из x_i в x_j и обратно, т. е.

$$(\forall x_i \in X) (\forall x_j \in X)$$

$$(\exists \mu (x_i \leftarrow x_j) \wedge (\exists \mu (x_j \rightarrow x_i)));$$

б) *односторонне связным* (рис. 8, б), если для любых двух вершин x_i и x_j существует путь из x_i в x_j или обратно, т. е.

$$(\forall x_i \in X) (\forall x_j \in X) (\exists \mu (x_i \rightarrow x_j) \vee$$

$$\exists \mu (x_j \rightarrow x_i));$$

в) *слабо связным* (рис. 8, в), если связным является неориентированный дубликат \bar{G} (рис. 8, г) графа G . Ориентированный граф G — *сильно связен* тогда и только тогда, когда в нем существует замкнутый остоновый маршрут; *слабо связан* тогда и только тогда, когда в нем существует остоновый маршрут. Всякий сильно связный граф является в то же время односторонне связным, а следовательно, и слабо связным.

Компоненты связности графов. *Компонента связности неориентированного графа G* — максимальный связный подграф графа G , т. е. компонента связности не является подграфом любого другого связного подграфа графа G . *Сильно связной компонентой* (сильной компонентой) ориентированного графа G называют максимальный сильно

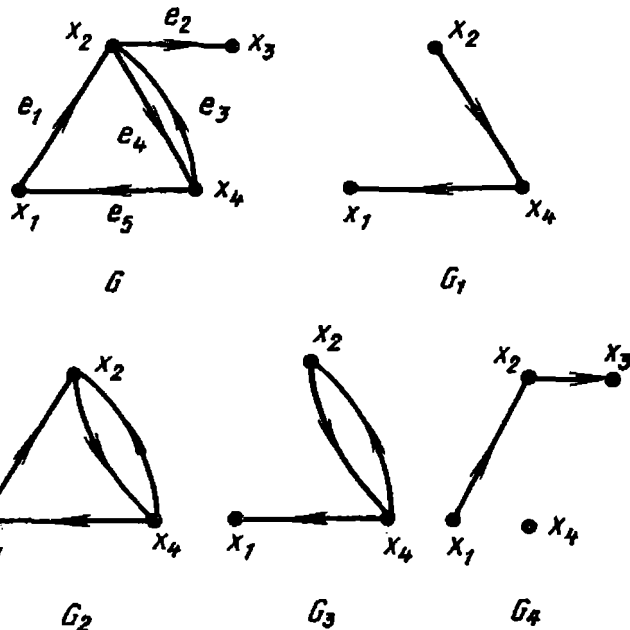


Рис. 7. Подграфы ориентированного графа

связный подграф графа G ; *слабо связной компонентой* (слабой компонентой) — максимальный слабо связный подграф графа G ; *односторонней компонентой* — максимальный односторонне связный подграф.

Компонента связности неориентированного графа — подграф, порожденный некоторым множеством вершин $X_i, i = 1, 2, \dots, p, p \in \mathbb{N}$ (p — число компонент связности G). Компонента сильной связности ориентированного графа G — подграф, порожденный некоторым множеством вершин $X_i^s, i = 1, 2, \dots, p_s, p_s \in \mathbb{N}$ (p_s — число компонент сильной связности G). Множества $\{X_i, i = 1, 2, \dots, p\}$ и $\{X_i^s, i = 1, 2, \dots, p_s\}$ образуют разбиение множества вершин соответствующих графов.

Из определений компонент связности следует, что односторонние компоненты графа G могут иметь общие

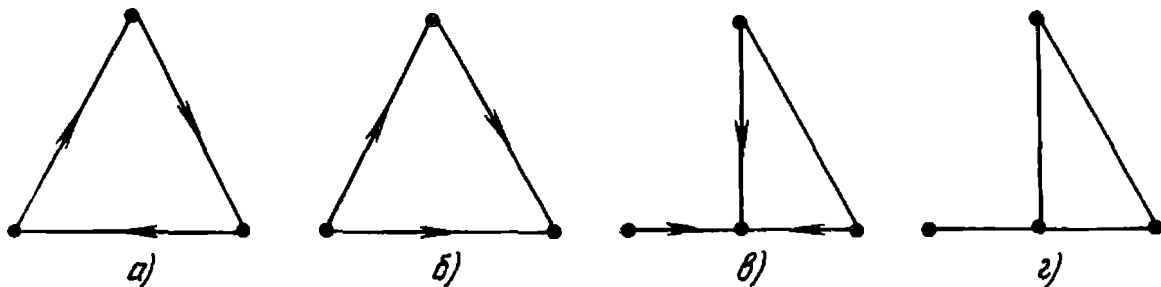


Рис. 8

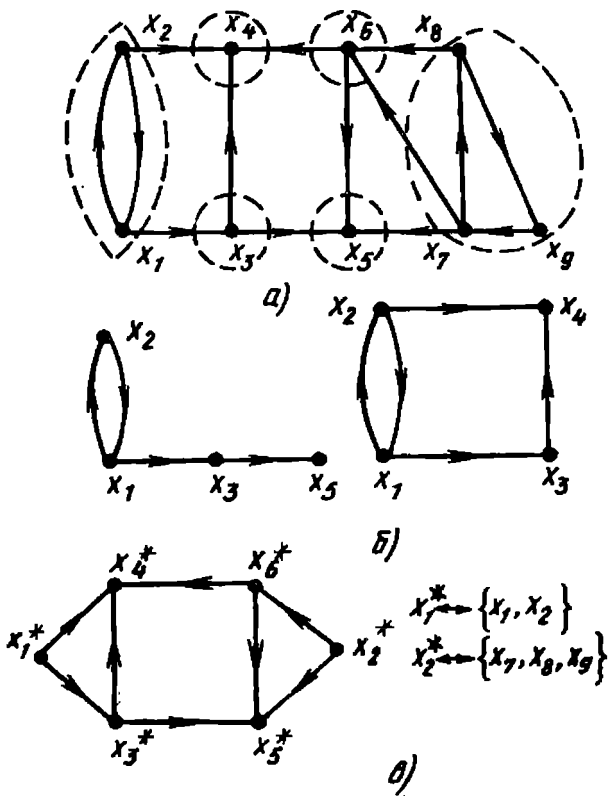


Рис. 9

вершины, сильная компонента содержится по крайней мере в одной одно-сторонней, а односторонняя компонента — в некоторой слабой компоненте данного графа G .

Для слабо связного графа G (рис. 9, а) все его сильные компоненты указаны штриховыми линиями, две его односторонние компоненты $\langle\{x_1, x_2, x_3, x_5\}\rangle$ и $\langle\{x_1, x_2, x_3, x_4\}\rangle$ изображены на рис. 9, б.

Свойство «связности» в неориентированном графе G влечет, помимо разбиения множества вершины $X(G)$, также и разбиение множества ребер

$$E(G) = \bigcup_{i=1}^p E_i, E_i \cap E_{i'} = \emptyset, i \neq i',$$

где E_i — множество ребер в i -й компоненте $\langle X_i \rangle$. Свойство «сильной связности», вообще говоря, не порождает разбиения множества дуг ориентированного графа. Последнее обстоятельство приводит к понятиям конденсации и базы ориентированного графа.

Конденсация и база ориентированного графа. Конденсацией ориентированного графа $G = (X, E)$ называют граф $G^* = (X^*, E^*)$, множество вер-

шин X^* и дуг E^* которого определяется следующим образом:

а) каждой сильной компоненте $\langle X_i^s \rangle$, $i = 1, 2, \dots, p_s$ графа G сопоставляется вершина $x_i^* \in X^*$, т. е. $|X^*| = p_s$;

б) дуга $(x_i^*, x_j^*) \in E^*$ тогда и только тогда, когда в графе G существует дуга $(x_i, x_j) \in E(G)$, такая, что $x_i \in X_i^s$, а $x_j \in X_j^s$.

Конденсация G^* графа G получается в результате стягивания (см. с. 37) всех дуг компонент сильной связности графа G .

Конденсация G^* графа G (см. рис. 9, а) изображена на рис. 9, в (вершинам x_1^* и x_2^* отвечают множества вершин графа G $\{x_1, x_2\}$ и $\{x_7, x_8, x_9\}$), не содержит контуров, т. е. G^* — ациклический граф. *Базой $B(G)$ графа G* называют минимальное множество вершин графа, из которых достижима любая вершина графа G . Для графа G (см. рис. 9, а) базой является множество $\{x_1, x_7\}$, другая база графа G — $\{x_2, x_9\}$.

Свойства а: в базе графа $B(G)$ нет вершины, которая достижима из другой вершины $B(G)$; в базе графа $B(G)$ нет двух вершин, которые при-

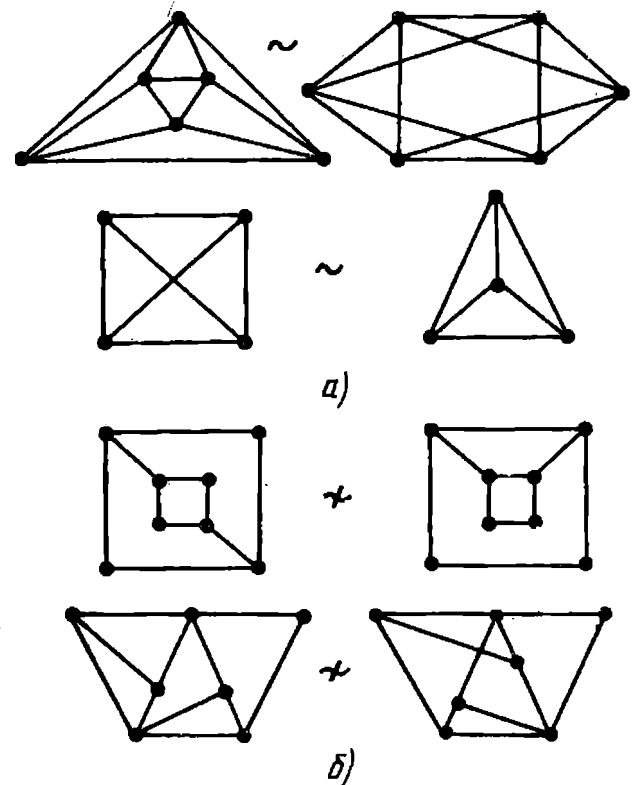


Рис. 10

надлежат одной и той же сильной компоненте связности графа G ; база $B^*(G^*)$ конденсации G^* графа G состоит из таких вершин графа G^* , полустепени захода которых равны нулю. Из свойств и определения конденсации графа следует способ построения одной из баз графа G : из каждой сильной компоненты $\langle X_i^s \rangle$ графа G , соответствующей вершине x_i^* в базе конденсации $B^*(G^*)$, надо взять любую вершину $x_{k_1} \in X_i^s$, тогда множество вершин $\{x_{k_1}, x_{k_2}, \dots, x_{k_q}\}$, где q — число вершин в базе конденсации $B^*(G^*)$, является базой исходного графа G .

Изоморфизм графов. При изображении двух графов на плоскости иногда можно путем непрерывной деформации ребер и перемещения вершин перевести один граф в другой, т. е. такие два графа в этом смысле эквивалентны. Точный смысл этой эквивалентности — изоморфизм графов.

Графы $G = (X, E)$ и $G' = (X', E')$ называют *изоморфными* ($G \sim G'$), если существует взаимно-однозначное соответствие между множеством вершин X и X' , сохраняющее отношение смежности вершин, т. е. такое взаимно-однозначное отображение $\varphi: X \rightarrow X'$, при котором справедлива импликация

$$(x_i, x_j) \in E(G) \Rightarrow$$

$$\Rightarrow (\varphi(x_i), \varphi(x_j)) \in E'(G').$$

Для псевдографов G и G' изоморфизм означает, что существует взаимно-однозначное соответствие между множествами вершин X и X' и множествами ребер (дуг) E и E' , сохраняющее отношение инцидентности.

Из определения изоморфизма графов можно вывести необходимые условия изоморфизма двух графов. Например, равенство количества вершин и дуг, числа вершин с одинаковыми полустепенями, числа контуров (простых циклов), совпадение других числовых и топологических характеристик изоморфных графов. Конечный набор признаков, достаточный для установления изоморфизма двух произвольных графов, неизвестен.

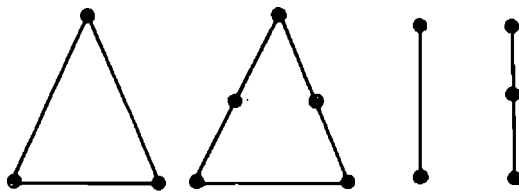


Рис. 11

На рис. 10, а приведены изоморфные друг другу графы, на рис. 10, б неизоморфные графы.

Графы G и G' называют *гомеоморфными* (рис. 11), если они изоморфны или могут стать изоморфными в результате применения повторных операций включения вершин или слияния двух смежных ребер (дуг) (см. ниже).

2. ОПЕРАЦИИ НАД ГРАФАМИ

С помощью различных операций — унарных (т. е. над одним графом) или бинарных (т. е. над парой графов) можно строить графы из более простых, переходить от одного графа к более простому, разбивать графы на более простые, в заданном классе графов переходить от одного графа к другому и т. д.

Унарные операции. Удаление ребра (дуги) (x_i, x_j) , при этом вершины x_i и x_j сохраняются (рис. 12, а);

добавление ребра (дуги) (рис. 12, б);

удаление вершины (вместе с инцидентными ее ребрами или дугами) (рис. 12, в);

добавление вершины x_i , которую можно соединить ребрами или дугами с другими вершинами графа (рис. 12, г);

включение вершины x_i на ребро (x_k, x_j) посредством введения ребер (x_k, x_i) и (x_i, x_j) (рис. 12, д);

слияние двух смежных ребер (дуг) $e_1 = (x_i, x_k)$ и $e_2 = (x_k, x_j)$; при этом вершина x_k удаляется, и пара e_1, e_2 заменяется одним ребром (дугой) (x_i, x_j) (рис. 12, е);

стягивание ребра (дуги) (x_i, x_j) ; при этом (x_i, x_j) удаляется, а концевые вершины x_i и x_j отождествляются (рис. 12, ж);

замыкание или отождествление пары смежных вершин x_i и x_j ; при этом они заменяются такой новой вершиной x_{ij} , что все ребра (дуги) в графе G ,

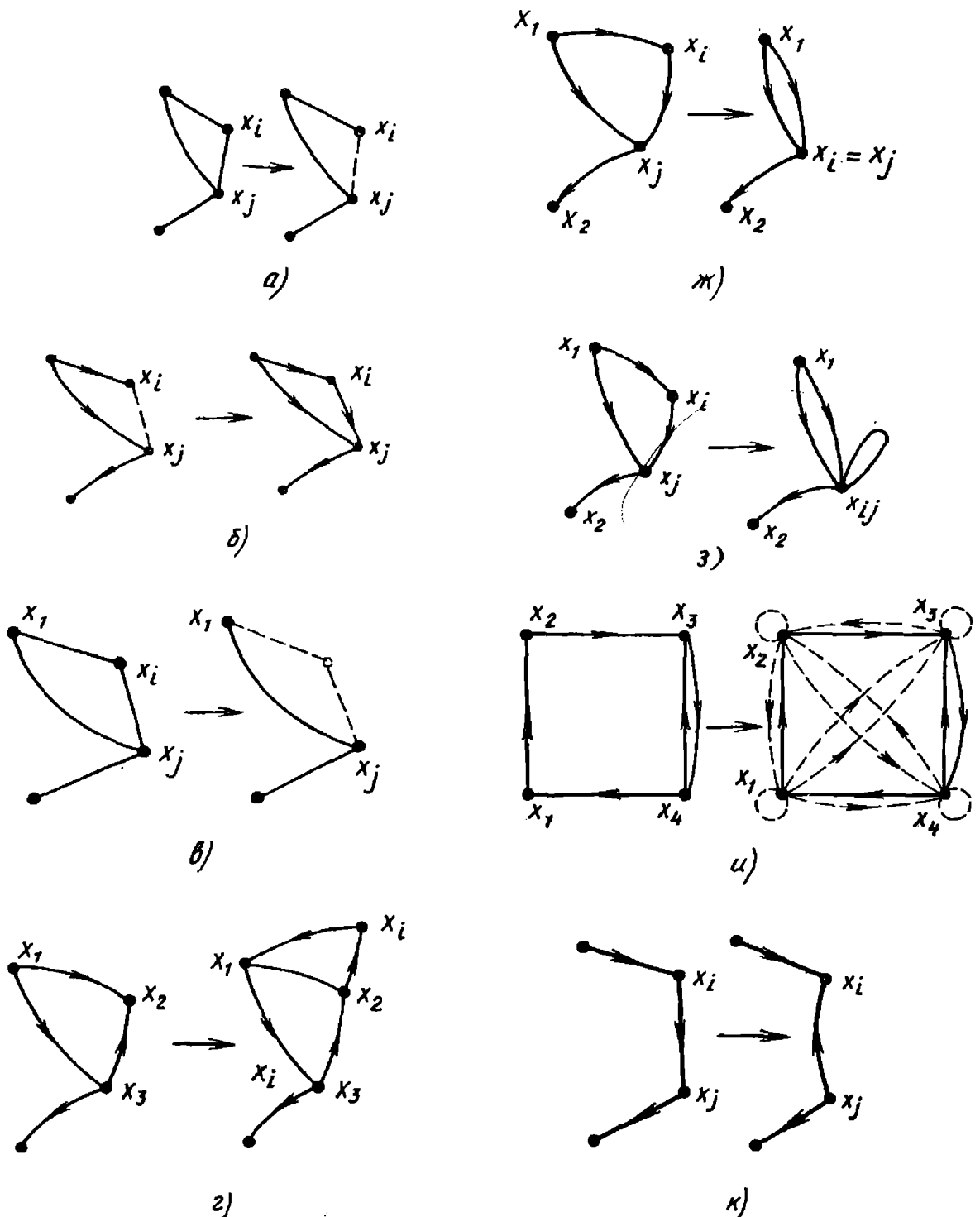
инцидентные x_i и x_j , становятся инцидентными новой вершине (рис. 12, э);

транзитивное замыкание ориентированного графа G — добавление к графу G минимально возможного числа дуг, соединяющих все такие пары вершин $\{x_i, x_j\}$ графа G , что вершина x_j достижима из x_i ($x_j \in D(x_i)$) (рис. 12, и);

Граф G^* , получающийся из G с помощью операции транзитивного замыкания, называют *транзитивным за-*

мыканием графа G . Исходный граф G является подграфом G^* . Матрица смежности графа G^* — матрица достижимости исходного графа G (см. с. 46);
переориентация дуги (x_i, x_j) — удаление (x_i, x_j) и добавление дуги (x_j, x_i) (рис. 12, к);

транзитивная ориентация неориентированного графа $G = (X, E)$ — ориентация всех ребер E , если это возможно, таким образом, что получившийся граф \vec{G} транзитивен (рис. 12, л).



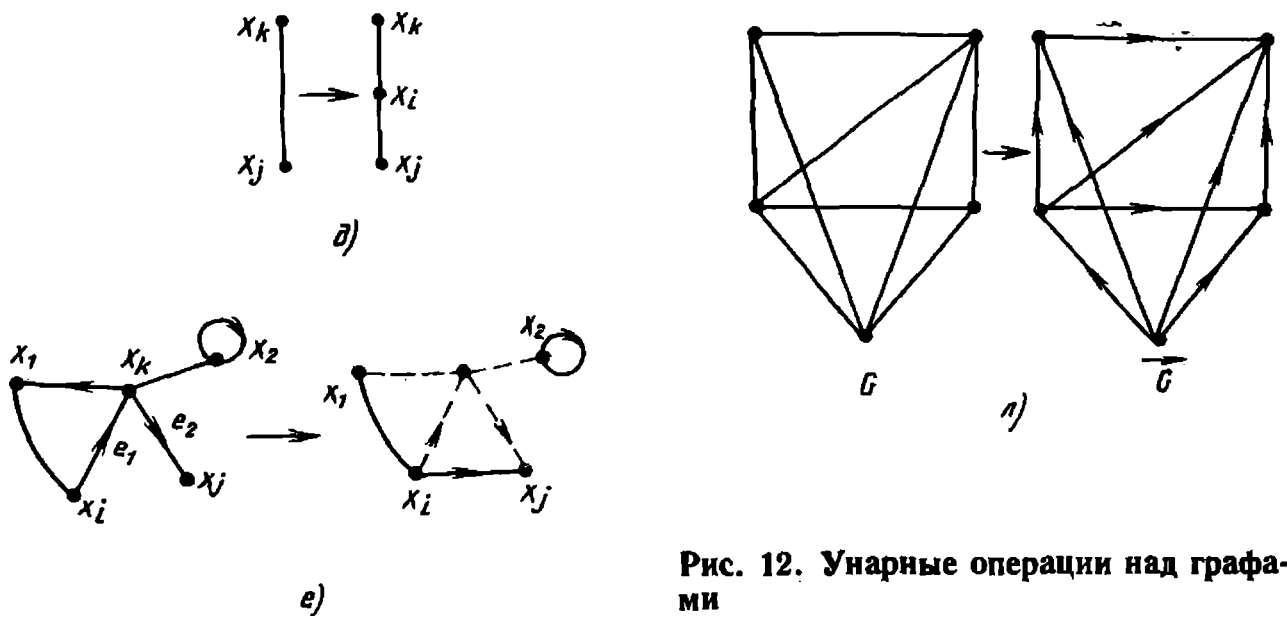


Рис. 12. Унарные операции над графами

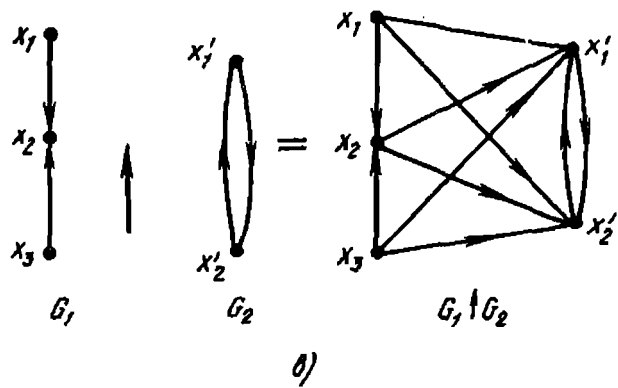
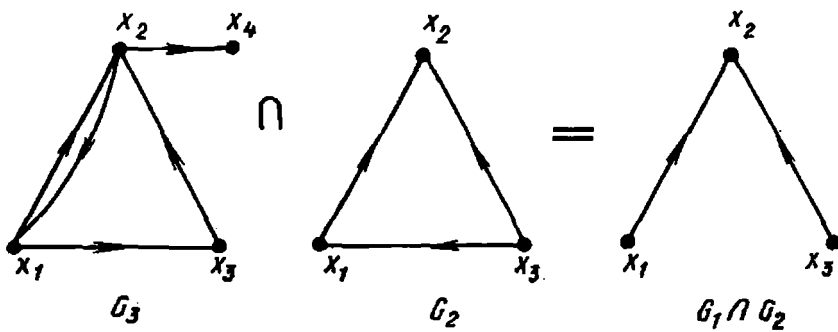
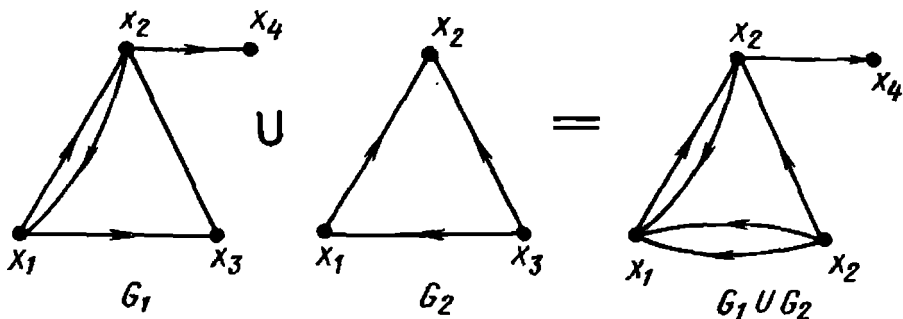


Рис. 13.

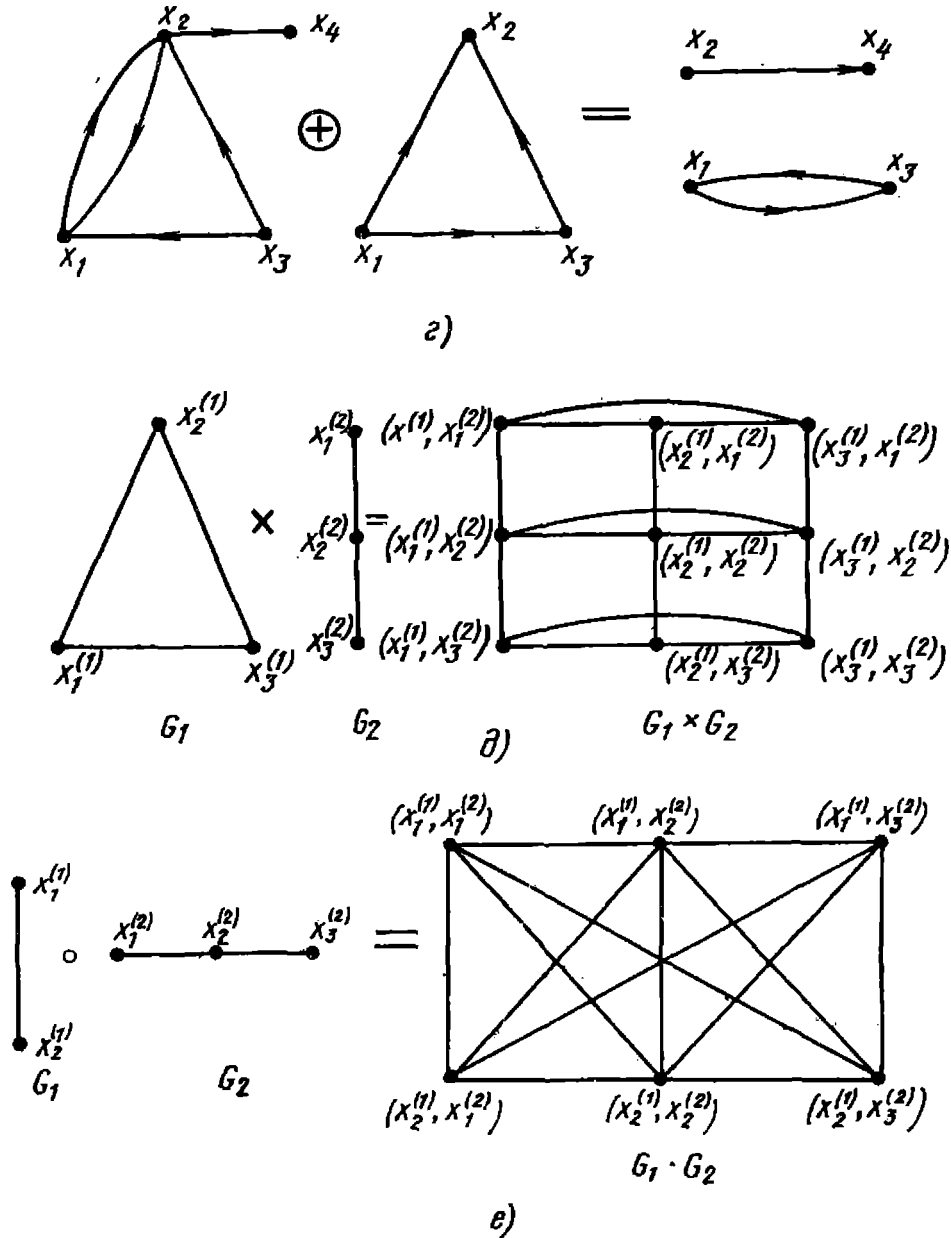


Рис. 13. Бинарные операции над графами

Замечание. Алгоритмы транзитивной ориентации графа см. в [5].

Бинарные операции: объединение графов $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемое как $G_1 \cup G_2$, представляет собой граф $G_3 = (X_1 \cup X_2, E_1 \cup E_2)$ (рис. 13, а);

пересечение графов $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемое как $G_1 \cap G_2$, представляет собой граф $G_3 = \{X_1 \cap X_2, E_1 \cap E_2\}$ (рис. 13, б);

соединение графов $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемое как $G_1 \uparrow G_2$, представляет собой граф $G_3 = \{X_1 \cup X_2, E_1 \cup E_2 \cup E_{1,2}\}$, где $E_{1,2}$ — множество дуг, соединяющих каждую вершину графа G_1 с каждой вершиной графа G_2 (рис. 13, в);

кольцевая сумма графов (или сумма графов по модулю 2) $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемая как $G_1 \oplus G_2$ представляет собой граф G_3 , порожденный множеством дуг (ребер) $(E_1 \cup E_2) \setminus (E_1 \cap E_2)$ (рис. 13, з);

произведение графов $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемое как $G_1 \times G_2$, представляет собой граф $G_3 = (X_1 \times X_2, E_3)$, причем две вершины $(x_i^{(1)}, x_j^{(1)})$ и $(x_k^{(1)}, x_m^{(1)})$ смежны в том и только том случае, когда либо $x_i^{(1)} = x_k^{(1)}$ и вершина $x_j^{(2)}$ смежна с $x_m^{(2)}$ в графе G_2 , либо $x_k^{(2)} = x_m^{(2)}$ и вершина $x_i^{(1)}$ смежна с $x_j^{(1)}$ в графе G_1 (рис. 13, д);

композиция графов $G_1 = (X_1, E_1)$ и $G_2 = (X_2, E_2)$, обозначаемая как $G_1 \cdot G_2$, представляет собой граф $G_3 = (X_1 \times X_2, E_3)$ (рис. 13, е), в котором две вершины $(x_i^{(1)}, x_j^{(2)})$ и $(x_k^{(1)}, x_m^{(2)})$ смежны тогда и только тогда, когда $x_i^{(1)}$ смежна с $x_k^{(1)}$ в графе G_1 или $x_i^{(1)} = x_k^{(1)}$ и вершина $x_j^{(2)}$ смежна с $x_m^{(2)}$ в графе G_2 .

3. СПЕЦИАЛЬНЫЕ ТИПЫ ГРАФОВ

Полные графы. Неориентированный граф, в котором каждая пара вершин смежна, называют **полным**. Полный граф с n вершинами обозначают через K_n . Число ребер в K_n равно $\frac{n(n-1)}{2}$.

Граф K_5 изображен на рис. 14, а. Ориентированный граф G называют **полным**, если его неориентированный дубликат \bar{G} — полный граф, т. е. каждая пара вершин x_i и x_j в графе G соединена по крайней мере одной дугой (x_i, x_j) либо (x_j, x_i) (рис. 14, б).

Регулярные или однородные графы. Неориентированный граф $G = (X, E)$, для которого степень любой вершины $x \in X(G)$ $d(x) = r, r \in N$, называют **r -регулярным** или **r -однородным** (рис. 15). При $r = 0$ граф G называют **нуль-графом** и обозначают $N_n, n = |X(G)|$. Граф K_5 — 4-регулярный граф. Связный 2-регулярный граф G называют **циклическим графом** и обозначают через $C_n, n = |X(G)|$ (рис. 15, а).

Двудольные графы. Неориентированный граф $G = (X, E)$ называется **двудольным**, если $X = X_1 \cup X_2, X_1 \cap X_2 = \emptyset$ и $(x_i, x_j) \in E$ тогда и только тогда, когда $x_i \in X_1, x_j \in X_2$.

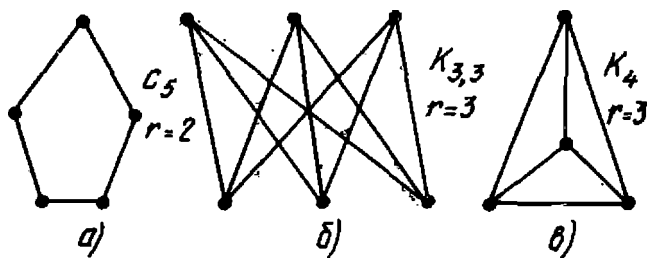


Рис. 15. Регулярные графы:

а — циклический; б — Куратовского; в — полный

Неориентированный граф G является **двудольным** тогда и только тогда, когда он не содержит циклов нечетной длины.

Двудольный граф G с разбиением вершин X_1 и X_2 называют **полным**, если для каждой пары вершин $x_i \in X_1$ и $x_j \in X_2$ существует ребро (x_i, x_j) . Если $|X_1| = n, |X_2| = m$, то полный двудольный граф обозначают через $K_{n,m}$.

На рис. 16, а изображен граф $K_{3,4}$, на рис. 15, б — $K_{3,3}$; граф $K_{1,n}$, называемый **звездой**, для $n = 5$ изображен на рис. 16, б.

Графы K_5 (см. рис. 14, а) и $K_{3,3}$ (см. рис. 15, б) называют **графами Куратовского**. Они играют важную роль в проблеме планарности графов.

Граф $K_{n,m}$ — соединение графов N_n и N_m . Объединение графов $K_{1,n}$ и $C_{n-1}, n \geq 3$ дает граф W_n , называемый **колесом** (рис. 16, в).

Планарные графы. Неориентированный граф $G = (X, E)$ называют **планарным**, если он может быть нарисован на плоскости (или сфере) таким образом, что произвольные два его ребра не пересекаются друг с другом (за исключением, быть может, вершины, инцидентной этим ребрам). Неориентированный граф, изоморфный планарному, называют **плоским**. Справедлива теорема Понтрягина—Куратовского: граф G планарен тогда и только тогда, когда он не содержит подграфа, гомеоморфного графу K_5 или $K_{3,3}$.

Двойственные графы. Пусть G — планарный граф и f_1, f_2, \dots, f_r — области (границы), на которые граф разбивает плоскость, включая внешнюю.

Геометрически двойственным графом G^* к планарному графу G называют

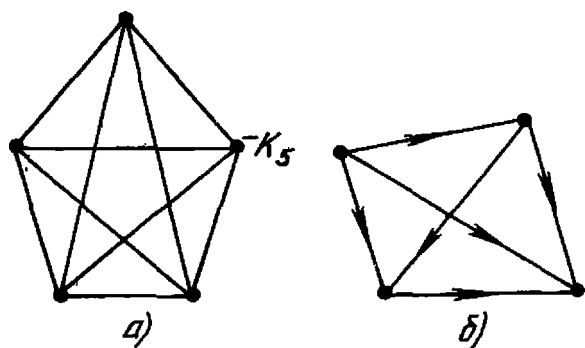


Рис. 14. Полные графы

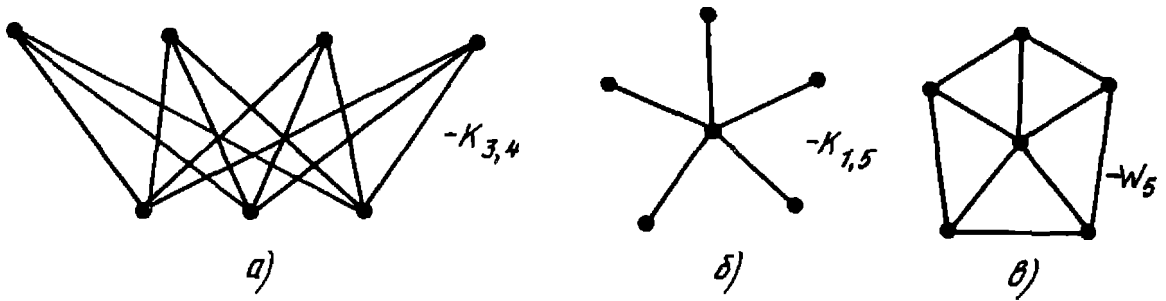


Рис. 16 Виды графов

граф (неориентированный псевдограф), определяемый следующим образом:

1) G^* имеет r вершин $x_1^*, x_2^*, \dots, x_r^*$; вершина x_i^* , $i = 1, 2, \dots, r$, соответствует области f_i ; 2) G^* имеет столько же ребер, сколько и G ; 3) если ребро e графа G является общим для областей f_i и f_j , то соответствующее ребро e^* графа G^* соединяет вершины x_i^* и x_j^* . Каждое ребро графа G является общим самое большее для двух областей, причем возможно, что ребро лежит в одной области. Процедура построения графа G^* , двойственного к графу G (рис. 17, а), показана на рис. 17, б.

Свойства G^* и G : если G — плоский и связный граф, то и G^* — плоский и связный; если G — плоский и связный граф, то $(G^*)^*$ изоморфен G ; если G — плоский и связный граф с n вершинами, m ребрами и r гранями, то для двойственного к нему графа G^* с n^* вершинами, m^* ребрами и r^* гранями справедливы соотношения: $n^* = r$, $m^* = m$, $r^* = n$.

Дополнительный граф. Дополнительным графом к неориентированному графу $G = (X, E)$ (или дополнением графа G) называют граф \bar{G} с множеством вершин $X = X(G)$, в котором

две вершины смежны тогда и только тогда, когда они не смежны в G .

Если G — граф с n вершинами рассматривать как подграф графа K_n , то дополнительный к нему граф \bar{G} можно построить удалив из графа K_n все ребра, принадлежащие G .

Эйлеровы графы. Эйлеровым циклом в неориентированном графе G называют замкнутую цепь, содержащую все ребра графа G . Граф G , допускающий эйлеров цикл, называют эйлеровым (рис. 18, а). Связный неориентированный граф G эйлеров тогда и только тогда, когда степень каждой его вершины четная (теорема Эйлера).

Ориентированным эйлеровым циклом в ориентированном графе G называют ориентированный цикл, содержащий все дуги графа. Граф, допускающий ориентированный цикл, называют ориентированным эйлеровым графом (рис. 18, б).

Слабо связный ориентированный граф G — ориентированный эйлеров граф тогда и только тогда, когда для любой вершины x графа G справедливо равенство $d_-(x) = d_+(x)$.

Гамильтоновы графы. Гамильтоновым циклом в неориентированном графе G называют замкнутый простой цикл, содержащий все вершины графа,

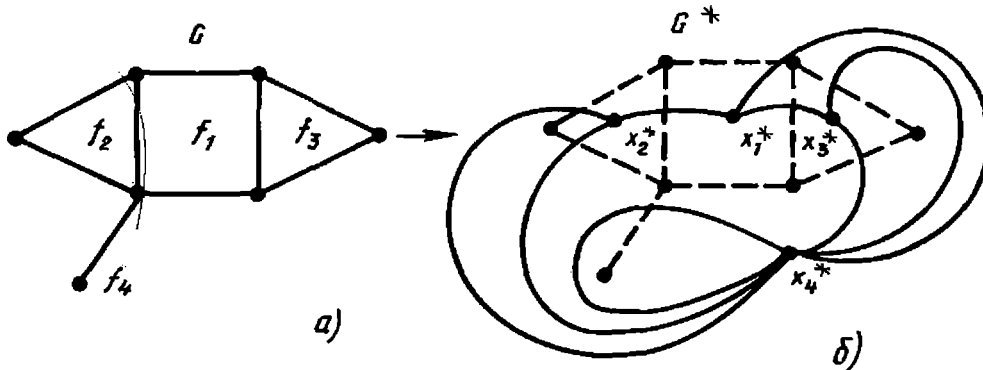


Рис. 17

т. е. маршрут, точно один раз проходящий через каждую вершину. Граф G , допускающий гамильтонов цикл, называют *гамильтоновым графом*. В общем случае довольно трудно установить, является ли граф гамильтоновым; в определенных случаях это удается сделать с помощью изоэдренных методов. Большинство известных теорем, в которых формулируются достаточные условия для существования гамильтонова цикла, справедливы для графов с достаточно большим числом ребер в G . Например: 1) G — гамильтонов граф, если $d(x) \geq n/2$ для любой вершины x графа G , $|X(G)| = n$; 2) G — гамильтонов граф, если $n \geq 3$, $m \geq (n^2 < 3n - 6)/2$. Отсюда: в графе K_n всегда существует гамильтонов цикл, причем число различных гамильтоновых циклов равно $(n - 1)!$

Ориентированным гамильтоновым графом называют ориентированный граф, имеющий гамильтонов контур, т. е. контур, содержащий все вершины графа G . Каждое из нижеследующих условий является достаточным условием того, что G — ориентированный гамильтонов граф:

1) G — сильно связный полный ориентированный граф;

2) G — сильно связный граф и $d_-(x) + d_+(x) \geq n$, $\forall x \in X(G)$;

3) $\min(\delta_-, \delta_+) \geq \frac{n}{2} - 1$, где

$$\delta_{\pm} = \min_{x \in X(G)} d_{\pm}(x).$$

Примеры гамильтоновых графов приведены на рис. 19.

Деревья, остовы, кодеревья. Неориентированное дерево может быть

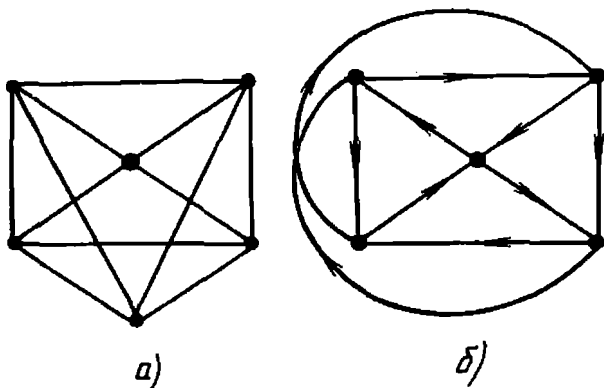


Рис. 18

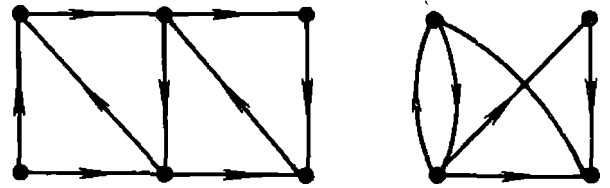


Рис. 19

определено одним из нижеследующих четырех эквивалентных определений. Неориентированное дерево есть: 1) связный ациклический граф; 2) граф, в котором существует лишь одна простая цепь между любыми двумя парами вершин графа; 3) связный граф, в котором $m = n - 1$; 4) ациклический граф, в котором при добавлении ребра, соединяющего две несмежные вершины, получается граф, имеющий точно один цикл.

Остовом $T(G)$ неориентированного графа G называют остовный подграф графа G , являющийся деревом. *Кодеревом $T^*(G)$* остова $T(G)$ называют граф, дополняющий $T(G)$ до G , т. е. граф, содержащий только те ребра из G , которые не входят в $T(G)$. Ребра остова $T(G)$ называют *ветвями $T(G)$* , а ребра кодерева $T^*(G)$ — *хордами или связями*.

Для графа G (рис. 20, а) два остова графа и соответствующие им кодеревья изображены на рис. 20, б.

Согласно определению дерева (см. п. 3) число ребер кодерева $T^*(G)$ остова $T(G)$ связного графа G равно $m - n + 1$, а из определения дерева (см. п. 4) следует, что добавление любой хорды $T^*(G)$ к $T(G)$ образует ровно один цикл c_i , $i = 1, 2, \dots, m - n + 1$, который называют *базисным циклом графа G относительно остова $T(G)$* . Множество всех таких циклов называют *базисным множеством циклов графа G относительно остова $T(G)$* . Любой цикл графа G можно выразить в виде кольцевой суммы (суммы по mod 2, см. с. 40) некоторых базисных циклов графа G по отношению к остову $T(G)$.

Пример представления цикла $c = \{3, 2, 1, 5, 4, 3\}$ графа G (см. рис. 20, а) через его базисные циклы c_1, c_2, c_3, c_4 (рис. 21, а) относительно остова T_1 (см. рис. 20, б) показан на рис. 21, б.

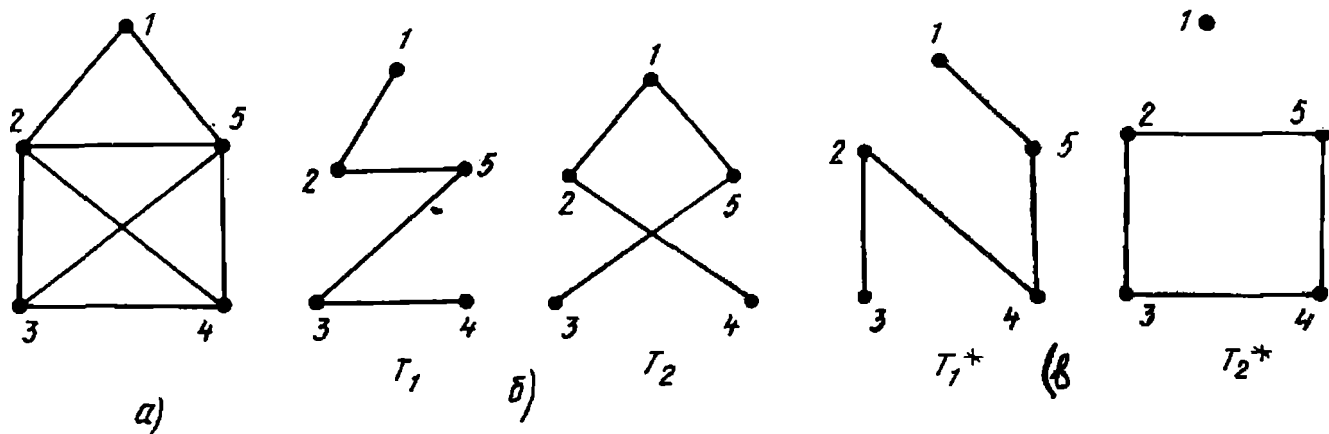


Рис. 20. Остова и кодеревья графа:

а — исходный граф; б — два остова; в — соответствующие кодеревья

Ориентированный граф G называют *деревом* (рис. 22, а), если его дубликат G также является деревом. Ориентированный граф G называют *ориентированным* или *корневым деревом*

(рис. 22, б), если он является деревом и имеет корень, т. е. вершину x_0 , из которой достижима любая другая вершина графа. Двоичным деревом называют ориентированное дерево, полустепень исхода каждой вершины которого не превышает 2 (рис. 22, в). Двоичные деревья полезны при анализе алгоритмов, в методах поиска информации и т. д.

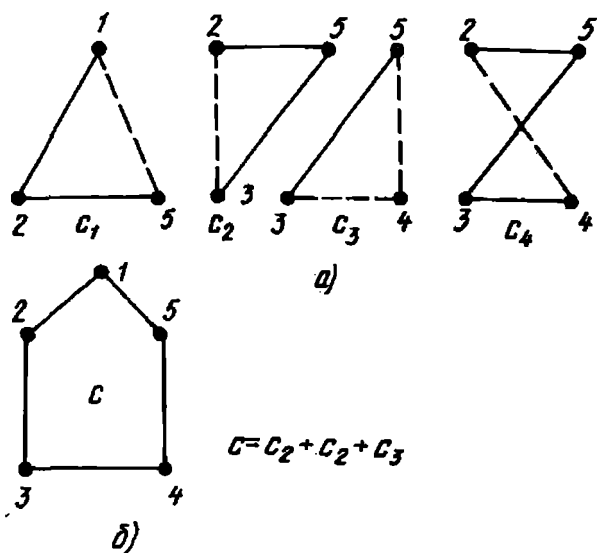


Рис. 21

Неразделимые графы. Вершину x_i графа G называют *точкой сочленения графа G* , если граф $G - x_i$ состоит из большего числа компонент связности, чем G . Изолированная вершина не есть точка сочленения. Связный граф без точек сочленения называют *неразделимым* (рис. 23, а); в противном случае — *разделимым* (рис. 23, б). *Блоком разделимого графа G* является максимальный неразделимый подграф графа G . Блоки разделимого графа G (см. рис. 23, б) изображены на рис. 23, в.

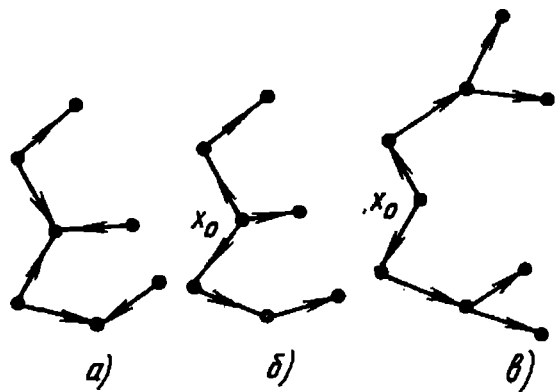


Рис. 22. Ориентированные деревья

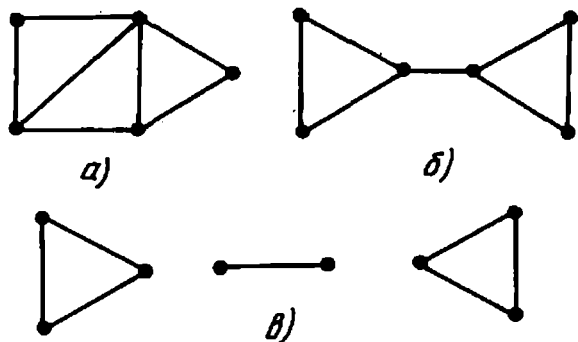


Рис. 23

4. МАТРИЦЫ ГРАФА. БИНАРНЫЕ ОТНОШЕНИЯ И ГРАФЫ

Многие проблемы теории графов могут быть сформулированы на языке матриц, ассоциированных с графами. Матричный язык особенно удобен при изучении структуры графа с использованием ЭВМ. В частности, матрицы инцидентий и смежности графа — один из способов его представления в ЭВМ.

Матрица инцидентий $B(G)$ ориентированного графа $G | X | = n, | E | = m$ является матрицей порядка $n \times m$ с элементами $b_{ij}, i = 1, 2, \dots, n; j = 1, 2, \dots, m$, определяемыми следующим образом;

$b_{ij} = 1$, если x_i — начальная вершина дуги e_j ;

$b_{ij} = -1$; если x_i — конечная вершина дуги e_j ;

$b_{ij} = 0$, если x_i не инцидентна дуге e_j .

Матрица инцидентий $B(G)$ неориентированного графа G — матрица порядка $n \times m$ с элементами $b_{ij}, i = 1, \dots, n; j = 1, 2, \dots, m, b_{ij} = 1$, если вершина x_i инцидентна ребру e_j ; $b_{ij} = 0$ в противном случае.

Для графов (рис. 24, а, б) матрицы инцидентий имеют вид соответственно

$$\begin{matrix} & e_1 & e_2 & e_3 & e_4 & e_5 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{vmatrix} +1 & +1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{vmatrix} \end{matrix},$$

$$\begin{matrix} & e_1 & e_2 & e_3 & e_4 & e_5 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{vmatrix} \end{matrix}.$$

Свойства $B(G)$ ориентированного графа G : $\text{rank } B(G) \leq n - 1$, т. е. n строк матрицы B линейно зависимы; $\text{rank } B(G) = n - p$, если G — слабо связный граф с p слабыми компонентами.

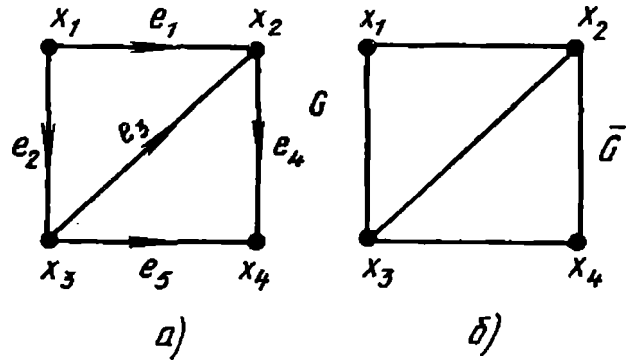


Рис. 24

Матрица смежности $A(G)$ графа $G, | X | = n$ — матрица порядка $n \times n$ с элементами $a_{ij}, i = 1, 2, \dots, n; j = 1, 2, \dots, n$, определяемыми следующим образом:

$$a_{ij} = \begin{cases} 1, & \text{если } (x_i, x_j) \in E(G); \\ 0, & \text{если } (x_i, x_j) \notin E(G). \end{cases}$$

Матрица смежности графа G (см. рис. 24, а) равна

$$\begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{vmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} \end{matrix}.$$

Симметрическая матрица смежности неориентированного графа (см. рис. 24, б) имеет вид

$$\begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{vmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{vmatrix} \end{matrix}.$$

Обозначим через G_n множество графов с n вершинами, в которых допускаются петли. Множество G_n находится во взаимно-однозначном соответствии с множеством A_n — булевых матриц порядка $n \times n$. Указанное соответствие определяется формулой: $(x_i, x_j) \in E(G) \Leftrightarrow a_{ij} = 1$.

Булеву матрицу P порядка $n \times n$ называют матрицей подстановки, если в каждой строке и каждом столбце матрицы P существует один и только

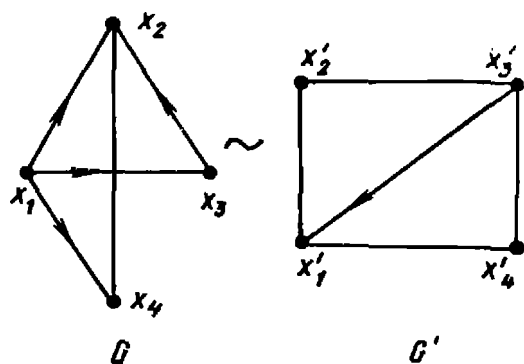


Рис. 25

один элемент, отличный от нуля.

Множество P_n всех матриц подстановок находится во взаимно-однозначном соответствии с множеством S_n всех подстановок длины n .

Указанное соответствие определяется формулами

$$\forall P = ((p_{ij})) \in P_n,$$

$$p_{ij} = 1 \iff \exists \pi \in S_n, j = \pi(i);$$

$$\forall \pi \in S_n, j = \pi(i) \iff P \in P_n, p_{ij} = 1.$$

Графы $G' = (X', E')$ и $G = (X, E)$ изоморфны тогда и только тогда, когда существует матрица подстановки $P \in P_n$, такая, что выполнено равенство

$$A(G) = P^t A(G') P.$$

Здесь $A(G)$, $A(G')$ — матрицы смежности графов соответственно G и G' ; P^t — матрица, транспонированная к P . Для изоморфных графов (рис. 25) матрица P имеет вид

$$\begin{array}{c|cccc}
 & 1 & 2 & 3 & 4 \\
 \hline
 1 & 0 & 0 & 1 & 0 \\
 2 & 1 & 0 & 0 & 0 \\
 3 & 0 & 0 & 0 & 1 \\
 4 & 0 & 1 & 0 & 0
 \end{array}$$

Пусть $[A^k(G)]_{ij}$ обозначает (i, j) — элемент k -й степени ($k \in \mathbb{N}$), $A^k(G)$ — матрицы смежности $A(G)$ ориентированного графа G . Тогда $[A^k(G)]_{ij}$ равен числу всех маршрутов длины k из вершины x_i в вершину x_j .

Матрица достижимости $D(G)$ графа G — матрица порядка $n \times n$ с элементами d_{ij} , $i = 1, 2, \dots, n$; $j = 1, 2, \dots, n$:

$$d_{ij} = \begin{cases} 1, & \text{если } x_j \text{ достижима из } x_i; \\ 0 & \text{в противном случае.} \end{cases}$$

Всегда $d_{ii} = 0$.

Матрица достижимости $D(G)$ графа G есть в то же время матрица смежности транзитивного замыкания $T^*(G)$ графа G .

Матрица расстояний $R(G)$ графа G — матрица порядка $n \times n$ с элементами r_{ij} , $i = 1, 2, \dots, n$; $j = 1, 2, \dots, n$, где r_{ij} — расстояние между вершинами x_i и x_j .

Матричный элемент $r_{ij} = \min_k \{[A^k(G)]_{ij} \neq 0\}$, если вершина x_j достижима из x_i . Для графа (см. рис. 24, а) матрицы достижимости и расстояний равны соответственно

$$\begin{array}{c|cccc}
 & x_1 & x_2 & x_3 & x_4 \\
 \hline
 x_1 & 1 & 1 & 1 & 1 \\
 x_2 & 0 & 1 & 0 & 1 \\
 x_3 & 0 & 1 & 1 & 1 \\
 x_4 & 0 & 0 & 0 & 0
 \end{array}$$

$$\begin{array}{c|cccc}
 & x_1 & x_2 & x_3 & x_4 \\
 \hline
 x_1 & 0 & 1 & 1 & 2 \\
 x_2 & \infty & 0 & \infty & 1 \\
 x_3 & \infty & 1 & 0 & 1 \\
 x_4 & \infty & \infty & \infty & 0
 \end{array}$$

Матрица полустепеней захода $K(G)$ ориентированного графа G , $|X| = n$ — матрица порядка $n \times n$, определяемая следующим образом:

$$k_{ij} = \begin{cases} -1, & \text{если } i \neq j \\ \text{и } \exists (x_i, x_j) \in E(G); \\ d_-(x_i), & i = j. \end{cases}$$

Обозначим через $K_{ij}(G)$ матрицу, полученную удалением i -й строки и j -го столбца из матрицы $K(G)$.

Число $n(T)$ ориентированных остовов графа, для которых вершина x_r — корневая, равно $\det K_{rr}$.

Ориентированный граф G является ориентированным деревом с корнем в вершине x_r тогда и только тогда, когда

$$k_{pp} = \begin{cases} 0, & p = r \\ 1, & p \neq r \end{cases}$$

и $\det K_{rr} = 1$.

Бинарные отношения и графы. На множестве R задано бинарное отношение ρ , если задано подмножество $M_\rho \subseteq R \times R$. Говорят, что элемент $a \in R$ находится в отношении ρ с элементом $b \in R$, если $(a, b) \in M_\rho$, и записывают в виде $a\rho b$. Пусть R — конечное множество с n элементами $R = X = \{x_1, \dots, x_n\}$.

Матрица A_ρ бинарного отношения ρ на X — булева матрица порядка $n \times n$ с элементами:

$$a_{ij} = \begin{cases} 1, & x_i\rho x_j; \\ 0, & \text{в противном случае.} \end{cases}$$

Множество всех бинарных отношений $R(X)$, заданных на множестве X с n элементами, находится во взаимно-однозначном соответствии с множеством A_n булевых матриц порядка n . Указанное соответствие определяется формулой

$$x_i\rho x_j \iff a_{ij} = 1.$$

Взаимно-однозначные соответствия, установленные выше, между множествами всех бинарных отношений $R(X)$, всех булевых матриц A_n и всех графов с n вершинами (и возможно с петлями) G_n позволяют изучать свойства элементов любого из трех множеств и преобразований (операций) над его элементами путем перехода к любому из оставшихся двух его эквивалентных реализаций. Например, ориентированный граф — удачный способ представления бинарного отношения ρ на конечном множестве X . Вершины такого графа представляют собой элементы множества X , а дуги — упорядоченные пары элементов из X , находящихся в отношении ρ . Известные операции над бинарными отношениями естественным образом порождают операции над соответствующими им графами.

Другим примером использования соответствия $R(X) \leftrightarrow G_n$ являются следующие (эквивалентные данным на с. 35 определениям) определения компоненты связности (сильной связности) графа:

1) i -я компонента связности неориентированного графа G , $i = 1, \dots, p$ — граф $\langle X_i \rangle$, порожденный множеством вершин X_i , где X_i — i -й класс эквивалентности, заданной на множестве

$X(G)$ отношением достижимости $\rho_g(x_i\rho_g x_j) \iff \in D(x_i)$;

2) i -я сильная компонента ориентированного графа G , $i = 1, 2, \dots, p_s$ — граф $\langle X_i^s \rangle$, порожденный множеством вершин X_i^s , где X_i^s — i -й класс эквивалентности, заданной на множестве $X(G)$ отношением сильной достижимости $\rho_{с.д}$

$$(x_i\rho_{с.д} x_j) \iff (x_j \in D(x_i)) \wedge \wedge (x_i \in D(x_j)).$$

Соответствие $G_n \leftrightarrow A_n$ позволяет естественным образом перенести операции над графами на матрицы смежности графов. Так, например, операциям объединения \cup и пересечения \cap двух графов G_1 и G_2 отвечают следующие операции над их матрицами смежности $A(G_1) = \|a_{ij}^1\|$, $A(G_2) = \|a_{ij}^2\|$: $A(G_1) \cup A(G_2) = \|a_{ij}^1 \vee a_{ij}^2\|$ и $A(G_1) \cap A(G_2) = \|a_{ij}^1 \wedge a_{ij}^2\|$, где \vee, \wedge — булевы операции сложения и умножения (см. гл. 2).

СПИСОК ЛИТЕРАТУРЫ

1. Басакер Р., Саати Т. Конечные графы и сети. М.: Наука, 1976.
2. Белов В. В., Воробьев Е. М., Шаталов В. Е. Теория графов. М.: Высшая школа, 1976.
3. Берж К. Теория графов и ее применение. М.: ИЛ, 1962.
4. Евстигнеев В. А. Применение теории графов в программировании. М.: Наука, 1985.
5. Кристофидес Н. Теория графов. Алгоритмический подход. М.: Мир, 1978.
6. Кузнецов О. П., Адельсон-Вельский Г. М. Дискретная математика для инженера. М.: Энергия, 1980.
7. Мальцев А. И. Алгебраические системы. М.: Наука, 1976.
8. Рейнгольд Э., Нивергельт Ю., Део Н. Комбинаторные алгоритмы. Теория и практика. М.: Мир, 1980.
9. Свами М., Тхуласираман К. Графы, сети и алгоритмы. М.: Мир, 1984.
10. Сигорский В. П. Математический аппарат инженера. Киев: Техника, 1975.
11. Харари Ф. Теория графов. М.: Мир, 1973.

1. ОБЫКНОВЕННЫЕ
ДИФФЕРЕНЦИАЛЬНЫЕ
УРАВНЕНИЯ

Основные понятия и определения. Обыкновенным дифференциальным уравнением n -го порядка относительно неизвестной функции $y(x)$ вещественной переменной x называют выражение вида

$$F(x, y, y', y'', \dots, y^{(n)}) = 0, \quad (1)$$

где $F(x, y, p_1, p_2, \dots, p_n)$ — заданная функция переменных x, y, p_1, \dots, p_n ; $y', y'', \dots, y^{(n)}$ — производные функции $y(x)$.

Решением уравнения (1) на подмножестве A числовой прямой R называют такую n раз дифференцируемую на A функцию, при которой функция $F(x, y(x), y'(x), \dots, y^{(n)}(x))$ как функция переменной x тождественно равна нулю при всех $x \in A$.

Уравнение (1) называют уравнением n -го порядка нормального вида, если оно разрешено относительно старшей производной, т. е.

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}). \quad (2)$$

Уравнение (1) называют линейным уравнением n -го порядка, если функция $F(x, y, p_1, \dots, p_n)$ — линейная функция переменных y, p_1, \dots, p_n , т. е.

$$F(x, y, p_1, \dots, p_n) = q(x) + a_0(x)y + a_1(x)p_1 + \dots + a_n(x)p_n \quad (3)$$

и при этом $a_n(x) \neq 0$.

Если функции $a_i(x)$, $0 \leq i \leq n$ постоянны, то уравнение (3) называют линейным уравнением n -го порядка с постоянными коэффициентами.

Линейное дифференциальное уравнение n -го порядка называют однородным, если $q(x) \equiv 0$, и неоднородным в противном случае.

Системой n дифференциальных уравнений первого порядка называют систему n дифференциальных уравнений первого порядка

$$\frac{dy_i}{dx} = f_i(x, y_1, \dots, y_n), \quad 1 \leq i \leq n \quad (4)$$

относительно неизвестных функций $y_1(x), \dots, y_n(x)$, где $f_i(x, p_1, \dots, p_n)$, $1 \leq i \leq n$ — заданные функции переменных x, p_1, \dots, p_n . Число n называют размерностью системы.

Вектор-функцию $Y(x) = (y_1(x), \dots, y_n(x))$ называют решением системы (4) на подмножестве A вещественной прямой R , если функции $y_i(x)$ дифференцируемы на A и функция $y'_i(x)$ тождественно совпадает с функцией $f_i(x, y_1(x), \dots, y_n(x))$ для всех $x \in A$.

Связь между уравнением n -го порядка и системой из n уравнений первого порядка. Уравнение n -го порядка нормального вида эквивалентно системе из n уравнений первого порядка. Положим $y_1 = y, y_2 = y', \dots, y_{n+1} = y^{(n)}$ и рассмотрим систему

$$\begin{aligned} \frac{dy_i}{dx} &= y_{i+1}, \quad 1 \leq i \leq n-1, \\ \frac{dy_n}{dx} &= f(x, y_1, \dots, y_n) \end{aligned} \quad (5)$$

вместе с уравнением

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}).$$

Тогда всякое решение $Y(x) = (y_1(x), \dots, y_n(x))$ системы дает решение уравнения, если положить $y = y_1(x)$. Обратно, если $y = y(x)$ — решение уравнения, то $Y(x) = (y(x), y'(x), \dots, y^{(n-1)}(x))$ — решение системы.

Симметричная форма системы дифференциальных уравнений. Обозначим в (4) независимую переменную x через y_{n+1} . Пусть функции $F_1(y_1, y_2, \dots, y_{n+1}), \dots, F_{n+1}(y_1, \dots, y_{n+1})$ таковы, что

$$f_i(y_{n+1}, y_1, \dots, y_n) = \frac{F_i(y_1, \dots, y_{n+1})}{F_{n+1}(y_1, \dots, y_{n+1})},$$

тогда система (4) принимает вид

$$\frac{dy_k}{dy_{n+1}} = \frac{F_k}{F_{n+1}}, \quad 1 \leq k \leq n. \quad (6)$$

Пусть $Y(y_{n+1}) = (y_1(y_{n+1}), \dots, y_n(y_{n+1}))$ — решение этой системы. Пусть $y_{n+1} = \varphi_i(y_i)$ — решение уравнения $y_i = y_i(y_{n+1})$ относительно y_{n+1} . Рассмотрим y_i в качестве новой независимой переменной, тогда (5) преобразуется к виду

$$\frac{dy_k}{dy_i} = \frac{F_k}{F_i}, \quad 1 \leq k \leq n+1, \quad k \neq i.$$

Таким образом, в системе (6) y_1, \dots, y_{n+1} в равной мере могут быть приняты за независимую переменную. Указанную симметрию можно представить, записав уравнение (5) в симметричной по отношению к выбору независимой переменной, симметричной форме:

$$\begin{aligned} \frac{dy_1}{F_1(y_1, \dots, y_{n+1})} &= \\ &= \frac{dy_2}{F_2(y_1, \dots, y_{n+1})} = \dots = \\ &= \frac{dy_{n+1}}{F_{n+1}(y_1, \dots, y_{n+1})}. \end{aligned} \quad (7)$$

Задача Коши для уравнений n -го порядка и систем уравнений. Задачей

Коши для обыкновенного дифференциального уравнения называют следующую задачу. Для заданных чисел $x_0, y_0, y_1, \dots, y_{n-1}$ найти решение уравнения

$$F(x, y, y', \dots, y^{(n)}) = 0,$$

удовлетворяющее начальным условиям

$$y(x_0) = y_0, \quad y'(x_0) = y_1,$$

$$y^{(n-1)}(x_0) = y_{n-1}. \quad (7')$$

Числа y_0, y_1, \dots, y_{n-1} называют начальными данными в точке x_0 . Функцию $y(x)$, являющуюся решением уравнения (1) и удовлетворяющую начальным условиям (7'), называют решением задачи Коши.

Задача Коши для системы из n дифференциальных уравнений первого порядка формулируется следующим образом: для заданных чисел x_0, y_1^0, \dots, y_n^0 найти решение системы

$$\frac{dy_1}{dx} = f_1(x, y_1, \dots, y_n),$$

$$\frac{dy_n}{dx} = f_n(x, y_1, \dots, y_n),$$

удовлетворяющее начальным условиям:

$$y_1(x_0) = y_1^0, \quad y_n(x_0) = y_n^0. \quad (8)$$

Задача Коши для уравнения (1) с начальными условиями (7) эквивалентна задаче Коши для системы (5) с начальными условиями

$$y_1(x_0) = y_0, \quad y_2(x_0) = y_1, \quad \dots, \quad y_n(x_0) = y_{n-1}.$$

Существование и единственность решения задачи Коши для системы дифференциальных уравнений первого порядка. Решение задачи Коши для системы (4) с начальными условиями (8):

а) существует, если функции $f_i(x, p_1, \dots, p'_n)$ непрерывны в замкнутой области $U_{a,b} = \{(x, p_1, \dots, p_n) : x_0 - a \leq x \leq x_0 + a, y_i^0 - b \leq p_i \leq y_i^0 + b; i = 1, 2, \dots, n\}$, $a > 0$, $b > 0$;

б) единственно, если функции $f_i(x, p_1, \dots, p_n)$ удовлетворяют условию Липшица в области $U_{a,b}$, т. е. существует константа $K > 0$, что

$$\begin{aligned} |f_i(x, p_1, \dots, p_n) - \\ - f_i(x, p'_1, \dots, p'_n)| &< \\ < K \sum_{i=1}^n |p_i - p'_i| \end{aligned} \quad (9)$$

для всех $1 \leq i \leq n$ и $(x, p_1, \dots, p_n), (x, p'_1, \dots, p'_n) \in U_{a,b}$.

Замечания. 1. Если $M = \max |f_i(x, p_1, \dots, p_n)|$, где максимум берется

по всем $1 \leq i \leq n$ и $(x, p_1, \dots, p_n) \in U_{a,b}$, то решение существует и единственно на множестве $(x_0 - h, x_0 + h)$, $h = \min(a, b/M)$.

2. Условие Липшица выполняется при существовании и непрерывности в $U_{a,b}$ частных производных

$$\frac{\partial f_i}{\partial p_k} \quad (i, k = 1, 2, \dots, n).$$

3. Решение задачи Коши для уравнения (2) с начальными условиями (7') существует и единственно, если функция $f(x, p_1, \dots, p_n)$ непрерывна в некоторой замкнутой окрестности начальных данных и удовлетворяет в этой окрестности условию Липшица по p_1, p_2, \dots, p_n .

4. Решение задачи Коши для линейного дифференциального уравнения n -го порядка существует и единственно, если коэффициенты уравнения $a_0(x), \dots, a_n(x), q(x)$ — непрерывные функции и $a_n^2(x) > 0$.

5. Существует непустое подмножество $U_{a,b}^0 \subset U_{a,b}$, что решение задачи Коши $Y(x) = (y_1(x), y_2(x), \dots, y_n(x))$ для системы (4) с начальными условиями (8) является непрерывной функцией переменных x_0, y_1^0, \dots, y_n^0 , принадлежащих области $U_{a,b}^0$. Таким образом, $Y(x) = Y(x, x_0, y_1^0, \dots, y_n^0)$. Задача Коши для системы (4) с начальными условиями (8) может быть переформулирована следующим образом: найти такую вектор-функцию $Y(x) = (y_1(x), \dots, y_n(x))$, что

$$y_k(x) = y_k^0 + \int_{x_0}^x f_k(\xi, y_1(\xi), \dots, y_n(\xi)) d\xi \quad (10)$$

для любого $1 \leq k \leq n$.

Система уравнений (10) имеет непрерывное решение $Y(x) = (y_1(x), \dots, y_n(x))$, если:

а) функции $f_k(x, p_1, \dots, p_n)$ определены для $a < x < b$, $-\infty < p_1, \dots, p_n < \infty$ при фиксированных значениях p_1, \dots, p_n измеримы по x , а при почти каждом фиксированном x непрерывны по p_1, \dots, p_n ;

б) существует такая интегрируемая по Лебегу на интервале $a < x < b$ функция $M(x)$, что

$$|f_k(x, p_1, \dots, p_n)| \leq M(x), \quad 1 \leq k \leq n.$$

Далее, решение системы (10) единственно, если выполнено обобщенное условие Липшица, т. е. существует интегрируемая по Лебегу функция $N(x)$, что

$$|f_k(x, p_1, \dots, p_n) - f_k(x, y_1(x), \dots, y_n(x))| \leq N(x) \sum_{i=1}^n |p_i - y_i(x)|$$

для любых p_1, \dots, p_n . При этом решение непрерывно зависит от x_0, y_1^0, \dots, y_n^0 .

Интегралы уравнения n -го порядка. *Общий интеграл.* Решение задачи Коши для уравнения n -го порядка нормального вида определяется заданием начальных данных y_0, y_1, \dots, y_{n-1} . Таким образом, решение является функцией числовых параметров y_0, y_1, \dots, y_{n-1} , т. е.

$$y = y(x, y_0, \dots, y_{n-1}). \quad (11)$$

Дифференцируя $n-1$ раз (11) по переменной x , получим систему алгебраических уравнений

$$y^{(k)} = y^{(k)}(x, y_0, \dots, y_{n-1}), \quad 0 \leq k \leq n-1. \quad (12)$$

Решая систему (12) относительно y_0, \dots, y_{n-1} и подставляя полученные решения в выражение $y^{(n)} = y^{(n)}(x, y_0, \dots, y_{n-1})$, получим

$$y^{(n)} = f(x, y, y^{(1)}, \dots, y^{(n-1)}),$$

т. е. по n -параметрическому семейству (11) мы восстановили уравнение, решением которого является любая функция семейства (11).

Функцию $\Phi(x, y, p_1, \dots, p_n)$ называют **общим интегралом уравнения (1)**, если, исключая из уравнений

$$\frac{d^k}{dx^k} (\Phi(x, y(x), p_1, \dots, p_n)) = 0, \quad 0 \leq k \leq n$$

параметры p_1, \dots, p_n , получим уравнение (1) или эквивалентное ему уравнение.

Промежуточный и первый интегралы. Функцию $\Psi(x, y(x), y^{(k)}(x), p_{k+1}, \dots, p_n)$ называют *промежуточным интегралом*, если уравнение (1) можно получить, исключая из системы

$$\frac{d^j}{dx^j} \Psi(x, y(x), \dots, y^{(k)}(x),$$

$$p_{k+1}, \dots, p_n) = 0, \quad 0 \leq j \leq n - k$$

параметры p_{k+1}, \dots, p_n . Если $k = n - 1$, то промежуточный интеграл называют *первым интегралом*. Пусть для уравнения (1) известен промежуточный интеграл, тогда уравнение

$$\Psi(x, y, y', \dots, y^{(k)},$$

$$p_{k+1}, \dots, p_n) = 0,$$

рассматриваемое как уравнение относительно неизвестной функции $y = y(x)$, является уравнением порядка $k < n$ и теоретически может считаться более простым.

Примеры построения решений уравнений, использующие общий интеграл. Рассмотрим уравнение

$$F(x, y^{(k)}, y^{(k+1)}, \dots, y^{(n)}) = 0. \quad (13)$$

Пусть $z = y^{(k)}$, тогда уравнение (13) примет вид

$$F(x, z, z^{(1)}, \dots, z^{(n-k)}) = 0, \quad (14)$$

т. е. является уравнением порядка $n - k$. Пусть $\Phi(x, z, p_1, \dots, p_{n-k})$ — общий интеграл уравнения (14), тогда уравнение (13) эквивалентно уравнению

$$\Phi(x, y^{(k)}, p_1, \dots, p_{n-k}) = 0, \quad (15)$$

и, следовательно, (15) является промежуточным интегралом уравнения (13). Решая (15) относительно $y^{(k)}$, получим

$$y^{(k)} = Q(x, p_1, \dots, p_{n-k}).$$

Решая последнее уравнение, получим общий интеграл уравнения (13).

Пусть уравнение (1) не содержит явно x , т. е. имеет вид

$$F(y, y', \dots, y^{(n)}) = 0. \quad (16)$$

Введем новую искомую функцию $p = dy/dx$, а за независимую переменную примем y , тогда

$$y^{(k)} = \Phi_k(y, p, p', \dots, p^{(k-1)}).$$

и, следовательно, уравнение (16) после подстановки относительно новых переменных оказывается уравнением $(n - 1)$ -го порядка

$$F_1(y, p, p^{(1)}, \dots, p^{(n-1)}) = 0.$$

Если его общий интеграл $\Phi(y, p, p_1, \dots, p_{n-1})$ удастся найти, то решение уравнения (16) сводится к решению уравнения

$$\Phi(y, y', p_1, \dots, p_{n-1}) = 0.$$

Пусть уравнение (1) таково, что

$$F(x, y, y', \dots, y^{(n)}) =$$

$$= \frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y} y' + \dots +$$

$$+ \frac{\partial \Phi}{\partial y^{(n-1)}} y^{(n)},$$

где $\Phi = \Phi(x, y, y', \dots, y^{(n-1)})$.

В этом случае уравнение

$$\Phi(x, y, y', \dots, y^{(n-1)}) - c = 0$$

является первым интегралом уравнения (1), и тем самым порядок уравнения понижен на единицу.

Первые интегралы системы обыкновенных дифференциальных уравнений. Решение задачи Коши для системы (4) при начальных условиях (8) является функцией n параметров (начальных данных $y_1^0, y_2^0, \dots, y_n^0$), т. е.

$$y_k = y_k(x, y_1^0, \dots, y_n^0), \quad 1 \leq k \leq n. \quad (17)$$

Пусть $y_j^0 = \omega_j(x, y_1, \dots, y_n)$, $1 \leq j \leq n$ — решение системы (17) относительно y_j^0 , $1 \leq j \leq n$, тогда

$$y_k' = y_k'(x, y_1^0, \dots, y_n^0) =$$

$$= y_k'(x, \omega_1(x, y_1, \dots, y_n),$$

$$\omega_n(x, y_1, \dots, y_n)) =$$

$$= f_k(x, y_1, \dots, y_n).$$

Следовательно, мы восстановили систему уравнений (4) по ее решению (17).

Первый интеграл системы. Функцию $F(x, p_1, \dots, p_n)$, не равную тождественно постоянной, называют **первым интегралом системы**, если для любого решения $Y(x) = (y_1(x), y_n(x))$ системы (4) функция

$$\frac{d}{dx} (F(x, y_1(x), y_2(x), \dots, y_n(x))) \equiv 0,$$

т. е. $F(x, y_1(x), \dots, y_n(x)) = c$, где c — константа.

Общий интеграл системы. Совокупность n первых интегралов $\{F_1(x, p_1, \dots, p_n), \dots, F_n(x, p_1, \dots, p_n)\}$ системы (4) называют **общим интегралом**, если функции F_1, F_2, \dots, F_n функционально независимы. Зная общий интеграл системы (4), ее решения можно получить как решения алгебраической, а не дифференциальной системы уравнений:

$$F_1(x, y_1, \dots, y_n) = c_1, \dots$$

$$F_n(x, y_1, \dots, y_n) = c_n.$$

Замечания: 1. Для любой непрерывной функции $\Phi(\eta_1, \dots, \eta_k)$, $1 \leq k \leq n$ функция

$$\Phi(F_1(x, p_1, \dots, p_n),$$

$$F_2(x, p_1, \dots, p_n), \dots,$$

$$F_k(x, p_1, \dots, p_n))$$

является первым интегралом системы (4), если F_1, \dots, F_k — ее первые интегралы.

2. Первые интегралы могут быть получены как решения системы (17) относительно начальных данных y_1^0, \dots, y_n^0 .

Свойство первого интеграла. Функция $F(x, p_1, \dots, p_n)$ является первым интегралом системы (4) тогда и только тогда, когда она удовлетворяет уравнению

$$\frac{\partial F}{\partial x} + f_1(x, p_1, \dots, p_n) \frac{\partial F}{\partial p_1} + \dots + f_n(x, p_1, \dots, p_n) \frac{\partial F}{\partial p_n} = 0. \quad (18)$$

Применение первых интегралов к понижению размерности системы. Пусть известно $k < n$ первых интегралов системы (4). Решая систему уравнений

$$F_i(x, y_1, \dots, y_k, y_{k+1}, \dots, y_n) = c_i, \quad 1 \leq i \leq k,$$

относительно, например, y_1, \dots, y_k и подставляя найденные решения как функции переменных x, y_{k+1}, \dots, y_n и постоянных c_1, c_2, \dots, c_k , получим систему

$$\frac{dy_j}{dx} = f_j(x, y_{k+1}, \dots, y_n, c_1, \dots, c_k), \quad k+1 \leq j \leq n$$

из $n - k$ уравнений относительно неизвестных функций y_{k+1}, \dots, y_n .

Зависимость решений задачи Коши от начальных данных и параметров. Если функции $f_k(x, p_1, \dots, p_n)$ непрерывны в области

$$U_{a,b} = \{(x, p_1, \dots, p_n) : |x - \bar{x}_0| \leq a, |p_i - \bar{y}_i^0| \leq b; i = 1, 2, \dots, n\}$$

и непрерывно дифференцируемы по переменным p_1, \dots, p_n в $U_{a,b}$, то существует такая область $U'_{a,b} \subset U_{a,b}$, что функции $y_k = y_k(x; x_0, y_1^0, \dots, y_n^0)$, являющиеся решениями задачи Коши

$$\frac{dy_k}{dx} = f_k(x, y_1, \dots, y_n), \quad y_k(x_0) = y_k^0, \quad k = 1, 2, \dots, n,$$

непрерывно дифференцируемы по переменным x_0, y_1^0, \dots, y_n^0 в области $U'_{a,b}$.

Если $M = \max |f_k(x, p_1, \dots, p_n)|$, где максимум берется по всем $1 \leq k \leq n$, $(x, p_1, \dots, p_n) \in U_{a,b}$, то в качестве области $U'_{a,b}$ можно выбрать множество

$$\{(x, p_1, \dots, p_n) : |x - \bar{x}_0| \leq \leq 1/2 \min(a, b/M), |p_k - \bar{y}_i^0| \leq b/2\}.$$

Дифференцируемость решений задачи Коши по параметру.

Если функции f_k в системе (4) зависят от числового параметра λ , т. е. система имеет вид

$$\frac{dy_k}{dx} = f_k(x, y_1, \dots, y_n, \lambda),$$

$$1 \leq k \leq n, \quad (19)$$

то, добавляя к системе (19) уравнение

$$\frac{d\lambda}{dx} = 0$$

и к начальным условиям (8) условие $\lambda(x_0) = \lambda_0$, приходим к утверждению, которое следует из дифференцируемости решений по начальным данным: если функции $f_k(x, y_1, \dots, y_n, \lambda)$ непрерывно дифференцируемы по параметру λ для $\lambda_0 - \Delta \leq \lambda \leq \lambda_0 + \Delta$, то решения системы (19) при начальных условиях (8) допускают непрерывные частные производные по параметру λ .

Оценки для приближенных решений системы обыкновенных дифференциальных уравнений. Пусть вместе с функциями $f_k(x, p_1, \dots, p_n)$ заданы функции $g_k(x, p_1, \dots, p_n)$, $1 \leq k \leq n$, причем как f_k , так и g_k непрерывны и удовлетворяют условию Липшица (10) по переменным p_1, \dots, p_n в некоторой области U . Рассмотрим вместе с системой (4) систему

$$\frac{dz_k}{dx} = g_k(x, z_1, \dots, z_n). \quad (20)$$

Если $Y(x) = (y_1(x), \dots, y_n(x))$ — решение системы (4) при начальных условиях (8), $Z(x) = (z_1(x), \dots, z_n(x))$ — решение системы (20) при начальных условиях $Z_k(x_0) = y_k^0$, $1 \leq k \leq n$, а функции f_k и g_k связаны соотношением

$$\max |f_k(x, p_1, \dots, p_n) - g_k(x, p_1, \dots, p_n)| \leq \varepsilon, \quad 1 \leq k \leq n; (x, p_1, \dots, p_n) \in U,$$

то имеет место оценка

$$\sum_{k=1}^m |y_k(x) - z_k(x)| \leq \frac{\varepsilon}{K} (e^{nK|x-x_0|} - 1) \quad (21)$$

для всех тех x , для которых одновременно существуют решения $Y(x)$ и $Z(x)$.

Замечание. Оценка (21) дает возможность приближенного решения сложных систем дифференциальных уравнений путем замены их соответствующим образом подобранными приближенными уравнениями, решающимися проще.

Пусть функции $f_k(x, p_1, \dots, p_n)$ удовлетворяют условиям существования и единственности решений задачи Коши (4), (8). Вектор-функции $W(x) = (w_1(x), \dots, w_n(x))$ и $V(x) = (v_1(x), \dots, v_n(x))$ являются приближенными решениями системы (4) на интервале $x_0 - h < x < x_0 + h$, т. е.

$$\left| \frac{d}{dx} w_k(x) - f_k(x, w_1(x), \dots, w_n(x)) \right| \leq \varepsilon_1;$$

$$\left| \frac{d}{dx} v_k(x) - f_k(x, v_1(x), \dots, v_n(x)) \right| \leq \varepsilon_2, \quad 1 \leq k \leq n,$$

и пусть известны значения вектор-функции $W(x)$ в точке $x = x_0$, $w_k = w_k(x_0)$ и вектор-функции $V(x)$ в точке $x = \bar{x}_0$, $v_k = v_k(\bar{x}_0)$, тогда для всех $x \in (x_0 - h, x_0 + h)$ имеет место оценка

$$\sum_{k=1}^n |w_k(x) - v_k(x)| \leq \frac{\varepsilon_1 + \varepsilon_2}{K} \times (e^{nK|x-x_0|} - 1) + \left\{ n(M + \varepsilon_1 + \varepsilon_2) |\bar{x}_0 - x_0| + \sum_{k=1}^n (v_k - w_k) \right\} \exp(nK|x - x_0|). \quad (22)$$

Замечание. Если $V(x)$ — точное решение задачи Коши (4), (8), то оценка (22) позволяет установить близость точного и приближенного решений.

Кроме того, оценка (22) позволяет оценить погрешности, вносимые в решение при неточном задании начальных условий. Эта оценка полезна при разработке численных методов решения систем дифференциальных уравнений.

Линейные дифференциальные уравнения. Свойства линейного дифференциального уравнения. Рассмотрим линейное дифференциальное уравнение

$$y^{(n)} + p_1 y^{(n-1)} + \dots + p_n y = q(x), \quad (23)$$

где $p_k = p_k(x)$, $1 \leq k \leq n$, $q(x)$ — непрерывные функции на интервале $a < x < b$.

Решение задачи Коши для уравнения (23) при начальных условиях

$$y(x_0) = y_0, \dots, y^{(n-1)}(x_0) = y_{n-1}, \quad a < x_1 < b$$

существует на всем интервале $a < x < b$ для любых значений y_0, y_{n-1} .

Для линейного уравнения (23) решение существует на всей области непрерывности коэффициентов и правой части $q(x)$ и непрерывно дифференцируемо по начальным данным.

Инвариантность по отношению к замене переменных. При замене переменной x на новую переменную формулой $x = \varphi(\xi)$ уравнение (23) остается линейным того же вида, если $\varphi(\xi)$ — n раз непрерывно дифференцируемая функция и $\varphi'(\xi)$ не обращается в ноль, т. е. существует функция $\psi(x)$, обратная $\varphi(\xi)$ при $a < x < b$. При замене функции $y(x)$ на новую функцию $\eta(x)$, связанную с $y(x)$ формулой $y(x) = v(x)\eta(x) + w(x)$, уравнение (23) переходит в аналогичное уравнение для $\eta(x)$, если функции $v(x)$ и $w(x)$ n раз непрерывно дифференцируемы и $v^2(x) > 0$ для всех $a < x < b$.

Композиция решений однородного линейного уравнения. Рассмотрим линейное уравнение без правой части

$$L[y] = y^{(n)} + p_1 y^{(n-1)} + \dots + p_n y = 0. \quad (24)$$

Если $y_1(x), \dots, y_k(x)$ — решения уравнения (24), то функция

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + \dots + c_k y_k(x) \quad (25)$$

также является решением уравнения (24) при любых значениях постоянных c_1, c_2, \dots, c_k .

Совокупность линейно независимых функций $y_1(x), y_2(x), \dots, y_k(x)$ называют фундаментальной системой решений уравнения (24), если любое решение этого уравнения может быть представлено в виде (25).

Фундаментальную систему называют **нормальной фундаментальной системой** (в точке $x = x_0$), если $y_k^{(j)}(x_0) = \delta_k^{j+1}$, где δ_k^j — символ Кронекера, $0 \leq j \leq n-1$, $1 \leq k \leq n$.

Критерий линейной независимости решений линейного дифференциального уравнения. Пусть $y_1(x), y_2(x), \dots, y_n(x)$ — n решений уравнения (24). Функцию

$$W[y_1, \dots, y_n] = W(x) = \det \|y_k^{(j)}\| = \begin{vmatrix} y_1(x) & y_n(x) \\ y_1'(x) & y_n'(x) \\ \dots & \dots \\ y_1^{(n-1)}(x) & y_n^{(n-1)}(x) \end{vmatrix},$$

$$1 \leq k \leq n, \quad 0 \leq j \leq n-1,$$

называют определителем Вронского или вронскианом системы функций $y_1(x), \dots, y_n(x)$.

Определитель Вронского, составленный для n решений уравнения (24), или тождественно равен нулю, и тогда и только тогда решения линейно зависимы, или не обращается в ноль ни в одной точке интервала $a < x < b$, и тогда и только тогда решения $y_1(x), y_2(x), \dots, y_n(x)$ линейно независимы.

Критерий фундаментальности системы решений уравнения. Если $y_1(x), \dots, y_n(x)$ — решения уравнения (24) и $W^2[y_1, \dots, y_n] > 0$ для всех $a < x < b$, то функции $y_1(x), \dots, y_n(x)$ являются фундаментальной системой решений уравнения (24).

Связь между фундаментальной системой и коэффициентами дифференциального уравнения. Пусть $y_1(x)$

.., $y_n(x)$ — фундаментальная система решений уравнения (24), тогда

$$p_{n-k}(x) = (-1)^{n-k} \times \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1^{(k-1)} & y_2^{(k-1)} & \dots & y_n^{(k-1)} \\ \dots & \dots & \dots & \dots \\ y_1^{(k+1)} & y_2^{(k+1)} & \dots & y_n^{(k+1)} \\ \dots & \dots & \dots & \dots \\ y_1^{(n)} & y_2^{(n)} & \dots & y_n^{(n)} \end{vmatrix} / W[y_1, \dots, y_n].$$

При $k = n - 1$ получим

$$p_1(x) = - \frac{W'[y_1, \dots, y_n]}{W[y_1, \dots, y_n]}.$$

Таким образом, определитель Вронского вычисляется через свое значение при $x = x_0$ формулой

$$W(x) = W(x_0) e^{-\int_{x_0}^x p_1(\xi) d\xi} \quad (26)$$

Построение решений задачи Коши для линейного однородного дифференциального уравнения. Пусть $y_1(x), \dots, y_n(x)$ — нормальная фундаментальная система решений уравнений (24) и пусть заданы начальные условия $y^{(k)}(x_0) = y_k, 0 \leq k \leq n - 1$, тогда функция

$$y(x) = y_0 y_1(x) + y_1 y_2(x) + \dots + y_{n-1} y_n(x)$$

является решением задачи Коши для уравнения (24).

Существование фундаментальной системы решений. Для линейного однородного дифференциального уравнения существует фундаментальная система решений.

Для любого линейного дифференциального уравнения существует нормальная фундаментальная система решений. В самом деле, пусть $y_1(x), \dots, y_n(x)$ — фундаментальная система. Построим функцию $z_k(x) = c_1^{(k)} y_1(x) + \dots + c_n^{(k)} y_n(x)$ так, чтобы $z_k^{(j-1)}(x_0) = \delta_k^j, 1 \leq k, j \leq n$. Для этого необходимо решить систему

линейных алгебраических уравнений относительно c_j^k :

$$c_1^{(k)} y_1^{(j-1)}(x_0) + c_2^{(k)} y_2^{(j-1)}(x_0) + \dots + c_n^{(k)} y_n^{(j-1)}(x_0) = \delta_k^j,$$

которая разрешима ввиду того, что ее определителем является $W(x_0) \neq 0$.

Общее решение неоднородного линейного уравнения. Общее решение уравнения (23) является суммой его частного решения и общего решения однородного уравнения (24). Таким образом, если $y_1(x), \dots, y_n(x)$ — фундаментальная система решений уравнения (24) и $\bar{y}(x)$ — какое-либо решение уравнения (23), то общее решение уравнения (23) определяется формулой

$$y(x) = c_1 y_1(x) + \dots + c_n y_n(x) + \bar{y}(x), \quad (27)$$

где c_1, c_2, \dots, c_n — произвольные постоянные.

Функция $\Phi(x, y, p_1, \dots, p_n) = y - [p_1 y_1(x) + \dots + p_n y_n(x) + \bar{y}(x)]$ является общим интегралом уравнения (23).

Методы построения решений неоднородного линейного дифференциального уравнения.

а) Метод вариации постоянных

Пусть $y_1(x), \dots, y_n(x)$ — фундаментальная система решений уравнения (24). Будем искать решение уравнения (23) в виде

$$y(x) = c_1(x) y_1(x) + \dots + c_n(x) \times y_n(x), \quad (28)$$

где $c_i(x)$ — неизвестные функции.

Составим систему уравнений

$$y_1^{(k)}(x) \frac{dc_1}{dx} + y_2^{(k)}(x) \frac{dc_2}{dx} + \dots + y_n^{(k)}(x) \frac{dc_n}{dx} = 0, \quad 0 \leq k \leq n - 2;$$

$$y_1^{(n-1)}(x) \frac{dc_1}{dx} + y_2^{(n-1)}(x) \frac{dc_2}{dx} + \dots + y_n^{(n-1)}(x) \frac{dc_n}{dx} = q(x). \quad (29)$$

Система (29) линейна относительно неизвестных функций $\frac{dc_i}{dx}$, определителем которой является вронскиан $W[y_1, \dots, y_n] \neq 0$. Следовательно, система (29) разрешима и $\frac{dc_i}{dx} = \varphi_i(x)$, где $\varphi_i(x)$ — известная функция. Таким образом, $c_i(x) = \int \varphi_i(x) dx + \bar{c}_i$, \bar{c}_i — константа. Подставив полученные функции в (28), общее решение получим в виде

$$y(x) = \sum_{i=1}^n \bar{c}_i y_i(x) + \sum_{i=1}^n y_i(x) \int \varphi_i(x) dx.$$

б) Метод Коши

Пусть $y_1(x|\xi), \dots, y_n(x|\xi)$ — нормальная фундаментальная система решений уравнения (24) в точке $x = \xi$, т. е. $y_k^{(j-1)}(\xi|\xi) = \delta_k^j$, $1 \leq k, j \leq n$. Тогда функция

$$y(x) = \int_{x_0}^x y_n(x|\xi) q(\xi) d\xi$$

является частным решением уравнения (23), удовлетворяющим нулевым начальным условиям.

в) Метод фундаментального решения

Фундаментальным решением уравнения (24) называют функцию $G(x, \xi)$, определенную при $a \leq x, \xi \leq b$ и обладающую свойствами:

в областях $a \leq x \leq \xi \leq b$ и $a \leq \xi \leq x \leq b$ функция $G(x, \xi)$ имеет частные производные по x до n -го порядка включительно;

в каждой из этих областей функция $G(x, \xi)$ как функция переменной x удовлетворяет уравнению (24);

в области $a \leq x, \xi \leq b$ функция $G(x, \xi)$ имеет непрерывные по x и ξ частные производные, а по x до $n - 2$ порядка включительно;

при $a < \xi < b$

$$\frac{\partial^{n-1}}{\partial x^{n-1}} G(\xi + 0, \xi) - \frac{\partial^{n-1}}{\partial x^{n-1}} G(\xi - 0, \xi) = 1.$$

Фундаментальным решением является, например, функция

$$G_0(x, \xi) = \frac{\text{sign}(x - \xi)}{2W(\xi)} \times \begin{vmatrix} y_1(\xi) & y_n(\xi) \\ y_1^{(n-2)}(\xi) & y_n^{(n-2)}(\xi) \\ y_1(x) & y_n(x) \end{vmatrix},$$

которая удовлетворяет соотношениям

$$G_0(\xi, \xi) = G'_x(\xi, \xi) = \dots = G_x^{(n-2)}(\xi, \xi) = 0.$$

Совокупность всех фундаментальных решений определяется формулой

$$G(x, \xi) = G_0(x, \xi) + c_1(\xi) y_1(x) + \dots + c_r(\xi) y_n(x),$$

где $c_i(\xi)$ — произвольные непрерывные функции, а y_1, \dots, y_n — фундаментальная система решений уравнения (24).

Если $G(x, \xi)$ — фундаментальное решение уравнения (24) [функция Грина уравнения (24)], то функция

$$y(x) = \int_a^b G(x, \xi) f(\xi) d\xi$$

является решением уравнения (23).

Фундаментальная система решений линейного дифференциального уравнения с постоянными коэффициентами. Рассмотрим вместе с уравнением

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = 0, \quad (30)$$

где a_1, \dots, a_n — числа, полином

$$Q_n(p) = p^n + a_1 p^{n-1} + \dots + a_n,$$

который называется характеристическим.

Пусть $p_1, \dots, p_m, m \leq n$ — корни этого полинома кратности соответственно $s_1, s_2, \dots, s_m, s_1 + s_2 + \dots + s_m = n$. Пусть $p_k = \alpha_k + i\beta_k$, где $i = \sqrt{-1}$, а α_k, β_k — вещественные числа.

Семейство функций

$$\{e^{(\alpha_k + i\beta_k)x}, xe^{(\alpha_k + i\beta_k)x}, \dots, x^{s_k-1} e^{(\alpha_k + i\beta_k)x}\}, \quad 1 \leq k \leq m,$$

является фундаментальной системой решений уравнения (30). Если числа a_1, \dots, a_n — вещественные, то фундаментальной системой решений являются функции

$$e^{\alpha_k x} \cos \beta_k x, xe^{\alpha_k x} \cos \beta_k x,$$

$$x^{s_k-1} e^{\alpha_k x} \cos \beta_k x;$$

$$e^{\alpha_k x} \sin \beta_k x, xe^{\alpha_k x} \sin \beta_k x,$$

$$x^{s_k-1} e^{\alpha_k x} \sin \beta_k x,$$

где мы берем только те k , для которых $\beta_k \geq 0$.

Приведение дифференциального уравнения второго порядка к простейшему виду.

Уравнение

$$y'' + p_1(x)y' + p_2(x)y = 0 \quad (31)$$

с помощью замены

$$y = ze^{-\int p_1(x) dx}$$

приводится к уравнению вида

$$z'' + Q(x)z = 0, \quad (32)$$

где

$$Q(x) = p_2(x) - (1/4)p_1^2(x) - (1/2)p_1'(x).$$

Любое уравнение, которое может быть получено из уравнения (31) посредством замены $y = u(x)z$, где $u(x)$ — заданная функция, приводятся к одному и тому же уравнению (32).

Свойства решений дифференциального уравнения второго порядка.

Свойство нулей фундаментальной системы решений. Если $y_1(x)$ и $y_2(x)$ — фундаментальная система решений уравнений (31), то между каждыми двумя соседними нулями одного ре-

шения (если такие существуют) лежит один и только один нуль другого решения, т. е. нули двух линейно независимых решений взаимно разделяют друг друга.

Сравнение характера колебаний решений двух уравнений. Если в интервале $a < x < b$ $Q_2(x) \geq Q_1(x)$, то между каждыми двумя нулями любого решения уравнения (32) с $Q(x) = Q_1(x)$ находится по крайней мере один нуль каждого решения уравнения (32) с $Q(x) = Q_2(x)$. Пусть $0 < m \leq Q(x) \leq M$ в интервале (a, b) , тогда данное утверждение приводит к следующему результату: расстояние Δ между двумя последовательными нулями любого решения уравнения (32) удовлетворяет неравенству

$$\frac{\pi}{\sqrt{M}} \leq \Delta \leq \frac{\pi}{\sqrt{m}}.$$

Поведение решений при $|x| \rightarrow \infty$. Если $Q(x) \rightarrow -\infty$ при $|x| \rightarrow \infty$, то каждое нетривиальное решение уравнения (32) имеет лишь конечное число нулей и

$$\left| \frac{z'(x)}{z(x)} \right| \rightarrow \infty \text{ при } |x| \rightarrow \infty;$$

существует два линейно независимых решения $z_1(x)$ и $z_2(x)$, таких, что либо

$$z_1, z_1' \rightarrow 0, z_2 \rightarrow \infty, z_2' \rightarrow -\infty$$

при $x \rightarrow -\infty$,

либо

$$z_1, z_1' \rightarrow 0, z_2, z_2' \rightarrow +\infty,$$

при $x \rightarrow +\infty$.

Если $Q(x) \geq \alpha^2 > 0$, то каждое решение уравнения (32), как и его производные, имеет бесконечное число нулей. Если $Q(x) = \lambda + q(x)$ и если $q(x) = o(x^{-k})$, $k > 1$, при $x \rightarrow +\infty$, то при каждом $\lambda > 0$ решение $z(x)$ уравнения (32) может быть представлено в виде

$$z(x) = \rho(x) \sin(\lambda x + \sigma(x)),$$

где $\rho(x)$, $\sigma(x)$ — дифференцируемые функции, а при соответствующем выборе ρ_0 и σ_0

$$\rho(x) = \rho_0 + o(1/x^{k-1}),$$

$$\sigma = \sigma_0 + o(1/x^{k-1}).$$

Системы линейных дифференциальных уравнений. Систему (4) называют **линейной**, если

$$f_k(x, \rho_1, \dots, \rho_n) = a_{k1}(x)\rho_1 + \dots + a_{kn}(x)\rho_n + q_k(x), \quad 1 \leq k \leq n. \quad (33)$$

Систему (33) называют **однородной**, если $q_k(x) \equiv 0$, $1 \leq k \leq n$, и **неоднородной** в противном случае. Таким образом, система (4) принимает вид

$$\frac{dy_k}{dx} = a_{k1}(x)y_1 + \dots + a_{kn}(x)y_n + q_k(x), \quad 1 \leq k \leq n. \quad (34)$$

Существование и единственность решения задачи Коши для линейной системы. Если функции $q_k(x)$, $a_{km}(x)$, $1 \leq k, m \leq n$ — непрерывны на $[a, b]$, то решение задачи Коши (34), (8) существует и единственно для любых y_1^0, \dots, y_n^0 , $x_0 \in (a, b)$ на всем интервале (a, b) .

Фундаментальная система решений однородной линейной системы уравнений. Совокупность решений $Y^{(1)}(x) = (y_1^1(x), \dots, y_n^1(x))$, $Y^{(k)}(x) = (y_1^k(x), \dots, y_n^k(x))$ однородной системы (34) называют **фундаментальной системой решений**, если решения $Y^{(1)}$, $Y^{(2)}$, \dots , $Y^{(n)}$ линейно независимы и любое решение $Y(x)$ системы (34) представимо в виде линейной комбинации решений $Y^{(1)}$, $Y^{(2)}$, \dots , $Y^{(n)}$.

Если $\bar{Y}(x) = (\bar{y}_1(x), \dots, \bar{y}_n(x))$ и $\bar{\bar{Y}}(x) = (\bar{\bar{y}}_1(x), \dots, \bar{\bar{y}}_n(x))$ — решения однородной системы (34), то вектор-функция

$$Y(x) = \alpha \bar{Y}(x) + \beta \bar{\bar{Y}}(x)$$

также является решением однородной системы (34) при любых $\alpha, \beta \in \mathbb{C}$.

Если $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — n -линейно независимые решения однородной системы (34), то для любых $a < x < b$ определитель

$$D(x) = D[Y^{(1)}, \dots, Y^{(n)}] = \begin{vmatrix} y_1^1(x) & \dots & y_1^n(x) \\ \dots & \dots & \dots \\ y_n^1(x) & \dots & y_n^n(x) \end{vmatrix} \neq 0. \quad (35)$$

Точнее, если $a < x_0$, $x < b$, то $D(x) = D(x_0) \exp \times$

$$\times \left[- \int_{x_0}^x (a_{11}(\xi) + a_{22}(\xi) + \dots + a_{nn}(\xi)) d\xi \right].$$

Из (35) следует, что если $Y^{(1)}(x_0), \dots, Y^{(n)}(x_0)$ линейно независимы, то для любого $a < x < b$ $Y^{(1)}(x), \dots, Y^{(n)}(x)$ также линейно независимые векторы.

Любые $(n+1)$ решения однородной системы (34) линейно зависимы.

Если функции $a_{km}(x)$, $1 \leq k, m \leq n$ непрерывны, то существует n линейно независимых решений системы (34).

Из сформулированных утверждений следует: любая линейная однородная система (34) обладает фундаментальной системой решений и такой системой является совокупность любых n линейно независимых решений.

Решение задачи Коши для однородной линейной системы. Пусть дана задача Коши (34), (8) и $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — фундаментальная система решений однородной системы (34). Ищем решение $Y(x)$ задачи Коши в виде $Y(x) = c_1 Y^{(1)}(x) + \dots + c_n Y^{(n)}(x)$, где $Y(x)$ — решение однородной системы (34). Найдем c_1, \dots, c_n из системы:

$$c_1 Y^{(1)}(x_0) + \dots + c_n Y^{(n)}(x_0) = (y_1^0, \dots, y_n^0). \quad (36)$$

Определитель системы (36) — это $D(x_0)$, определяемый формулой (35). Так как $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — фундаментальная система решений, то система (36) при любых y_1^0, \dots, y_n^0 имеет решение, и оно единственно.

Если $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — нормальная фундаментальная система решений в точке $x = x_0$, т. е. $Y^{(k)}(x_0) = e_k$, $e_k = (\delta_1^k, \dots, \delta_n^k)$, то решение задачи Коши представимо в виде

$$Y(x) = y_1^0 Y^{(1)}(x) + \dots + y_n^0 Y^{(n)}(x).$$

Для каждой точки x_0 существует нормальная фундаментальная система решений. Ее можно построить, полагая $Y_0 = (y_1^0, \dots, y_n^0) = e_k$ и решая n раз систему (36).

Общее решение неоднородной линейной системы. Если $\bar{Y}(x)$ — частное решение системы (34), то любое решение $Y(x)$ этой системы представимо в виде

$$Y(x) = c_1 Y^{(1)}(x) + \dots + c_n Y^{(n)}(x) + \bar{Y}(x), \quad (37)$$

где $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — фундаментальная система решений однородной системы (34), а c_1, \dots, c_n — постоянные.

Если $\bar{Y}(x_0) = 0$, то решение задачи Коши (34), (8), представимое в виде (37), может быть найдено посредством решения системы (36).

Построение частного решения неоднородной линейной системы методом вариации постоянных. Будем искать решение задачи Коши для уравнения (34) с нулевыми начальными условиями в виде

$$\bar{Y}(x) = c_1(x) Y^{(1)}(x) + \dots + c_n(x) Y^{(n)}(x), \quad (38)$$

где $c_1(x), \dots, c_n(x)$ — неизвестные функции, а $Y^{(1)}(x), \dots, Y^{(n)}(x)$ — фундаментальная система решений однородной системы (34).

Подставляя (38) в (34), видим, что если $c_k(x)$, $1 \leq k \leq n$ будут удовлетворять системе

$$\frac{dc_1}{dx} Y^{(1)}(x) + \dots + \frac{dc_n}{dx} Y^{(n)}(x) = Q(x) = (q_1(x), \dots, q_n(x))^T \quad (39)$$

и начальным условиям $c_k(x_0) = 0$, то $\bar{Y}(x)$ будет искомым решением. Си-

стема (39) разрешима для всех $a < x < b$, и, следовательно, $c_k(x)$ является решением задачи Коши

$$\frac{dc_k}{dx} = \varphi_k(x), \quad c_k(x_0) = 0, \quad 1 \leq k \leq n, \quad (40)$$

где $\varphi_k(x) = D_k(x)/D(x)$; здесь $D(x)$ определяется формулой (35), а $D_k(x)$ получается из определителя $D(x)$ заменой k -го столбца на столбец $(q_1(x), \dots, q_n(x))^T$.

Из (40) очевидно, что

$$c_k(x) = \int_{x_0}^x \varphi_k(\xi) d\xi,$$

и, следовательно,

$$\bar{Y}(x) = Y^{(1)}(x) \int_{x_0}^x \frac{D_1(\xi)}{D(\xi)} d\xi + \dots + Y^{(n)}(x) \int_{x_0}^x \frac{D_n(\xi)}{D(\xi)} d\xi. \quad (41)$$

Из (41) следует, что, имея фундаментальную систему решений однородной системы (34), можно построить любое решение неоднородной системы свести к вычислению интегралов от известных функций [определяется формулами (41) и (37)].

Системы линейных уравнений с постоянными коэффициентами. Если в системе (34) все функции $a_{km}(x)$, $1 \leq k, m \leq n$ суть постоянные, то систему называют линейной системой с постоянными коэффициентами.

Для построения общего решения линейной системы достаточно найти фундаментальную систему ее решений. Для линейной системы с постоянными коэффициентами задачи нахождения ее фундаментальной системы решений оказывается чисто алгебраической и эквивалентна приведению матрицы $A = \|a_{ij}\|$ к нормальной жордановой форме.

Пусть e_k , $1 \leq k \leq n$, — базис, в котором матрица A имеет вид, определяемый уравнениями (34), а x_k , $1 \leq k \leq n$, — базис Жордана, в кото-

ром матрица A приводится к блочно-диагональному виду

$$A = \left\| \begin{array}{ccc} \boxed{M_1} & & 0 \\ & \boxed{M_2} & \\ 0 & & \dots \\ & & & \boxed{M_n} \end{array} \right\|, \quad (42)$$

где вне блоков все элементы нулевые, а каждый блок есть квадратная матрица размера $s_j \times s_j$ вида

$$M_j = \left\| \begin{array}{ccc} \lambda_j & & 0 \\ & \ddots & \\ & & 1 \\ 0 & & & \lambda_j \end{array} \right\|.$$

Числа λ_j суть различные корни характеристического уравнения $\det(A - \lambda I) = 0$ кратности s_j ,

$$I = \|\delta_j^i\|, \quad s_1 + s_2 + \dots + s_k = n.$$

Решение однородной системы (34) с постоянными коэффициентами представимо в виде

$$Y(x) = e^{Ax} Y_0,$$

где матрица $e^{Ax} = \sum_{k=0}^{\infty} (Ax)^k / k!$, $Y_0 =$

$$= (y_1^0, \dots, y_n^0).$$

Если A имеет вид (42), то

$$e^{Ax} = \left\| \begin{array}{ccc} e^{M_1 x} & & 0 \\ & e^{M_2 x} & \\ 0 & & \dots \\ & & & e^{M_n x} \end{array} \right\|,$$

$$\text{где } e^{M_j x} = \left\| \begin{array}{cccc} 1 & x & x^2/2 & \dots & \frac{x^{s_j-1}}{(s_j-1)!} \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & x \\ 0 & \dots & \dots & \dots & 1 \end{array} \right\| e^{\lambda_j x}.$$

Фундаментальной системой решений системы (34) в базисе κ_k , $1 \leq k \leq n$, являются столбцы матрицы e^{Ax} . Обозначим эти решения $Z^{(m')}(x) = (z_1^{(m')}(x), \dots, z_n^{(m')}(x))$; здесь $(z_1^{(m')}(x), \dots, z_n^{(m')}(x))$ — m' -й столбец матрицы e^{Ax} в жордановом ба-

зисе, $1 \leq m' \leq n$. Если $B = \|b_{ij}\|$, $1 \leq i, j \leq n$ — матрица, преобразующая координаты вектора из базиса κ_k , $1 \leq k \leq n$, в базис e_k , $1 \leq k \leq n$, то фундаментальной системой решений уравнения (34) в базисе e_k , $1 \leq k \leq n$, будут функции

$$Y^{(m)}(x) = \begin{pmatrix} y_1^{(m)}(x) \\ \vdots \\ y_n^{(m)}(x) \end{pmatrix} = \begin{pmatrix} b_{11} & b_{1n} \\ b_{n1} & b_{nn} \end{pmatrix} \begin{pmatrix} z_1^{(m)}(x) \\ \vdots \\ z_n^{(m)}(x) \end{pmatrix}. \quad (43)$$

Вектор-функцию $Y^{(m)}(x)$ можно найти, не вычисляя заранее матрицу B . Вычислив λ_j и s_j , фундаментальное решение будем искать в виде

$$Y_{j,r}(x) = e^{\lambda_j x} c_1 + x e^{\lambda_j x} c_2 + \dots + \frac{x^r}{r!} e^{\lambda_j x} c_r, \quad 1 \leq j \leq k, \quad (44)$$

$$0 \leq r \leq s_j - 1,$$

где c_1, c_2, \dots, c_r — неизвестные постоянные векторы, которые могут быть найдены как нетривиальные решения цепочки уравнений

$$(A - \lambda_j) c_r = 0, \quad (A - \lambda_j) c_{r-1} = c_r, \dots; \quad (A - \lambda_j) c_1 = c_2.$$

Устойчивость решений систем уравнений. Решение $\bar{Y}(x) = (\bar{y}_1(x), \dots, \bar{y}_n(x))$ системы (4), удовлетворяющее начальным условиям $\bar{y}_k(x_0) = \bar{y}_k^0$, называют устойчивым по Ляпунову, если для любого $\epsilon > 0$ найдется такое $\delta(\epsilon) > 0$, что любое решение $Y(x)$ системы (4) с начальными условиями $y_k(x_0) = y_k^0$ удовлетворяет для всех $x \geq x_0$ неравенству

$$|\bar{y}_k(x) - y_k(x)| < \epsilon, \quad 1 \leq k \leq n, \quad \text{при } |\bar{Y}_k^0 - y_k^0| < \delta(\epsilon), \quad 1 \leq k \leq n.$$

Решение $\bar{Y}(x)$ называют асимптотически устойчивым, если оно устойчиво и

$$\lim_{x \rightarrow \infty} (y_k(x) - \bar{y}_k(x)) = 0, \quad 1 \leq k \leq n.$$

Устойчивость по Ляпунову называют устойчивостью по начальным данным.

Достаточные условия устойчивости.

Пусть $\bar{Y}(x) = (\bar{y}_1(x), \dots, \bar{y}_n(x))$ — решение системы (4), удовлетворяющее начальным условиям $\bar{y}_k(x_0) = \bar{y}_k^0$, $1 \leq k \leq n$. Предположим, что имеет место разложение

$$\begin{aligned} f_k(x, \bar{y}_1(x) + w_1, \dots, \bar{y}_n(x) + w_n) = \\ = f_k(x, \bar{y}_1(x), \dots, \bar{y}_n(x) + \\ + \sum_{m=1}^n a_{km}(x) w_m + \varphi_k(x, w_1, \dots, w_n), \\ 1 \leq k \leq n. \end{aligned} \quad (45)$$

Решение $\bar{Y}(x)$ устойчиво по Ляпунову, если:

существуют $\sigma > 0$, $M > 0$, что

$$|\varphi_k(x, w_1, \dots, w_n)| \leq M(w_1^2 + \dots + w_n^2)^{1/2 + \delta}, \quad 1 \leq k \leq n;$$

существует такое $q > 0$, что для каждого $x \geq x_0$ любой корень λ уравнения $\det \|a_{km} - \lambda \delta_k^m\|$ удовлетворяет условию $\operatorname{Re} \lambda(x) \leq -q < 0$, где a_{km} и φ_k — функции, входящие в разложение (45).

Краевые задачи для линейных дифференциальных уравнений. *Постановка краевых задач.* Пусть на интервале $a \leq x \leq b$ уравнение

$$L_n[y] = y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = 0 \quad (46)$$

имеет фундаментальную систему решений $y_1(x), y_2(x), \dots, y_n(x)$. Пусть $S_j(q_1, \dots, q_n, p_1, \dots, p_n)$, $1 \leq j \leq m$ — линейные функции переменных $q_1, \dots, q_n, p_1, \dots, p_n$, т. е.

$$\begin{aligned} S_j(q_1, \dots, q_n, p_1, \dots, p_n) = \\ = \sum_{k=1}^n (a_{kj}q_k + b_{kj}p_k). \end{aligned}$$

Краевой задачей для уравнения (46) является следующая задача: найти на интервале $a \leq x \leq b$ решение

уравнения (46), удовлетворяющее условиям: $1 \leq j \leq m$,

$$\begin{aligned} S_j[y] = S_j(y(a), \dots, y^{(n-1)}(a), \\ y(b), \dots, y^{(n-1)}(b)) = 0. \end{aligned} \quad (47)$$

Условия (47) называют краевыми условиями.

Существование решений краевой задачи. Если ранг матрицы

$$S = \|S_j[y_k]\|, \quad 1 \leq j \leq m, \quad 1 \leq k \leq n$$

равен r , то краевая задача (46), (47) имеет ровно $n - r$ линейно независимых решений.

Ранг матрицы S называют рангом краевой задачи.

Собственные значения и собственные функции краевой задачи. Рассмотрим краевую задачу

$$L[y] - \lambda y = 0, \quad S_j[y] = 0, \quad 1 \leq j \leq m, \quad (48)$$

где λ — числовой параметр.

Число λ называют собственным значением краевой задачи (48), если существует решение $y(x) \not\equiv 0$ этой краевой задачи, которое называют собственной функцией. Пусть $y_{1,\lambda}(x), \dots, y_{n,\lambda}(x)$ — фундаментальная система решения уравнения

$$L[y] - \lambda y = 0.$$

Рассмотрим определитель

$$\Delta(\lambda) = \det \|S_j[y_{i,\lambda}]\|, \quad 1 \leq j \leq m, \quad 1 \leq i \leq n.$$

Собственные числа краевой задачи (48) суть нули функции $\Delta(\lambda)$. Если $\Delta(\lambda)$ тождественно равна нулю, то любое λ — собственное значение краевой задачи. Если $\Delta(\lambda)$ не обращается в нуль тождественно, то $\Delta(\lambda)$ — целая аналитическая функция параметра λ и имеет не более счетного числа нулей, которые не имеют конечной предельной точки. Совокупность всех собственных значений функций, отвечающих одному и тому же собственному значению, есть конечномерное пространство размерности не более n .

2. УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ ПЕРВОГО ПОРЯДКА

Основные определения. Уравнением в частных производных первого порядка относительно неизвестной функции $u(x_1, \dots, x_n)$ независимых переменных x_1, \dots, x_n называется выражение вида

$$F(x_1, \dots, x_n, u, u'_{x_1}, \dots, u'_{x_n}) = 0, \quad (49)$$

где $F(x_1, \dots, x_n, u, p_1, \dots, p_n)$ — заданная функция переменных $x_1, \dots, x_n, u, p_1, \dots, p_n$, а $u'_{x_i} = \partial u / \partial x_i$, $1 \leq i \leq n$.

Уравнение (49) называют квазилинейным, если

$$\begin{aligned} F(x_1, \dots, x_n, u, p_1, \dots, p_n) = \\ = f_1(x_1, \dots, x_n, u) p_1 + \dots + \\ + f_n(x_1, \dots, x_n, u) p_n - \\ - f(x_1, \dots, x_n, u). \end{aligned} \quad (50)$$

Квазилинейное уравнение (49) называют линейным, если функции $f_k(x_1, \dots, x_n, u)$, $1 \leq k \leq n$, не зависят от u , а $f(x_1, \dots, x_n, u) \equiv 0$. Функцию $u = u(x_1, \dots, x_n)$ называют решением уравнения (49) в области $\Omega \subset R^n$, если функция

$$\begin{aligned} F(x_1, \dots, x_n, u(x_1, \dots, x_n), \\ \frac{\partial u}{\partial x_1}(x_1, \dots, x_n), \\ \dots, \frac{\partial u}{\partial x_n}(x_1, \dots, x_n)) \equiv 0 \end{aligned}$$

как функция переменных x_1, \dots, x_n , как только $(x_1, \dots, x_n) \in \Omega$.

Задача Коши для уравнения в частных производных первого порядка. Пусть Γ — $(n-1)$ -мерная поверхность в n -мерном пространстве R^n , заданная уравнениями

$$\begin{aligned} x_i = x_i(\alpha_1, \dots, \alpha_{n-1}), \quad 1 \leq i \leq n, \\ (\alpha_1, \dots, \alpha_{n-1}) \in R^{n-1}. \end{aligned} \quad (51)$$

Задачей Коши для уравнения (51) называют следующую задачу: найти решение $u = u(x_1, \dots, x_n)$ уравнения (49), удовлетворяющее условию

$$\begin{aligned} u(x_1(\alpha_1, \dots, \alpha_{n-1}), \\ x_n(\alpha_1, \dots, \alpha_{n-1})) = \varphi(\alpha_1, \dots, \alpha_{n-1}), \end{aligned} \quad (52)$$

где $\varphi(\alpha_1, \dots, \alpha_{n-1})$ — заданная функция переменных $\alpha_1, \alpha_2, \dots, \alpha_{n-1}$.

Замечания: 1. Условие (52) означает, что решение $u(x_1, \dots, x_n)$ уравнения (49) является решением задачи Коши (49), (51), (52), если сужение $u(x_1, \dots, x_n)$ на поверхность Γ совпадает с заданной на Γ функцией φ .

2. Часто поверхность Γ является $(n-1)$ -мерной гиперплоскостью, задаваемой уравнением $x_i = x_i^0$, где x_i^0 — заданное число, т. е. $\Gamma = \{(x_1, \dots, x_n) \in R^n \mid x_i = x_i^0\}$.

Общий метод решения задачи Коши для уравнения в частных производных первого порядка. Построение решения задачи Коши (49), (51), (52) сводится к интегрированию обыкновенных дифференциальных уравнений и решению систем алгебраических уравнений. Пусть $F(x_1, \dots, x_n, u, p_1, \dots, p_n)$ имеет непрерывные производные по всем переменным, а функции

$$\begin{aligned} x_i = x_i(\tau, x_1^0, \dots, x_n^0, u^0, p_1^0, \dots, p_n^0), \\ p_i = p_i(\tau, x_1^0, \dots, x_n^0, u^0, p_1^0, \dots, p_n^0), \\ 1 \leq i \leq n, u = u(\tau, x_1^0, \dots, x_n^0, \\ u^0, p_1^0, \dots, p_n^0) \end{aligned}$$

являются решением следующей задачи Коши:

$$\begin{aligned} d\tau = \frac{dx_1}{\partial F / \partial p_1} = \dots = \frac{dx_n}{\partial F / \partial p_n} = \\ = \frac{du}{\sum_{i=1}^n p_i \partial F / \partial p_i} = \\ = - \frac{dp_1}{\partial F / \partial x_1 + p_1 \partial F / \partial u} = \dots \end{aligned} \quad (53)$$

$$\begin{aligned} x_i(0) = x_i^0, \quad p_i(0) = p_i^0, \quad 1 \leq i \leq n, \\ u(0) = u^0. \end{aligned}$$

Положим

$$x_i^0 = \kappa_i(\alpha_1, \dots, \alpha_{n-1}), \quad 1 \leq i \leq n, \\ u^0 = \varphi(\alpha_1, \dots, \alpha_{n-1}). \quad (54)$$

Пусть $p_i^0 = \psi_i(\alpha_1, \dots, \alpha_{n-1})$ — решение алгебраической системы уравнений

$$F(x_1^0, \dots, x_n^0, u^0, p_1^0, \dots, p_n^0) = 0; \quad (55)$$

$$\frac{\partial \varphi}{\partial \alpha_k} = p_1^0 \frac{\partial x_1^0}{\partial \alpha_k} + \dots + p_n^0 \frac{\partial x_n^0}{\partial \alpha_k}, \\ 1 \leq k \leq n-1,$$

где x_i^0, u^0 определяются формулами (54).

Обозначим

$$\tilde{x}_i(\alpha_1, \dots, \alpha_{n-1}, \tau) = \\ = x_i(\tau, x_1^0, \dots, x_n^0, u^0, p_1^0, \dots, p_n^0); \\ \tilde{u}(\alpha_1, \dots, \alpha_{n-1}, \tau) = \\ = u(\tau, x_1^0, \dots, x_n^0, u^0, p_1^0, \dots, p_n^0).$$

Здесь вместо x_i^0, p_i^0 подставлены соответственно функции $\kappa_i(\alpha_1, \dots, \alpha_{n-1}), \psi_i(\alpha_1, \dots, \alpha_{n-1}), 1 \leq i \leq n$, а вместо u^0 — функция $\varphi(\alpha_1, \dots, \alpha_{n-1})$.

Если функции $\tau(x_1, \dots, x_n), \alpha_k = \alpha_k(x_1, \dots, x_n), 1 \leq k \leq n-1$, являются решением системы алгебраических уравнений

$$x_i = \tilde{x}_i(\alpha_1, \dots, \alpha_{n-1}, \tau), \quad 1 \leq i \leq n, \\ \text{то функция}$$

$$u(x_1, \dots, x_n) = \tilde{u}(\alpha_1(x_1, \dots, x_n), \dots, \\ \alpha_{n-1}(x_1, \dots, x_n), \tau(x_1, \dots, x_n))$$

есть решение задачи Коши (49), (51), (52).

Решение задачи Коши для квазилинейного уравнения. Рассмотрим квазилинейное уравнение

$$f_1(x_1, \dots, x_n, u) \frac{\partial u}{\partial x_1} + \dots + \\ + f_n(x_1, \dots, x_n, u) \frac{\partial u}{\partial x_n} = \\ = f(x_1, \dots, x_n, u). \quad (56)$$

Пусть $\Phi_1(x_1, \dots, x_n, u); \dots; \Phi_n(x_1, \dots, x_n, u)$ — n первых интегралов системы обыкновенных дифференциальных уравнений

$$\frac{dx_1}{f_1(x_1, \dots, x_n, u)} = \dots = \\ = \frac{dx_n}{f_n(x_1, \dots, x_n, u)} = \\ = \frac{du}{f(x_1, \dots, x_n, u)}.$$

Для любой непрерывно дифференцируемой функции $F(y_1, \dots, y_n)$ переменных y_1, \dots, y_n функция $u = u(x_1, \dots, x_n)$, являющаяся решением алгебраического уравнения

$$F(\Phi_1(x_1, \dots, x_n, u), \\ \Phi_n(x_1, \dots, x_n, u)) = 0, \quad (57)$$

является также решением дифференциального уравнения (56). Решение задачи Коши (56), (51), (52) строится следующим образом:

пусть функции

$$\alpha_k = \alpha_k(y_1, \dots, y_n), \quad 1 \leq k \leq n-1, \\ u = \beta(y_1, \dots, y_n)$$

суть решения алгебраической системы уравнений

$$y_k = \Phi_k(\kappa_1(\alpha_1, \dots, \alpha_{n-1}), \dots, \\ \kappa_n(\alpha_1, \dots, \alpha_{n-1}), u), \quad 1 \leq k \leq n.$$

Положим

$$F(y_1, \dots, y_n) = \beta(y_1, \dots, y_n) - \\ - \varphi(\alpha_1(y_1, \dots, y_n), \dots,$$

$$\alpha_{n-1}(y_1, \dots, y_n)).$$

Функция $u = u(x_1, \dots, x_n)$, являющаяся решением алгебраического уравнения

$$F(\Phi_1(x_1, \dots, x_n, u),$$

$$\Phi_n(x_1, \dots, x_n, u)) = 0,$$

является решением задачи Коши (56), (51), (52).

Пример. Рассмотрим уравнение

$$\frac{\partial u}{\partial x_1} + u \frac{\partial u}{\partial x_2} = 1$$

и поставим задачу Коши, положив

$$x_1 = \alpha_1, \quad x_2 = \alpha_1^2, \quad u = \alpha_1.$$

Составим систему

$$\frac{dx_1}{1} = \frac{dx_2}{u} = \frac{du}{1}.$$

Ее два первых интеграла, например, следующие:

$$\Phi_1(x_1, x_2, u) = x_1 - u;$$

$$\Phi_2(x_1, x_2, u) = x_2 - u^2/2.$$

Решая алгебраическую систему

$$y_1 = \alpha_1 - u; \quad y_2 = \alpha_1^2 - u^2/2,$$

получим

$$\beta(y_1, y_2) = 2y_1 \pm \sqrt{2y_1^2 + 2y_2},$$

$$\alpha_1(y_1, y_2) = 3y_1 \pm \sqrt{2y_1^2 + 2y_2}.$$

Так как $\varphi(\alpha_1) = \alpha_1$, то

$$F(y_1, y_2) = \beta(y_1, y_2) - \varphi(\alpha_1(y_1, y_2)) = -y_1,$$

и решение найдем из уравнения

$$F(\Phi_1(x_1, x_2, u), \Phi_2(x_1, x_2, u)) = -\Phi_1(x_1, x_2, u) = u - x_1 = 0.$$

Таким образом, $u = x_1$.

1. ОСНОВНЫЕ ПОНЯТИЯ

Эксперимент. Вероятностный эксперимент. Теория вероятностей как математическая дисциплина имеет предметом своих исследований математические модели, описывающие с определенной степенью адекватности реальные явления. Основу таких моделей составляет мыслимый или реально осуществляемый эксперимент. Эксперимент — научно поставленный опыт, наблюдение исследуемого явления в точно учитываемых условиях, позволяющих следить за ходом явления и воссоздавать его каждый раз при повторении этих условий. Результатом эксперимента является некоторый исход, событие. Эксперимент, в котором нельзя однозначно предсказать исход, называют вероятностным экспериментом.

Примеры:

1. Бросание игральной кости, на гранях которой написаны числа от 1 до 6. Результат эксперимента — выпадение одного из этих чисел.

2. Измерение характеристик орбиты искусственного спутника Земли. Результат эксперимента — значение искомым характеристик.

3. Измерение сроков безотказной работы некоторого изделия (например, электрической лампочки). Результат — число часов, проработанных изделием до отказа.

1.1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ
ВЕРОЯТНОСТНОГО ЭКСПЕРИМЕНТА

Математическая модель вероятностного эксперимента определяется тремя объектами $(\Omega, \mathfrak{A}, P)$, где Ω — множество элементарных событий; \mathfrak{A} — σ — алгебра подмножеств множе-

ства Ω ; P — вероятностная мера, определенная на элементах A σ -алгебры \mathfrak{A} ($A \in \mathfrak{A}$).

Тройку $(\Omega, \mathfrak{A}, P)$ называют вероятностным пространством.

Вероятностное пространство — это измеримое пространство с определенной на нем вероятностной мерой.

Множеством элементарных событий Ω называют множество всех мыслимых исходов эксперимента.

Замечание 1. Как следует из определения, исход (событие) есть результат эксперимента. Однако результат эксперимента может быть сформулирован неоднозначно.

Например, при бросании игральной кости (пример 1), если выпало число 6, то мы можем сказать, что выпало число очков: четное, большее 3, меньшее или равное 6, делящееся на 3, 6 очков.

Поэтому, чтобы подчеркнуть, что элементы ω множества Ω есть элементарные мыслимые исходы (события), заметим, что для любого события A (исхода эксперимента) элементарное событие $\{\omega\}$ либо влечет A ($\omega \in A$), либо не влечет A ($\omega \notin A$). Так как указанное свойство элементарного события ω выполняется для любого события A , то тем самым гарантируется неразложимость множества $\{\omega\}$, т. е. его нельзя представить в виде суммы событий, отличных от \emptyset .

В приведенном примере событие «выпало 6 очков» является элементарным. Остальные события — неэлементарные.

Замечание 2. Для одного и того же физического эксперимента множество элементарных событий может быть определено по-разному в зависимости от целей, для которых создается математическая модель.

Так, в примере 1 эксперимент можно считать экспериментом с шестью элементарными исходами $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$, ω_i — есть выпадение i очков, а можно считать с двумя элементарными исходами $\Omega = \{\omega'_1, \omega'_2\}$, ω'_1 — выпадение нечетного числа очков, ω'_2 — выпадение четного числа очков.

σ -алгебра событий. Событие, которое осуществляется каждый раз, когда осуществляется эксперимент, называют достоверным. Достоверное событие есть множество Ω всех мыслимых элементарных исходов.

Невозможное событие — событие, которое не может произойти ни при каком результате эксперимента. Невозможное событие есть пустое множество \emptyset .

Произвольное событие как результат эксперимента есть некоторое подмножество A множества Ω ($A \subset \Omega$).

Операции над событиями. Для событий можно определить те же операции, которые были определены для множеств (см. с. 13):

сумма событий A и B (записывается $A \cup B$) есть событие, состоящее в том, что произошло либо A , либо B , либо оба эти события;

произведение событий A и B (записывается $A \cdot B$ или $A \cap B$) есть событие, состоящее в том, что произошло A и B одновременно;

разность событий A и B (записывается $A \setminus B$) есть событие, состоящее в том, что произошло A и не произошло B ;

противоположное событие (записывается \bar{A}) определяется как разность $\bar{A} = \Omega \setminus A$, т. е. это есть событие, состоящее в том, что произошло не A или A не произошло.

События A и B называют несовместными, если их произведение есть невозможное событие, $A \cdot B = \emptyset$, т. е. одновременное осуществление этих событий невозможно.

События A_1, A_2, \dots, A_n образуют полную группу несовместных событий, если:

$$1) \text{ для любых } i \neq j \ A_i \cdot A_j = \emptyset;$$

$$2) \bigcup_{k=1}^n A_k = \Omega.$$

σ -алгебра событий. При построении математической модели вероятностного эксперимента определяется некоторая σ -алгебра подмножеств A множества Ω ($A \subset \Omega$) (см. с. 16), и в дальнейшем только элементы этой σ -алгебры называют событиями, т. е. результатами эксперимента.

Замечания: 1. Выбор σ -алгебры событий с математической точки зрения, вообще говоря, произволен. Он определяется физической сущностью эксперимента и адекватностью вероятностной модели реальной ситуации.

2. Часто выбирают σ -алгебру событий как минимальную σ -алгебру, порожденную счетным разбиением $\Omega_1, \Omega_2, \dots, \Omega_n$ множества Ω ($\Omega_i \cap \Omega_j = \emptyset, i \neq j$ и $\bigcup_k \Omega_k = \Omega$).

Каждому событию A (элементу σ -алгебры \mathfrak{A} , $A \in \mathfrak{A}$) ставится в соответствие некоторое число $P(A) = p$ ($0 \leq p \leq 1$), которое называют вероятностью события A , т. е. на множестве \mathfrak{A} определяется действительная функция, обладающая следующими свойствами:

1. Для любого $A \in \mathfrak{A}$ справедливы неравенства

$$0 \leq P(A) \leq 1.$$

$$2. P(\Omega) = 1.$$

3. Если события $A_1, A_2, \dots, A_n, \dots$ попарно несовместны, т. е. для любых $i \neq j \ A_i \cdot A_j = \emptyset$, то

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k).$$

Свойства функции $P(A)$ постулируются, и поэтому их называют системой аксиом.

Замечания: 1. Определение вероятности полностью совпадает с определением вероятностной меры (см. с. 20).

2. Функция $P(A)$ определена на σ -алгебре \mathfrak{A} подмножеств Ω , поэтому вероятность приписывается каждому

событию, и по аксиоме 3 $\bigcup_{k=1}^{\infty} A_k$ есть

событие $\left(\bigcup_{k=1}^{\infty} A_k \in \mathfrak{A}\right)$, следовательно.

в последнем равенстве определены все элементы.

Элементарные свойства вероятности

1) $P(A) = 1 - P(\bar{A})$;

2) $P(\emptyset) = 0$;

3) $P(A \cup B) = P(A) + P(B) - P(A \cdot B)$;

4) $P\left(\bigcup_{k=1}^m A_k\right) \leq \sum_{k=1}^m P(A_k)$ для

любых событий $\{A_k, k = 1, m\}$;

5) для любых событий A_1, A_2, \dots, A_n имеет место равенство (теорема сложения)

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{i_1 > i_2} P(A_{i_1 i_2}) + \sum_{i_1 < i_2 < i_3} P(A_{i_1 i_2 i_3}) + \dots + (-1)^{k-1} \sum_{i_1 < i_2 < \dots < i_k} P(A_{i_1 i_2 \dots i_k}) + \dots + (-1)^{n-1} P(A_{123 \dots n}),$$

где $A_{i_1 i_2 \dots i_k} = A_{i_1} A_{i_2} \dots A_{i_k}$;

6) $P(A) \leq P(B)$, если $A \subseteq B$.

Свойство непрерывности вероятности. Для последовательности вложенных событий $A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots$ справедливо равенство

$$\lim_{n \rightarrow \infty} P(A_n) = P = \left(\lim_{n \rightarrow \infty} A_n\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right).$$

Для последовательности расширяющихся событий $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$ справедливо равенство

$$\lim_{n \rightarrow \infty} P(A_n) = P = \left(\lim_{n \rightarrow \infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right).$$

Замечание. Определенное выше свойство отождествляется со свойством непрерывности, потому что для вероятности как функции, определен-

ной на множествах, можно менять символ функции и предельный переход: предел функции равен значению функции в предельной «точке».

1.2. ДИСКРЕТНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

Вероятностное пространство $(\Omega, \mathfrak{A}, P)$, в котором множество элементарных событий Ω конечно или счетно, называют дискретным.

Способы построения σ -алгебры:

если Ω — счетно, то строится минимальная σ -алгебра, порожденная всевозможными конечными подмножествами E множества Ω (см. с. 16), в частности, системой множеств $\{\omega_i\}, i = \overline{1, \infty}$;

если Ω — конечно, то рассматривается, как правило, алгебра, образованная всевозможными подмножествами множества Ω .

Задание вероятностной меры на дискретном измеримом пространстве (Ω, \mathfrak{A}) . При определяемом выше выборе σ -алгебры \mathfrak{A} вероятностную меру в дискретном измеримом пространстве (Ω, \mathfrak{A}) можно определить, задавая вероятности множеств $\{\omega_i\}$: любой элемент A σ -алгебры \mathfrak{A} определяется набором (конечным или бесконечным) несовпадающих индексов $n_1, n_2, \dots, n_N \subset \mathbb{N}$, т. е. $A = \{\omega_{n_1}, \omega_{n_2}, \dots, \omega_{n_N}\}$. Тогда по определению (см. свойства вероятностной меры)

$$P(A) = \sum_{i \in N} p_i, \text{ где } p_i = P(\omega_{n_i}).$$

Примеры дискретных вероятностных пространств. 1. Бросание игральной кости (пример 1): $\Omega = \{\omega_1, \dots, \omega_6\}$;

$$p_i = \frac{1}{6}; i = \overline{1, 6}.$$

2. $\Omega = \{\omega_1, \omega_2\}, p_1 = p_2 = \frac{1}{2}$

(эксперимент можно трактовать как бросание однородной монеты).

3. Измерение сроков безотказной работы изделия (пример 3): $\Omega = \{\omega_1, \omega_2, \dots, \omega_k, \dots\}$, где ω_k есть событие, состоящее в том, что изделие откажет в интервале $(k, k + 1]$, $p_k = P(\omega_k) = p^k (1 - p), 0 < p < 1$.

Классическое определение вероятности. Если в эксперименте возможно только конечное N число элементарных исходов и каждый исход может осуществляться с одинаковой вероятностью, то вероятность произвольного события A равна $P(A) = \frac{M}{N}$, где N — число элементарных исходов, M — число элементарных событий, благоприятствующих появлению события A . (Элементарное событие ω называют благоприятствующим появлению события A , если при реализации события ω происходит событие A , т. е. $\omega \in A$, M — число элементов множества A .)

Замечания: 1. Классическое определение вероятности базируется на двух предположениях: конечность числа элементарных событий $N < \infty$ и равновероятность (принцип равных вероятностей).

2. Классическое определение вероятности есть частный случай задания вероятностной меры на дискретном (конечном) вероятностном пространстве (см. п. 1.2). Для указанного определения вероятности выполняются все аксиомы (см. с. 66).

Примеры: 1. Эксперимент: однократное бросание правильной игральной кости, на гранях которой написаны числа от 1 до 6. Число элементарных событий шесть ($N = 6$): ω_i — выпало i очков, $i = 1, 2, \dots, 6$. Если кость правильная, то предполагаем $P(\omega_i) = \frac{1}{6}$. Тогда вероятность выпадения менее трех очков (событие A) $P(A) = \frac{2}{6} = \frac{1}{3}$ (событию A благоприятствуют два элементарных события ω_1 и ω_2); вероятность выпадения четного числа очков (событие B) $P(B) = \frac{3}{6} = \frac{1}{2}$ (событию B благоприятствуют три элементарных события ω_2, ω_4 и ω_6).

2. Эксперимент — выборка с возвращением. Из урны, в которой находятся $2K$ шаров с номерами от 1 до $2K$, выбирается последовательно k шаров (после каждого выбора вынутый шар возвращается в урну). Элементарное событие есть последовательность k чисел (n_1, n_2, \dots, n_k) , $1 \leq n_i \leq 2K$, число элементарных событий $N = (2K)^k$. Предполагаем

равновероятность появления каждого элементарного события.

Тогда вероятность появления последовательности только из нечетных чисел (событие A)

$$P(A) = \frac{K^k}{2^k K^k} = \frac{1}{2^k}$$

(при выборке с возвращением на каждом шаге возможно K вариантов, благоприятствующих появлению события A ; при k шагах число таких событий равно K^k);

вероятность того, что в последовательности (n_1, \dots, n_k) не будет единицы (событие B)

$$P(B) = \frac{(2K-1)^k}{(2K)^k} = \left(1 - \frac{1}{2K}\right)^k$$

(при k -кратном выборе с возвращением число элементарных событий, благоприятствующих появлению события B , равно $(2K-1)^k$).

3. Эксперимент — выборка без возвращения. Из урны (см. пример 2) выбирается k шаров без возвращения. После выбора номера шаров упорядочиваются. Элементарное событие — последовательность $n_1 < n_2 < \dots < n_k$ (выборки отличаются только составом элементов, порядок их выбора неважен). Число элементарных событий $N = C_{2K}^k = \frac{(2K)!}{k!(2K-k)!}$. Выбор фиксированной последовательности осуществляется с вероятностью $\frac{1}{N}$. Тогда вероятность того, что в выборке m четных чисел, $m \leq k$ (событие A)

$$P(A) = \frac{C_K^m C_K^{k-m}}{C_{2K}^k},$$

в частности, вероятность того, что в выборке нет четных чисел, $m = 0$ (событие B)

$$P(B) = \frac{C_K^k}{C_{2K}^k};$$

вероятность того, что в выборке нет числа 1 (событие C)

$$P(C) = \frac{C_{2K-1}^k}{C_{2K}^k}.$$

2. ЭЛЕМЕНТАРНЫЕ ФОРМУЛЫ

Условная вероятность. Если для события B $P(B) > 0$, то для произвольного события A условная вероятность события A при условии выполнения события B (записывается $P(A/B)$) определяется как отношение

$$P(A/B) = \frac{P(A \cdot B)}{P(B)}. \quad (1)$$

Условная вероятность в классическом случае. Пусть N — число элементарных событий, событию B благоприятствует n событий, событию $A \cdot B$ благоприятствуют m событий ($m \leq n$, так как $A \cdot B \subseteq B$). Тогда

$$P(A/B) = \frac{P(A \cdot B)}{P(B)} = \frac{m/N}{n/N} = \frac{m}{n}. \quad (2)$$

Равенство (2) показывает: при определении условной вероятности вся вероятностная мера (единица) приписывается только элементарным событиям, благоприятствующим событию B .

Теорема умножения. Из равенства (1) имеем

$$\begin{aligned} P(A \cdot B) &= P(A/B) P(B), \\ P(A \cdot B) &= P(B/A) P(A). \end{aligned} \quad (3)$$

Для произведения n событий A_1, A_2, \dots, A_n справедливо равенство

$$\begin{aligned} P\left(\bigcap_{k=1}^n A_k\right) &= \\ &= P(A_1) P(A_2/A_1) P(A_3/A_1 A_2) \dots \quad (4) \\ &= P\left(A_{n-1} / \bigcap_{k=1}^{n-2} A_k\right) P\left(A_n / \bigcap_{k=1}^{n-1} A_k\right). \end{aligned}$$

Независимость событий. Два события A и B называют независимыми, если $P(A \cdot B) = P(A) \cdot P(B)$.

Для независимых событий A и B $P(A/B) = P(A)$, $P(B/A) = P(B)$.

Попарная независимость событий. События A_1, A_2, \dots, A_n называют попарно независимыми, если для любой пары индексов i, j ($1 \leq i < j \leq n$)

$$P(A_i A_j) = P(A_i) P(A_j).$$

Независимость в совокупности. События A_1, A_2, \dots, A_n называют независимыми в совокупности, если для любого набора несовпадающих индексов $1 \leq i_1 < i_2 < \dots < i_k \leq n$ ($k \leq n$)

$$P\left(\bigcap_{s=1}^k A_{i_s}\right) = \prod_{s=1}^k P(A_{i_s}).$$

Замечание. Независимые в совокупности события попарно независимы. Обратное неверно.

Пример. Пусть эксперимент заключается в равновероятном выборе одного из четырех шаров с номерами: 1 — на первом шаре, 2 — на втором, 3 — на третьем, (1, 2, 3) — на четвертом шаре. Пусть A_i ($i = 1, 2, 3$) есть событие, состоящее в том, что на наудачу выбранном шаре имеется цифра i . Тогда

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2};$$

$$\begin{aligned} P(A_1 \cdot A_2) &= P(A_1 \cdot A_3) = \\ &= P(A_2 \cdot A_3) = \frac{1}{4}; \end{aligned}$$

$$P(A_1 \cdot A_2 \cdot A_3) = \frac{1}{4},$$

т. е. события A_i ($i = 1, 2, 3$) попарно независимы, но не независимы в совокупности

$$P(A_1 \cdot A_2 \cdot A_3) \neq P(A_1) P(A_2) P(A_3).$$

Формула полной вероятности. Пусть события H_1, H_2, \dots, H_n образуют полную группу несовместных событий (см. с. 66), тогда для произвольного события A

$$P(A) = \sum_{k=1}^n P(A/H_k) P(H_k). \quad (5)$$

Формула Байеса. Пусть события H_1, H_2, \dots, H_n образуют полную груп-

пу несовместных событий, тогда для произвольного события A справедлива формула

$$P(H_k/A) = \frac{P(A/H_k)P(H_k)}{\sum_{s=1}^n P(A/H_s)P(H_s)}. \quad (6)$$

Замечание. Вероятности $P(H_s)$ называют **априорными**, вероятности $P(H_s/A)$ — **апостериорными**.

Формулу (6) можно трактовать следующим образом: пусть заранее (до опыта) известны вероятности различных гипотез (событий) H_k , и пусть в результате опыта (эксперимента) реализовалось событие A . Реализация события A меняет априорную информацию о событиях H_k . Формула (6) есть формула определения вероятностей событий H_k при указанном условии (реализация события A).

3. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Случайной величиной на вероятностном пространстве $(\Omega, \mathfrak{A}, P)$ называют измеримую функцию $\xi = \xi(\omega)$, определенную на элементах Ω и принимающую действительные значения ($\xi \in R^{(1)}$).

Замечания: 1. Измеримость функций $\xi(\omega)$ (или ξ) необходима для того, чтобы определить вероятность того, что случайная величина ξ принадлежит некоторому борелевскому множеству B на прямой $R^{(1)}$. Это делается следующим образом:

$$P(\xi \in B) = P(\omega : \xi(\omega) \in B); \quad (7)$$

так как $\xi(\omega)$ — измеримая функция, то $\{\omega : \xi(\omega) \in B\} \in \mathfrak{A}$, и, следовательно, правая часть равенства (7) определена. Таким образом, порождается вероятностная мера (вероятность) на борелевских множествах $B \in \mathfrak{B}$.

2. Определение случайной величины обязательно связывается с некоторым вероятностным пространством как с моделью вероятностного эксперимента.

3.1. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

Функцией распределения случайной величины $\xi = \xi(\omega)$ называют функцию

$$F_\xi(x) = P(\xi < x) = P(\omega : -\infty < \xi(\omega) < x). \quad (8)$$

Замечания:

так как множество $B = (-\infty, x)$ есть борелевское при любом x , то правая часть равенства (8) определена при любом x ;

так как вероятностная мера на борелевских множествах на прямой, т. е. на измеримом пространстве $(R^{(1)}, \mathfrak{B})$, однозначно определяется мерой, заданной на борелевских множествах вида $B = (-\infty, x)$, то для задания таковой достаточно определить функцию распределения;

согласно последнему замечанию часто отвлекаются от исходного вероятностного пространства $(\Omega, \mathfrak{A}, P)$ и рассматривают вероятностную меру на борелевских множествах из $R^{(1)}$, т. е. задают случайную величину функцией распределения (хотя можно построить разные случайные величины — измеримые функции, у которых одна и та же функция распределения).

Свойства функции распределения:

область определения $R^{(1)} = (-\infty, +\infty)$;

для любого $x \in R^{(1)}$ $0 \leq F_\xi(x) \leq 1$;

функция $F_\xi(x)$ — неубывающая,

непрерывная слева;

$$\lim_{x \rightarrow -\infty} F_\xi(x) = 0;$$

$$\lim_{x \rightarrow +\infty} F_\xi(x) = 1;$$

функция распределения имеет не более чем счетное число скачков (точек разрыва);

имеет место равенство $P(a \leq \xi < b) = F_\xi(b) - F_\xi(a)$ при $a < b$;

если в точке x функция $F_\xi(x)$ непрерывна, то $P(\xi = x) = 0$;

если в точке x функция $F_\xi(x)$ имеет разрыв, то $P(\xi = x) = F_\xi(x+0) - F_\xi(x) > 0$;

любая функция распределения представляется в виде суммы

$$F_\xi(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x) + \alpha_3 F_3(x);$$

$$\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_3 \geq 0, \alpha_1 + \alpha_2 + \alpha_3 = 1,$$

где $F_1(x)$ — ступенчатая функция, обладающая всеми перечисленными свойствами функции распределения;

$F_2(x)$ — непрерывная функция, обладающая всеми перечисленными свойствами функции распределения и для которой существует такая функция $f_2(x)$, что справедливо представление

$$F_2(x) = \int_{-\infty}^x f_2(y) dy;$$

$F_3(x)$ — непрерывная неубывающая функция, для которой $F_3(-\infty) = 0$, $F_3(+\infty) = 1$ и $\frac{dF_3(x)}{dx} = 0$, за исключением, быть может, борелевского множества из $R^{(1)}$, для которого вероятностная мера равна нулю (сингулярная компонента)*.

3.2. ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дискретной случайной величиной называют измеримую функцию, принимающую конечное или счетное число значений.

Замечания: 1. Если (x_1, x_2, \dots, x_n) — значения, которые принимает дискретная случайная величина $\xi = \xi(\omega)$, то множество элементарных событий Ω разбивается на пересекающиеся подмножества:

$$A_i = \{\omega : \xi(\omega) = x_i\}, \quad i = 1, 2,$$

$$A_i A_j = \emptyset \text{ при } i \neq j, \quad \bigcup_i A_i = \Omega.$$

2. Для любого борелевского множества B в соответствии с (7) положим

$$\begin{aligned} P(\xi \in B) &= P(\omega : \xi(\omega) \in B) = \\ &= P\left(\bigcup_{i: x_i \in B} A_i\right) = \sum_{i: x_i \in B} P(A_i); \end{aligned}$$

последнее равенство справедливо, так как множества A_i непересекающиеся.

* В дальнейшем сингулярную компоненту рассматривать не будем, считая $\alpha_3 = 0$, так как в распределениях, встречающихся в теории надежности и эффективности, сингулярная компонента отсутствует. Таким образом, будем исследовать смеси дискретных и абсолютно непрерывных распределений, или просто дискретные распределения ($\alpha_2 = 0$), или абсолютно непрерывные распределения ($\alpha_1 = 0$).

3. В частном случае при $B = \{x_i\}$ имеем

$$P(\xi = x_i) = P(A_i) = p_i.$$

Таким образом, случайную дискретную величину можно задавать набором ее значений $(x_1, x_2, \dots, x_n, \dots)$ и вероятностями $p_i = P(\xi = x_i)$. Набор таких пар (x_i, p_i) называют дискретным распределением.

Функция распределения дискретной случайной величины. В соответствии с общим определением для дискретной случайной величины функция распределения определяется равенством для любого $x \in R^{(1)}$

$$F_\xi(x) = P(\xi < x) = \sum_{i: x_i < x} p_i. \quad (9)$$

Свойства:

функция распределения (9) обладает всеми свойствами, перечисленными в п. 3.1;

функция распределения (9) представляет собой ступенчатую функцию со скачками величины p_i в точках x_i , $i = 1, 2$,

Решетчатое распределение. Случайная величина имеет решетчатое распределение, если она дискретная, и для тех значений $(x_1, x_2, \dots, x_n, \dots)$, которые принимаются с положительной вероятностью $p_i = P(\xi = x_i) > 0$, существует a, h и набор целых чисел $k_1, k_2, \dots, k_n, \dots$, таких, что справедливо представление $x_i = a + k_i h$ для всех $i = 1, 2$

Примеры. 1. *Биномиальное распределение:*

$$\begin{aligned} p_k &= P(\xi = k) = \\ &= \begin{cases} 0, & k < 0; \\ C_n^k p^k (1-p)^{n-k}, & 0 \leq k \leq n; \\ 0, & k > n, \end{cases} \end{aligned}$$

k, n — целые числа; $n > 0; 0 < p < 1$.

2. *Геометрическое распределение:*

$$\begin{aligned} p_k &= P(\xi = k) = \\ &= \begin{cases} 0 & \text{при } k < 0; \\ p^{k-1} (1-p) & \text{при } k \geq 0, 0 < p < 1, \end{cases} \end{aligned}$$

k — целое число.

3. Распределение Пуассона:

$$p_k = P(\xi = k) = \begin{cases} 0 & \text{при } k < 0; \\ \frac{\lambda^k}{k!} e^{-\lambda} & \text{при } k \geq 0; \lambda > 0, \end{cases}$$

k — целое число.

4. Отрицательное биномиальное распределение:

$$p_k = P(\xi = k) = \begin{cases} 0 & \text{при } k < n; \\ C_{k-1}^{n-1} p^n (2-p)^{k-n} & \text{при } k \geq n; \\ 0 < p < 1, \end{cases}$$

k, n — целые числа.

5. Равномерное распределение:

$$p_k = P(\xi = k) = \begin{cases} 0 & k \leq 0; \\ \frac{1}{n} & 0 < k \leq n; \\ 0 & k > n, \end{cases}$$

k, n — целые числа, $n > 0$.

6. Гипергеометрическое распределение:

$$p_m = P(\xi = m) = \begin{cases} 0 & m < 0; \\ \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} & 0 \leq m \leq n; \\ 0 & m > n, \end{cases}$$

$N \geq M, N > n; N, M, n, m$ — целые, положительные числа.

3.3. НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Если функция распределения $F_\xi(x)$ непрерывна при любом $x \in R^{(1)}$, то говорят, что $\xi = \xi(\omega)$ непрерывная случайная величина.

Случайная величина ξ имеет абсолютно непрерывное распределение, если существует такая функция $f_\xi(x)$, что

$$F_\xi(x) = \int_{-\infty}^x f_\xi(y) dy.$$

Функцию $f_\xi(x)$ называют плотностью распределения случайной величины ξ .

Свойства плотности:

$$f_\xi(x) \geq 0 \text{ почти для всех } x \in R^{(1)};$$

$$\int_{-\infty}^{\infty} f_\xi(x) dx = 1;$$

$$P(a \leq \xi < b) = \int_a^b f_\xi(x) dx = F_\xi(b) - F_\xi(a), \quad a \leq b.$$

Примеры. 1. *Нормальное распределение.* Случайная величина ξ распределена по нормальному закону с параметрами a и σ ($\sigma > 0$), если

$$F_\xi(x) = P(\xi < x) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^x e^{-\frac{(u-a)^2}{2\sigma^2}} du, \\ -\infty < x < \infty.$$

Плотность нормального распределения определяется равенством

$$f_\xi(x) = F'_\xi(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

2. *Усеченное нормальное распределение.* Плотность усеченного нормального распределения определяется равенством

$$f(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{1}{c\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} & \text{при } x \geq 0, \end{cases}$$

$a \in R^{(1)}, \sigma > 0; c$ — нормирующий множитель.

3. *Логарифмически нормальное распределение.* Такое распределение имеет по определению случайная величина ξ , логарифм которой $\eta = \ln \xi$ распределен по нормальному закону. Плотность распределения определяется равенством

$$f_{\xi}(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{1}{\sqrt{2\pi} \sigma x} e^{-\frac{(\ln x - a)^2}{2\sigma^2}} & \text{при } x > 0. \end{cases}$$

4. *Равномерное распределение.* Случайную величину ξ называют равномерно распределенной на отрезке $[a, b]$, если ее плотность $f_{\xi}(x)$ задается равенством

$$f_{\xi}(x) = \begin{cases} \frac{1}{b-a} & \text{при } a \leq x \leq b; \\ 0 & \text{при } x < a \text{ или } x > b. \end{cases}$$

Функция распределения имеет вид

$$F_{\xi}(x) = \begin{cases} 0 & \text{при } x < a; \\ \frac{x-a}{b-a} & \text{при } a \leq x \leq b; \\ 1 & \text{при } x > b. \end{cases}$$

5. *Гамма-распределение (Г-распределение).* Двухпараметрическое семейство распределений, для которого плотность задается равенством

$$f(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{при } x \geq 0, \end{cases}$$

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-\lambda x} dx,$$

где $\alpha > 0$ — параметр формы; $\lambda > 0$ — параметр масштаба.

6. *Экспоненциальное распределение.* При $\alpha = 1$ и произвольном $\lambda > 0$ получаем экспоненциальное распределение

$$f(x) = \begin{cases} 0 & \text{при } x < 0; \\ \lambda e^{-\lambda x} & \text{при } x \geq 0. \end{cases}$$

Функция распределения имеет вид

$$F(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ 1 - e^{-\lambda x} & \text{при } x > 0. \end{cases}$$

7. *Распределение Эрланга.* Если α — целое, то из гамма-распределения получаем распределение Эрланга.

Функция распределения случайной величины ξ , распределенной по закону Эрланга, равна

$$F_{\xi}(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ 1 - \sum_{k=0}^{\alpha-1} \frac{(\lambda x)^k}{k!} e^{-\lambda x} & \text{при } x > 0. \end{cases}$$

8. *Хи-квадрат распределение (χ^2 -распределение).* Плотность χ^2 -распределения определяется равенством

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & \text{при } x > 0. \end{cases}$$

Параметр n определяет число степеней свободы (n — целое, положительное).

9. *Распределение Вейбулла — Гнеденко.* Двухпараметрическое семейство распределений, для которого плотность задается равенством

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{\alpha}{\lambda} x^{\alpha-1} e^{-\frac{x^{\alpha}}{\lambda}} & \text{при } x > 0, \end{cases}$$

где $\alpha > 0$ — параметр формы, $\lambda > 0$ — параметр масштаба.

10. *Бета-распределение (В-распределение).* Двухпараметрическое семейство функций распределения, для которых плотность задается равенством

$$f(x) = \begin{cases} 0 & \text{при } x < 0 \text{ или } x > 1; \\ \frac{x^{a-1} (1-x)^{b-1}}{B(a, b)} & \text{при } 0 \leq x \leq 1, \end{cases}$$

где $B(a, b) =$

$$= \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)} \text{ — бета-функция.}$$

11. *Распределение Стьюдента (t-распределение)*. Плотность распределения Стьюдента определяется равенством

$$f(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \times \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

Параметр n называется числом степеней свободы (n — целое, положительное число).

12. *Распределение Фишера*. Двухпараметрическое семейство распределений, для которого плотность задается равенством

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{\Gamma\left(\frac{n+m}{2}\right) x^{\frac{m}{2}-1}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{m}{2}\right) (1+x)^{\frac{m+n}{2}}} & \text{при } x > 0, \end{cases}$$

m, n — целые, положительные числа.

3.4. ПОЛОЖИТЕЛЬНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Случайная величина $\xi = \xi(\omega)$ называется **положительной**, если $\{\omega : \xi < 0\} = \emptyset$, т. е. областью значений измеримой функции $\xi(\omega)$ является положительная полупрямая $R^+ = [0, +\infty)$.

Свойства:

для положительной случайной величины $F_\xi(0) = 0$;

для положительной случайной величины, имеющей абсолютно непрерывное распределение,

$$f_\xi(x) = 0 \text{ при } x < 0;$$

для дискретной положительной случайной величины, принимающей значения (x_1, x_2, \dots) с вероятностями $p_k > 0$, справедливы неравенства $x_k \geq 0, k = 1, 2,$

Примеры:

примеры п. 3.2 — примеры положительных дискретных случайных величин;

примеры 2, 3, 5, 6, 7, 8, 9, 10, 12 п. 3.3 — примеры положительных случайных величин.

4. СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

Многомерной случайной величиной (системой случайных величин) на вероятностном пространстве $(\Omega, \mathfrak{A}, P)$ называют измеримую функцию $\xi(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$, определенную на множестве Ω и принимающую действительные значения из пространства $R^{(n)}$ ($\xi \in R^{(n)}$). Таким образом порождается вероятностная мера на борелевских множествах B в $R^{(n)}$.

В частности, если $B = (-\infty < x_1 < b_1, \dots, -\infty < x_n < b_n)$, то определена функция

$$F_\xi(b_1, b_2, \dots, b_n) = P(\omega : -\infty < \xi_1(\omega) < b_1, \dots, -\infty < \xi_n(\omega) < b_n) = \dots$$

$$F_\xi(b_1, b_2, \dots, b_n) = P(\xi_1 < b_1, \xi_2 < b_2, \dots, \xi_n < b_n). \quad (10)$$

Многомерной функцией распределения случайных величин $(\xi_1, \xi_2, \dots, \xi_n)$ или **совместной функцией распределения** называют функцию $F_{\xi_1, \dots, \xi_n}(b_1, b_2, \dots, b_n)$, определенную равенством (10).

Свойства многомерной функции распределения $F_{\xi_1, \dots, \xi_n}(b_1, \dots, b_n)$:

область определения $R^{(n)}$;

для любого $(x_1, x_2, \dots, x_n) \in R^{(n)}$

$$0 \leq F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) \leq 1;$$

функция распределения неубывающая по каждому аргументу;

функция распределения непрерывная слева по каждому аргументу;

при любом $1 \leq k \leq n$

$$\lim_{x_k \rightarrow -\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = 0;$$

$$\lim_{\substack{x_k \rightarrow +\infty \\ k=1, n}} F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = 1;$$

для любого набора $0 < i_1 < i_2 < \dots < i_m \leq n, m \leq n$

$$F_{\xi_{i_1} \xi_{i_2} \dots \xi_{i_m}}(x_{i_1}, x_{i_2}, \dots, x_{i_m}) = \lim_{\substack{x_k \rightarrow +\infty \\ k \neq i_s \\ s=1, m}} F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n).$$

В частности, при $n = 1, i_1 = k$ имеем

$$F_{\xi_k}(x_k) = F_{\xi_1 \xi_2 \dots \xi_n}(\underbrace{+\infty, +\infty, \dots, +\infty}_{k-1}, x_k, +\infty, \dots, +\infty).$$

Дискретным случайным вектором называют систему случайных величин, принимающую конечное или счетное число значений.

Пусть случайные величины ξ_k принимают значения $(x_{k1}, x_{k2}, \dots), k = 1, 2, \dots, n$, тогда совместное распределение задается набором вероятностей

$$P_{i_1 i_2 \dots i_n} = P(\xi_1 = x_{1i_1}, \xi_2 = x_{2i_2}, \dots, \xi_n = x_{ni_n}).$$

Совместная функция распределения для этого случая задается формулой

$$F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = \sum_{\substack{i_1: x_{1i_1} < x_1 \\ \dots \\ i_n: x_{ni_n} < x_n}} P_{i_1 i_2 \dots i_n}.$$

Случайный вектор $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ имеет абсолютно непрерывное распределение, если существует такая функция $f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$, что

$$F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f_{\xi_1 \xi_2 \dots \xi_n}(y_1, y_2, \dots, y_n) dy_1 dy_2 \dots dy_n.$$

Функцию $f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$ называют плотностью совместного

распределения случайных величин $(\xi_1, \xi_2, \dots, \xi_n)$.

Свойства плотности совместного распределения:

$f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) \geq 0$ для почти всех (x_1, x_2, \dots, x_n) ;

$$f_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) = \frac{\partial^n F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

для почти всех (x_1, x_2, \dots, x_n) ;

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1;$$

$$\underbrace{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) \times dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n}_{n-1} = f_{\xi_k}(x_k)$$

[свойство согласованности маргинальных (частных) распределений];

$$\int_B f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = P((\xi_1, \dots, \xi_n) \in B).$$

Пример. Случайный вектор $(\xi_1, \xi_2, \dots, \xi_n), n > 1$, имеет нормальное распределение (многомерное), если оно абсолютно непрерывное распределение с плотностью вида

$$f(x) = f(x_1, x_2, \dots, x_n) = \frac{c}{(2\pi)^{n/2}} e^{-\frac{1}{2} \{(x-a)^T B (x-a)\}}$$

где c — нормирующая константа; $c = \sqrt{|B|}$; B — матрица с определителем $|B| > 0$; a — вектор-столбец (a_1, \dots, a_n) ; x — вектор-столбец (x_1, \dots, x_n) ; $(x-a)^T$ — транспонированный вектор (вектор-строка).

Вектор (ξ_1, ξ_2) имеет двумерное нормальное распределение, если его плотность $f(x_1, x_2)$ равна

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ \times \exp \left\{ -\frac{1}{2} \frac{1}{1-\rho^2} \left[\frac{(x_1 - a_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - a_1)(x_2 - a_2)}{\sigma_1\sigma_2} + \frac{(x_2 - a_2)^2}{\sigma_2^2} \right] \right\}, |\rho| < 1, \sigma_1 > 0, \sigma_2 > 0.$$

Независимость двух случайных величин. Две случайные величины ξ и η называют независимыми, если

$$F_{\xi\eta}(x, x_2) = F_{\xi}(x_1) F_{\eta}(x_2)$$

для любых x_1 и x_2 .

Независимость двух групп случайных величин. Две группы случайных величин (ξ_1, \dots, ξ_n) и (η_1, \dots, η_m) называют независимыми, если

$$F_{\xi_1 \dots \xi_n \eta_1 \dots \eta_m}(x_1, \dots, x_n, y_1, \dots, y_m) = \\ = F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) \times \\ \times F_{\eta_1 \dots \eta_m}(y_1, \dots, y_m).$$

Независимость в совокупности. Случайные величины называют независимыми в совокупности, если

$$F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) = \\ = \prod_{k=1}^n F_{\xi_k}(x_k).$$

Замечания: 1. Из независимости в совокупности следует независимость любых двух групп различных случайных величин из (ξ_1, \dots, ξ_n) , в частности попарная независимость. Обратное неверно.

2. Если случайные величины (ξ_1, \dots, ξ_n) имеют абсолютно непрерывное распределение, то для независимых в совокупности случайных величин совместная плотность представляется

в виде произведения одномерных плотностей

$$f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = \\ = \prod_{k=1}^n f_{\xi_k}(x_k).$$

5. ФУНКЦИИ ОТ СЛУЧАЙНЫХ ВЕЛИЧИН

Пусть задан случайный вектор $(\xi_1, \xi_2, \dots, \xi_n)$ и измеримая функция $\varphi(x_1, \dots, x_n)$, определенная в пространстве $R^{(n)}$ и принимающая значения из $R^{(1)}$.

Тогда:

функция $\eta = \varphi(\xi_1, \xi_2, \dots, \xi_n)$ есть случайная величина (см. свойства измеримых функций, с. 19);

функция распределения случайной величины $\eta = \varphi(\xi_1, \dots, \xi_n)$ определяется равенством

$$F_{\eta}(x) = P(\eta < x) = \\ = P(\varphi(\xi_1, \dots, \xi_n) < x) = \\ = P(\{\omega : \varphi(\xi_1(\omega), \dots, \xi_n(\omega)) < x\});$$

если случайные величины ξ_k , $k = \overline{1, n}$ имеют дискретное распределение, которое определяется константами

$$p_{i_1 i_2 \dots i_n} = \\ = P(\xi_1 = x_{1i_1}, \dots, \xi_n = x_{ni_n}),$$

то функция распределения определяется равенством

$$F_{\eta}(x) = P(\eta < x) = \\ = \sum p_{i_1 i_2 \dots i_n} \\ \{ (i_1, i_2, \dots, i_n) : \varphi(x_{1i_1}, x_{2i_2}, \dots, \\ \dots, x_{ni_n}) < x \};$$

если случайные величины имеют абсолютно непрерывное распределение и $f_{\xi_1 \dots \xi_n}(x_1, \dots, x_n)$ — совместная плотность этого распределения, то

$$F_{\eta}(x) = \int \int f_{\xi_1 \dots \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ \{ (x_1, \dots, x_n) : \varphi(x_1, \dots, x_n) < x \}$$

Сумма случайных величин. Пусть $\eta = \xi_1 + \xi_2$, тогда

$$F_\eta(x) = P(\{\omega : \xi_1(\omega) + \xi_2(\omega) < x\}); \quad (11)$$

если ξ_i ($i = 1, 2$) — дискретные случайные величины и $p_{km} = P(\xi_1 = x_{1k}, \xi_2 = x_{2m})$ ($k = 1, 2, \dots, m = 1, 2, \dots$), то

$$F_\eta(x) = \sum_{\{(k, m) : x_{1k} + x_{2m} < x\}} p_{km} = \sum_{k=1}^{\infty} \sum_{\{m : x_{2m} < x - x_{1k}\}} p_{km},$$

причем если $\{m : x_{2m} < x - x_{1k}\} = \emptyset$, то внутренняя сумма считается равной нулю;

если ξ_i — целочисленные случайные величины, принимающие значения $x_k = 0, \pm 1, \pm 2, \dots, \pm n, \dots$, то

$$P_m = P(\xi_1 + \xi_2 = m) = \sum_{k=-\infty}^{+\infty} p_{m-k, k};$$

если в условиях предыдущего пункта случайные величины ξ_1 и ξ_2 независимы $p_{ij} = p_i q_j$, $p_i = P(\xi_1 = i)$, $q_j = P(\xi_2 = j)$, то

$$P_m = \sum_{k=-\infty}^{+\infty} p_{m-k} q_k;$$

если совместное распределение случайных величин (ξ_1, ξ_2) имеет плотность $f_{\xi_1 \xi_2}(x_1, x_2)$, то

$$F_\eta(x) = \int_{-\infty}^{x+\infty} \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(z-y, y) dy dz,$$

$$f_\eta(x) = \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(x-y, y) dy;$$

если совместное распределение (ξ_1, ξ_2) имеет плотность и случайные величины ξ_1 и ξ_2 независимы, то

$$f_{\xi_1 \xi_2}(x_1, x_2) = f_{\xi_1}(x_1) f_{\xi_2}(x_2);$$

$$F_\eta(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f_{\xi_1}(z-y) f_{\xi_2}(y) dy dz;$$

$$f_\eta(x) = \int_{-\infty}^{+\infty} f_{\xi_1}(x-y) f_{\xi_2}(y) dy;$$

если совместное распределение (ξ_1, ξ_2) имеет плотность и случайные величины ξ_1 и ξ_2 положительные, то

$$F_\eta(x) = \begin{cases} \int_0^x \int_0^z f_{\xi_1 \xi_2}(z-y, y) dy dz & \text{при } x > 0; \\ 0 & \text{при } x \leq 0, \end{cases}$$

$$f_\eta(x) = \begin{cases} \int_0^x f_{\xi_1 \xi_2}(x-y, y) dy & \text{при } x > 0; \\ 0 & \text{при } x \leq 0; \end{cases}$$

если совместное распределение (ξ_1, ξ_2) имеет плотность и случайные величины независимы и положительны, то при $x \geq 0$

$$F_\eta(x) = \int_0^x \int_0^z f_{\xi_1}(z-y) f_{\xi_2}(y) dy;$$

$$f_\eta(x) = \int_0^x f_{\xi_1}(x-y) f_{\xi_2}(y) dy.$$

Замечания: 1. Приведенные выше зависимости обобщаются на случай произвольного числа слагаемых $\eta = \xi_1 + \xi_2 + \dots + \xi_n$. Общее определение (11) остается в силе

$$F_\eta(x) = P(\{\omega : \xi_1(\omega) + \xi_2(\omega) + \dots + \xi_n(\omega) < x\}); \quad (12)$$

в остальных формулах изменяется только область суммирования или интегрирования в соответствии с определением (12).

2. При определении характеристик распределения суммы независимых в совокупности случайных величин можно использовать приведенные выше формулы рекуррентно, считая

$$\eta_{k+1} = \eta_k + \xi_{k+1}, \quad k = 1, 2, \dots, n;$$

$$\eta_1 = \xi_1 \text{ и } \eta = \eta_n.$$

6. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Математическим ожиданием $M\xi$ случайной величины ξ с функцией распределения $F_\xi(x)$ называют интеграл

$$M\xi = \int_{-\infty}^{+\infty} x dF_\xi(x) \stackrel{\text{def}}{=}$$

$$\stackrel{\text{def}}{=} \alpha_1 \sum_{k=0}^{\infty} x_k p_k + \alpha_2 \int_{-\infty}^{+\infty} x f(x) dx,$$

если ряд и интеграл, входящие в правую часть равенства, абсолютно сходятся, где

$$F_\xi(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x); \quad \alpha_i \geq 0,$$

$$\alpha_1 + \alpha_2 = 1;$$

$$F_1(x) = \sum_{s: x_s < x} p_s \text{ — дискретная компонента (ступенчатая функция);}$$

$$F_2(x) = \int_{-\infty}^x f(y) dy \text{ — абсолютно непрерывная компонента, } f(x) \text{ — ее плотность.}$$

Математическое ожидание дискретной случайной величины. Пусть дискретная случайная величина ξ принимает значения $(x_0, x_1, \dots, x_n, \dots)$ с вероятностями $p_k = P(\xi = x_k)$, $p_k \geq 0$,

$\sum_{k=0}^{\infty} p_k = 1$. Тогда математическим ожиданием $M\xi$ называют сумму ряда

$$M\xi = \sum_{k=0}^{\infty} x_k p_k,$$

если этот ряд абсолютно сходится.

Математическое ожидание абсолютно непрерывной случайной величины. Если ξ — абсолютно непрерывная случайная величина с плотностью распределения $f_\xi(x)$, то математическое ожидание $M\xi$ есть интеграл

$$M\xi = \int_{-\infty}^{\infty} x f_\xi(x) dx,$$

если он абсолютно сходится.

Свойства математического ожидания: если существуют математические ожидания $M\xi_i$, $i = 1, 2, \dots, n$, то существует математическое ожидание суммы

$$M \left(\sum_{k=1}^n \xi_k \right) \text{ и справедливо равенство}$$

$$M \left(\sum_{k=1}^n \xi_k \right) = \sum_{k=1}^n M\xi_k, \text{ т. е. математическое ожидание суммы равно сумме математических ожиданий;}$$

константу можно выносить за знак математического ожидания, т. е.

$$M(a\xi) = aM\xi;$$

из приведенных выше свойств следует, что математическое ожидание линейной комбинации случайных величин равно линейной комбинации математических ожиданий

$$M \left(\sum_{k=1}^n a_k \xi_k \right) = \sum_{k=1}^n a_k M\xi_k;$$

математическое ожидание константы C ($P(\xi = C) = 1$) равно этой константе, т. е. $MC = C$;

если случайные величины ξ и η независимы, то $M\xi\eta = M\xi \cdot M\eta$.

Математическое ожидание функции от случайных величин. Пусть задана случайная величина ξ с функцией распределения $F_\xi(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x)$ и измеримая функция $\varphi(x)$, тогда $\eta = \varphi(\xi)$ есть случайная величина

и

$$M\eta = \int_{-\infty}^{\infty} \varphi(x) dF_{\xi}(x) =$$

$$= \alpha_1 \sum_{k=0}^{\infty} \varphi(x_k) p_k +$$

$$+ \alpha_2 \int_{-\infty}^{\infty} \varphi(x) f(x) dx,$$

если ряд и интеграл абсолютно сходятся.

Моменты старших порядков. Моментом m -го порядка $M\xi^m$ случайной величины ξ с функцией распределения $F_{\xi}(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x)$ называют интеграл

$$M\xi^m = \int_{-\infty}^{\infty} x^m dF_{\xi}(x).$$

Центральным моментом m -го порядка называют m -й момент центрированной (с математическим ожиданием ноль) случайной величины $\eta = \xi - M\xi$, т. е.

$$M(\xi - M\xi)^m =$$

$$= \int_{-\infty}^{\infty} (x - M\xi)^m dF_{\xi}(x).$$

Дисперсией $D\xi$ случайной величины ξ называют центральный момент второго порядка ($m = 2$)

$$D\xi = M(\xi - M\xi)^2 =$$

$$= \int_{-\infty}^{\infty} (x - M\xi)^2 dF_{\xi}(x) = M\xi^2 - (M\xi)^2.$$

Если распределение $F_{\xi}(x)$ дискретно ($\alpha_2 = 0$), то

$$D\xi = \sum_{k=0}^{\infty} x_k^2 p_k - \left(\sum_{k=0}^{\infty} x_k p_k \right)^2$$

Если распределение $F_{\xi}(x)$ абсолютно непрерывно ($\alpha_1 = 0$), то

$$D\xi = \int_{-\infty}^{+\infty} x^2 f_{\xi}(x) dx -$$

$$- \left(\int_{-\infty}^{+\infty} x f_{\xi}(x) dx \right)^2.$$

Величину $\sqrt{D\xi}$ называют **средним квадратическим отклонением** случайной величины ξ .

Свойства дисперсии: дисперсия суммы попарно независимых случайных величин равна сумме дисперсий

$$D\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n D\xi_i; \quad (13)$$

константа выносится за знак дисперсии в квадрате

$$D(a\xi) = a^2 D\xi;$$

если $D\xi = 0$, то $P(\xi = M\xi) = 1$; дисперсия — величина неотрицательная $D\xi \geq 0$;

дисперсия константы равна нулю $DC = 0$;

$$D(a\xi + b) = a^2 D\xi.$$

Пусть задан случайный вектор $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ с функцией распределения $F_{\xi}(x) = F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n)$.

Тогда смешанным моментом k -го порядка случайных величин ξ_1, \dots, ξ_n называют интеграл

$$M(\xi_1^{k_1} \xi_2^{k_2} \dots \xi_n^{k_n}) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} dF_{\xi}(x_1, x_2, \dots, x_n),$$

где $k_i \geq 0$; $\sum_{i=1}^n k_i = k$.

Если распределение $F_{\xi}(x_1, \dots, x_n)$ дискретно, т. е. определяется кон-

Числовые характеристики случайных величин

Распределение	Математическое ожидание	Дисперсия
Биномиальное	np	$np(1-p)$
Отрицательное биномиальное	$\frac{n}{p}$	$\frac{n(1-p)}{p^2}$
Геометрическое	$\frac{1}{1-p}$	$\frac{p}{(1-p)^2}$
Пуассона	$\frac{1}{\lambda}$	$\frac{1}{\lambda}$
Равномерное дискретное	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Гипергеометрическое	$\frac{M}{N}$	$\frac{M(N-M)}{N^2}$
Нормальное	a	σ^2
Усеченное нормальное	$\frac{a}{2c}$	$\frac{\sigma^2}{2c} + \frac{a^2}{2c} \left(\frac{2c-1}{2c} \right)^2$
Логарифмически нормальное	$e^{(a-\sigma^2)}$	$e^{(2a+\sigma^2)} (e^{\sigma^2} - 1)$
Равномерное	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Гамма-распределение	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
Экспоненциальное	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Эрланга	$\frac{\alpha}{\lambda}$, α — целое	$\frac{\alpha}{\lambda^2}$
Вейбулла—Гнеденко	$\left(\frac{1}{\lambda} \right)^{\frac{1}{\alpha}} \Gamma \left(1 + \frac{1}{\alpha} \right)$	$\left(\frac{1}{\lambda} \right)^{2/\alpha} \left[\Gamma \left(1 + \frac{2}{\alpha} \right) - \Gamma^2 \left(1 + \frac{1}{\alpha} \right) \right]$
Бета-распределение	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2 (a+b+1)}$
χ^2 -распределение	n	$2n$
t -распределение	0	$\frac{n}{n-2}$
Фишера	$\frac{m}{n-2}$ ($n > 2$)	$\frac{2m(m+n-2)}{(n-2)^2 (n-4)}$, $n > 4$

стантами $p_{i_1 i_2 \dots i_n} = P(\xi_1 = y_{1i_1}, \xi_2 = y_{2i_2}, \dots, \xi_n = y_{ni_n}), p_{i_1 i_2 \dots i_n} \geq 0, i_s \in I_s (I_s - \text{некоторое множество индексов}), \text{причем } \sum_{\substack{i_s \in I_s \\ s=1, n}} p_{i_1 \dots i_n} =$

$= 1$, то

$$M \xi_1^{k_1} \xi_2^{k_2} \dots \xi_n^{k_n} =$$

$$= \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \dots$$

$$\sum_{i_n \in I_n} y_{1i_1}^{k_1} y_{2i_2}^{k_2} \dots y_{ni_n}^{k_n} p_{i_1 i_2 \dots i_n}$$

$$\cdot y_{ni_n}^{k_n} p_{i_1 i_2 \dots i_n}$$

Если распределение $F_{\xi}(x)$ абсолютно непрерывно, т. е. имеет совместную плотность $f_{\xi}(x) = f_{\xi}(x_1, x_2, \dots, x_n)$, то

$$M \xi_1^{k_1} \xi_2^{k_2} \dots \xi_n^{k_n} =$$

$$= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} f_{\xi}(x_1, \dots, x_n) dx_1 \dots dx_n$$

$$\cdot x_n^{k_n} f_{\xi}(x_1, \dots, x_n) dx_1 \dots dx_n$$

Ковариацией $\text{cov}(\xi, \eta)$ двух случайных величин ξ и η называют смешанный момент второго порядка ($n = 2$) центрированных случайных величин $\xi - M\xi$ и $\eta - M\eta$

$$\text{cov}(\xi, \eta) = M[(\xi - M\xi)(\eta - M\eta)]$$

Для произвольно зависимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$

$$D \left(\sum_1^n \xi_i \right) =$$

$$= \sum_{i=1}^n D\xi_i + 2 \sum_{i>j} \text{cov}(\xi_i \xi_j)$$

Коэффициентом корреляции $R_{\xi\eta}$ случайных величин ξ и η называют отношение

$$R_{\xi\eta} = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi D\eta}}$$

Свойства: $|R_{\xi\eta}| \leq 1; R_{\xi\xi} = 1$; если $|R_{\xi\eta}| = 1$, то с вероятностью

единица случайные величины ξ и η связаны линейной зависимостью

$$\eta = R_{\xi\eta} \sqrt{\frac{D\eta}{D\xi}} (\xi - M\xi) + M\eta$$

Ковариационной матрицей случайных величин $(\xi_1, \xi_2, \dots, \xi_n)$ называют квадратную, симметричную относительно главной диагонали матрицу C с элементами $C_{ij} = \text{cov}(\xi_i \xi_j)$. Ковариационная матрица положительно определена.

Некоррелированность. Случайные величины $\xi_1, \xi_2, \dots, \xi_n$ называют попарно некоррелированными, если для любых $i, j (i \neq j) R_{\xi_i \xi_j} = 0$.

Если случайные величины попарно независимы, то они попарно некоррелированы. Обратное в общем случае неверно.

Если случайные величины имеют многомерное нормальное распределение и попарно некоррелированы, то они независимы.

Если случайные величины попарно некоррелированы, то выполняется равенство (13) для дисперсий.

В таблице на с. 80 приведены примеры числовых характеристик (математических ожиданий и дисперсий) случайных величин, распределения которых приведены на с. 71—74.

7. НЕРАВЕНСТВА ЧЕБЫШЕВА И КОЛМОГОРОВА

Неравенство Чебышева. *Общий случай.* Пусть ξ — произвольная случайная величина, $g(x)$ — четная функция, $g(x) \geq 0$ и неубывающая при $x \in [0, +\infty)$. Тогда при любом $\epsilon > 0$

$$P(|\xi| \geq \epsilon) \leq \frac{Mg(\xi)}{g(\epsilon)}$$

Частные случаи. Пусть $g(x) = |x|^r, r > 0$, тогда

$$P(|\xi| \geq \epsilon) \leq \frac{M|\xi|^r}{\epsilon^r}$$

При $r = 2$ имеем классическое неравенство Чебышева

$$P(|\xi| \geq \epsilon) \leq \frac{M\xi^2}{\epsilon^2}$$

Если ξ — центрированная случайная величина, $\xi = \eta - M\eta$, то

$$P(|\eta - M\eta| \geq \varepsilon) \leq \frac{D\eta}{\varepsilon^2}$$

или

$$P(|\eta - M\eta| < \varepsilon) > 1 - \frac{D\eta}{\varepsilon^2}.$$

Неравенство Колмогорова. Если случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимые и имеют ограниченную дисперсию $D\xi_k < \infty$, то при любом $\varepsilon > 0$

$$P\left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k (\xi_i - M\xi_i) \right| \geq \varepsilon\right) \leq \frac{\sum_{i=1}^n D\xi_i}{\varepsilon^2}.$$

Замечание. Неравенство Колмогорова при $n = 1$ совпадает с неравенством Чебышева для центрированной случайной величины, поэтому оно является в определенном смысле обобщением неравенства Чебышева.

8. СХОДИМОСТЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ СЛУЧАЙНЫХ ВЕЛИЧИН

Сходимость по вероятности. Последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n \dots$ сходится по вероятности к случайной величине ξ , если для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\xi_n - \xi| > \varepsilon) = 0.$$

Последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n$ сходится по вероятности к константе a , если для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\xi_n - a| > \varepsilon) = 0.$$

В частности, при $a = 0$ получаем определение сходимости последовательности к нулю по вероятности.

Записывается: $\underset{P}{\xi_n} \rightarrow \xi$, $\underset{P}{\xi_n} \rightarrow a$, $\underset{P}{\xi_n} \rightarrow 0$.

Сходимость в среднем квадратическом. Последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n$ сходится в среднем квадратическом к случайной величине ξ , если $\lim_{n \rightarrow \infty} M|\xi_n - \xi|^2 = 0$.

Замечание. Из сходимости в среднем квадратическом следует сходимость по вероятности. Обратное в общем случае неверно. Нужны некоторые дополнительные условия на последовательность $(\xi_n, n \geq 1)$.

Сходимость с вероятностью единица (почти наверное, почти всюду). Последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n$ с вероятностью единица сходится к случайной величине ξ , если

$$P(\{\omega : |\xi_n(\omega) - \xi(\omega)| \rightarrow 0 \text{ при } n \rightarrow \infty\}) = 1.$$

Записывается: $\xi_n \xrightarrow{пн} \xi$ или $\xi_n \xrightarrow{пв} \xi$.

Слабая сходимость случайных величин. Последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$, имеющих распределения $F_i(x) = P(\xi_i < x)$, слабо сходится к случайной величине ξ с распределением $F(x)$, если $F_i(x) \rightarrow F(x)$ $i \rightarrow \infty$ для всех точек x непрерывности функции $F(x)$.

9. ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ

Законом больших чисел для последовательности $\xi_1, \xi_2, \dots, \xi_n$ называют утверждения о сходимости по вероятности последовательности $\eta_k =$

$$= \frac{1}{k} \sum_{i=1}^k \xi_i, \quad k \geq 1, \text{ составленной из}$$

средних арифметических последовательностей исходной последовательности.

Ниже приведен ряд теорем, формулирующих условия справедливости закона больших чисел.

Достаточные условия выполнения закона больших чисел. *Теорема 1.* Если последовательность независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$, такова, что дисперсии $D\xi_k$ существуют и $\lim_{n \rightarrow \infty} \frac{1}{n} D\xi_n = 0$, то для последовательности $(\xi_n, n \geq 1)$

справедлив закон больших чисел, т. е. для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n M\xi_k \right| \geq \varepsilon \right) = 0.$$

Последовательность

$$\eta_n = \frac{1}{n} \sum_{k=1}^n (\xi_k - M\xi_k)$$

по вероятности сходится к нулю.

Теорема 2. (Закон больших чисел в форме Чебышева).

Если последовательность независимых случайных величин $\{\xi_n, n \geq 1\}$ такова, что существуют дисперсии $D\xi_n$ и они равномерно ограничены, $D\xi_n <$

$< C$ при любом $n \geq 1$, то $\frac{1}{n} \sum_{k=1}^n \xi_k -$

$-\frac{1}{n} \sum_{k=1}^n M\xi_k \rightarrow 0$ по вероятности.

Теорема 3. (Закон больших чисел для последовательности одинаково распределенных случайных величин).

Пусть $\{\xi_n, n \geq 1\}$ — последовательность независимых одинаково распределенных случайных величин, для которых существует математическое ожидание $M\xi_n = a < \infty$, тогда последовательность средних арифметических $\eta_n = \frac{1}{n} \sum_{k=1}^n \xi_k$ по вероятности сходится к a , т. е. при любом $\varepsilon > 0$

$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n \xi_k - a \right| > \varepsilon \right) = 0.$

Теорема 4 (теорема Бернулли). Пусть последовательность $\{v_i, i \geq 1\}$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n \xi_k - a \right| > \varepsilon \right) = 0.$$

есть последовательность независимых одинаково распределенных случайных величин, принимающих два значения 0 и 1,

$P(v_i = 0) = q, P(v_i = 1) = p, p > 0, q > 0, p + q = 1.$

Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n v_i - p \right| \geq \varepsilon \right) = 0.$$

Замечание. Случайную величину v_i можно трактовать как число успехов в i -м испытании, $\sum_{i=1}^n v_i$ есть число успехов в n испытаниях, величина $\frac{1}{n} \sum_{i=1}^n v_i$ есть частота успехов в n испытаниях.

Теорема Бернулли утверждает, что при неограниченном увеличении числа испытаний частота успехов сходится к вероятности успеха p в одном испытании по вероятности.

Усиленным законом больших чисел для последовательности ξ_1, \dots, ξ_n , называют утверждения о сходимости с вероятностью единица последова-

тельности $\eta_k = \frac{1}{k} \sum_{i=1}^k \xi_i$.

Ниже приведены некоторые достаточные условия выполнения усиленного закона больших чисел.

Достаточные условия выполнения усиленного закона больших чисел.

Теорема 5. Пусть $\{\xi_n, n \geq 1\}$ есть последовательность независимых случайных величин, для которой

$\sum_{k=1}^{\infty} \frac{1}{k^2} D\xi_k < \infty$, тогда с вероятностью единица при $n \rightarrow \infty$

$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n M\xi_k \rightarrow 0.$

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n M\xi_k \rightarrow 0.$$

Теорема 6. (Усиленный закон больших чисел для последовательности одинаково распределенных случайных величин.)

Для последовательности $\{\xi_n, n \geq 1\}$ независимых одинаково распределенных случайных величин с конечным математическим ожиданием $M\xi_n = a < \infty$ с вероятностью единица при $n \rightarrow \infty$

$$\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow a.$$

Пример. Для схемы Бернулли (см. теорему 4) случайные величины v_i независимы, одинаково распределены, $Mv_i = p < \infty$, поэтому частота успехов

$$\frac{1}{n} \sum_{k=1}^n v_k$$

с вероятностью единица сходится к вероятности успеха p при $n \rightarrow \infty$ (неограниченное увеличение числа испытаний).

10. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Локальная теорема (нормальное приближение распределения Бернулли). Для последовательности независимых одинаково распределенных случайных величин v_i ($i = 1, 2, \dots$), принимающих два значения 0 или 1 с вероятностями $p = P(v_i = 1)$, $q = P(v_i = 0)$, $p + q = 1$, $0 < p < 1$, при $\left| \frac{m - np}{\sqrt{npq}} \right| < c < \infty$ имеет место предельное равенство

$$\lim_{n \rightarrow \infty} \frac{\sqrt{npq} P\left(\sum_{k=1}^n v_k = m\right)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{(m - np)^2}{2npq}}} = 1.$$

Замечание. Так как $P\left(\sum_{i=1}^n v_i = m\right) = C_n^m p^m q^{n-m}$, то приведенную выше теорему можно рассматривать как приближение величин $C_n^m p^m q^{n-m}$ при $n \rightarrow \infty$ величинами

$$\frac{1}{\sqrt{2\pi npq}} e^{-\frac{(m - np)^2}{2npq}}$$

Указанное приближение справедливо не только в области $\left| \frac{m - np}{\sqrt{npq}} \right| < c < \infty$, но и в более широкой области изменения параметра m , а именно $|m - np| = o(npq)^{2/3}$.

Интегральная теорема Муавра—Лапласа. Если $S_n = \sum_{i=1}^n v_i$, где v_i $i = 1, 2, \dots$ — последовательность независимых случайных величин, определенных выше, и $0 < p < 1$, то для любых $-\infty \leq a < b < +\infty$

$$\lim_{n \rightarrow \infty} \left[P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy \right] = 0.$$

Теорема Пуассона. Если при $n \rightarrow \infty$ $p = p(n) \rightarrow 0$, причем $np(n) \rightarrow a$, то при любом m

$$\lim_{n \rightarrow \infty} P\left(\sum_{k=1}^n v_k = m\right) = \frac{a^m}{m!} e^{-a}.$$

Замечание. Теорема Пуассона дает приближение распределения случайных величин $S_n = \sum_{k=1}^n v_k$ в области малых значений вероятности $p = P(v_i = 1)$.

Центральная предельная теорема для сумм одинаково распределенных слагаемых. Если независимые случайные величины ξ_1, ξ_2, \dots одинаково распределены и имеют конечную отличную от нуля дисперсию, то при $n \rightarrow \infty$ равномерно по x

$$P\left(\frac{1}{\sqrt{D\xi_1 n}} \sum_{k=1}^n (\xi_k - M\xi_k) < x\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

Центральная предельная теорема для сумм произвольно распределенных слагаемых (достаточность условия Линдберга). Если последовательность независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ с распределениями $F_n(x)$ при любом x удовлетворяет условию Линдберга

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{\substack{|x - M\xi_k| > \\ > \tau B_n}} (x - M\xi_k)^2 \times \\ \times dF_k(x) = 0, \quad (14)$$

где $B_n^2 = D\left(\sum_{k=1}^n \xi_k\right)$, то при $n \rightarrow \infty$ равномерно по x

$$P\left(\frac{1}{B_n} \sum_{k=1}^n (\xi_k - M\xi_k) < x\right) \rightarrow \\ \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

Замечание. Условие Линдберга (14) означает равномерную малость каждого слагаемого, входящего в предельную сумму, т. е.

$$P\left(\max_{1 \leq k \leq n} |\xi_k - M\xi_k| \geq \tau B_n\right) \rightarrow 0 \\ \text{при } n \rightarrow \infty.$$

Теорема Ляпунова. Если для некоторого $\delta > 0$ выполняется условие Ляпунова

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M|\xi_k - M\xi_k|^{2+\delta} \rightarrow 0 \\ \text{при } n \rightarrow \infty, \quad (15)$$

то для последовательности независимых случайных величин при $n \rightarrow \infty$ равномерно по x справедливо предельное равенство

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{B_n} \sum_{k=1}^n (\xi_k - M\xi_k) < x\right) = \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

1. ОСНОВНЫЕ ПОНЯТИЯ И КЛАССИФИКАЦИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

Случайным процессом $\{X_t, t \in T\}$ называют семейство случайных величин, определенных на некотором вероятностном пространстве $\{\Omega, \mathfrak{B}, P\}$, $\Omega = \{\omega\}$.

Пример 1. Пусть $\omega = (\omega_0, \omega_1, \dots, \omega_{2k})$ — последовательность значений стандартных случайных величин z_i , $i = 0, \dots, 2k$, т. е. взаимно независимых равномерно распределенных на $[0, 1]$ случайных величин, а $C_0, C_1, \dots, C_{2k}, \lambda_1, \dots, \lambda_k$ — набор вещественных чисел. Каждому вещественному числу t сопоставим случайную величину

$$X_t = C_0 Z_0 + C_1 Z_1 \sin \lambda_1 t + C_2 Z_2 \cos \lambda_1 t + C_{2k-1} Z_{2k-1} \sin \lambda_k t + C_{2k} Z_{2k} \cos \lambda_k t. \quad (1)$$

Случайные величины $X_t, t \in R^1$ определяются последовательностью случайных взаимно независимых величин $\{Z_i\}$ и, следовательно, определены на одном вероятностном пространстве $\Omega = \{\omega\}$. Если характер зависимости между Z_i не определен, то (1) нельзя считать набором случайных величин.

Пример 2. Пусть $\{Z_t, t = 0, 1, 2, \dots\}$ — последовательность стандартных случайных величин, $\psi(x, y)$ — неслучайная функция двух переменных. Последовательность случайных величин $X_0 = \psi(0, Z_0), X_1 = \psi(x_0, Z_1), \dots, X_{t+1} = \psi(X_t, Z_{t+1})$ является семейством случайных величин и, следовательно, случайным процессом $\{X_t, t \in T = \{0, 1, 2, \dots\}\}$.

Семейство случайных величин $X_t = X(t, \omega)$, составляющих случай-

ный процесс, можно рассматривать как одну функцию двух переменных — времени t и элементарного события ω , определяющего значения всех случайных величин этого семейства. Если ω фиксировано, то значения $X(t, \omega)$, когда t пробегает значения из T , называют реализацией случайного процесса. В примере 1 реализациями являются тригонометрические полиномы от t , в примере 2 — последовательности вещественных чисел.

Множество \mathcal{E} значений случайных величин X_t , составляющих случайный процесс, называют пространством состояний. Множество значений T называют временем. Время называют дискретным, если $T \subseteq \{0, \pm 1, \pm 2, \dots\}$. Если $\mathcal{E} = R^1$, а $T = \{(k, l), k, l = 0, \pm 1, \pm 2, \dots\}$, то соответствующий случайный процесс X_t называют вещественным случайным полем на двумерной решетке. Если $\mathcal{E} = R^1, T = R^1$, то X_t — вещественный случайный процесс с непрерывным временем. Если $\mathcal{E} = R^1, T = \{0, \pm 1, \pm 2, \dots\}$, то X_t называют вещественным случайным процессом с дискретным временем, и т. д. Задание пар \mathcal{E}, T определяет различные классы случайных процессов. В ряде задач удобно рассматривать комплекснозначные случайные процессы, у которых множество состояний $\mathcal{E} = Z$ — множество комплексных чисел. Для класса ветвящихся случайных процессов T — так называемые деревья (см. гл. 3).

С каждым случайным процессом $\{X_t, t \in T\}$, реализации которого непрерывны справа, можно связать наборы (наблюдаемых) событий. Пусть $T \subseteq R^1$, а \mathfrak{F}_t — минимальная σ -алгебра (см. гл. 1) событий, порожденная событиями $\{\omega: X_s = X(s, \omega) \in A\}$,

где A — измеримое подмножество в \mathcal{E} , а $s \leq t$. Например, если $\mathcal{E} = R^1$, то в качестве A можно взять $(-\infty, a]$. В таком случае \mathfrak{F}_t порождается событиями $\{\omega: X_s = X(s, \omega) \leq a\}$. Так как число событий в \mathfrak{F}_t может лишь увеличиться с ростом t , то $\mathfrak{F}_{t_1} \subseteq \mathfrak{F}_{t_2}$ для $t_1 > t_2$, $\mathfrak{F}_t = \mathfrak{F}_{t+} = \lim_{h \downarrow 0} \mathfrak{F}_{t+h}$.

Для любого t $\mathfrak{F}_t \subseteq \mathfrak{B}$. Последовательность \mathfrak{F}_t называют потоком событий, порожденным случайным процессом X_t .

Пусть $\{X_t, Y_t, t \in T\}$ — случайные процессы, определенные на общем вероятностном пространстве $\{\Omega, \mathfrak{B}, P\}$. Случайный процесс Y_t называют согласованным с X_t , если для любого t событие $\{Y_t \in A\} \in \mathfrak{F}_t$, где A — измеримое множество в \mathcal{E}' — пространстве состояний процесса Y_t , \mathfrak{F}_t — поток событий, порожденный X_t .

Пусть $T \subseteq R^1$. Случайную величину $\tau = \tau(\omega)$ со значениями из T называют моментом остановки, если для любого t событие $\{\omega: \tau = \tau(\omega) \leq t\} \in \mathfrak{F}_t$.

2. МАРКОВСКИЕ ЦЕПИ

Случайный процесс $\{X_t, t \in T\}$ называют марковской цепью, если:

а) множество состояний является не более чем счетным $\mathcal{E} = \{E_1, E_2, \dots\}$, ниже используется запись $E_i = i$;

б) время $T = \{0, 1, 2, \dots\}$ или $T = \{0, \pm 1, \pm 2, \dots\}$;

в) для любых последовательных моментов времени $t, t+1$ условные вероятности для всех $i, j \in \mathcal{E}$;

$$P\{X_{t+1} = j | X_s = i_s, s \leq t\} = P\{X_{t+1} = j | X_t = i_t\}. \quad (2)$$

$$\sum_{t \in T} P(X_t = j) \neq 0.$$

Если число состояний \mathcal{E} конечно (равно $l < \infty$), то марковскую цепь называют конечной марковской цепью. В противном случае ($l = \infty$) ее называют счетной марковской цепью.

Условные вероятности $P\{X_{t+1} = j | X_t = i\} = p_{i,j}(t, t+1)$ называют переходными вероятностями. Если они не зависят от t ($p_{i,j}(t, t+1) = p_{i,j}$), то марковскую цепь

называют однородной (по времени). Переходные вероятности $p_{i,j}$ можно записать в виде матрицы переходных вероятностей $P = (p_{i,j})$, где i — номер строки, j — номер столбца. Матрица P является стохастической, т. е. все ее элементы $p_{i,j}$ таковы, что

$$0 \leq p_{i,j} \leq 1 \text{ и } \sum_{j=1}^l p_{i,j} = 1, \quad l \leq \infty.$$

Реализацию однородной марковской цепи $X_t = X(t, \omega)$ с матрицей переходных вероятностей P можно получить на основе последовательности стандартных случайных величин $\{Z_t\}$ и соответствующим образом подобранной функции $\psi(x, y)$, где $y \in [0, 1]$, $x \in \mathcal{E} = \{1, 2, \dots, l\}$. Полагаем $\psi(i, y) = j$, если

$$p_{i,0} + p_{i,1} + \dots + p_{i,j-1} \leq y < p_{i,1} + \dots + p_{i,j}, \quad j = 1, \dots, l, \\ p_{i,0} = 0, \quad X_{t+1} = \psi(X_t, Z_t).$$

Если X_t — однородная марковская цепь, событие $B_t \in \mathfrak{F}_t$, то $P(X_{t+1} = j | B_t, X_t = i) = p_{i,j}$. Для любого момента остановки

$$P\{X_{\tau+1} = j | B_\tau, X_\tau = i\} = p_{i,j}. \quad (3)$$

Соотношение (2) означает, что для марковской цепи при фиксированном «настоящем» события в «прошлом» и в «будущем» взаимно независимы. Соотношение (3) означает, что это свойство независимости событий для однородных марковских цепей верно и относительно случайного момента остановки. Кратко такое свойство называют строгой марковостью.

Анализ эволюции марковской цепи упрощается, если рассматривается ее граф (рис. 1), вершинам которого соответствуют состояния $i \in \mathcal{E}$. Если за один шаг возможен переход из состояния i в состояние j ($p_{i,j} > 0$), то вершины i и j соединяются ребром со стрелкой, ведущей из i в j . Замкнутая стрелка из i в i указывает положительную вероятность перехода из i в i за один шаг, $p_{i,i} > 0$. Около ребер, соединяющих i с j , пишутся вероятности $p_{i,j}$.

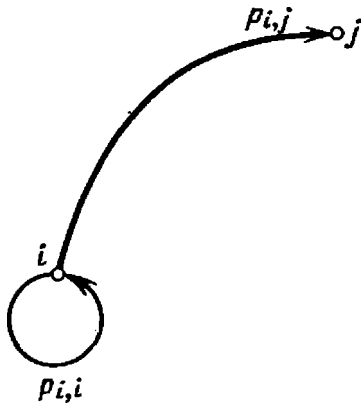


Рис. 1. Фрагмент графа марковской цепи

Вероятность $p_{i,j}(s)$ перехода из i в j за s шагов есть (i, j) -й элемент матрицы $P(s) = P^s$, т. е. элемент в i -й строке и j -м столбце s -й степени матрицы переходных вероятностей P .

Для возведения матриц P в степень можно использовать спектральное разложение, а также формулу Перрона. Для марковской цепи с двумя состояниями

$$P = \begin{pmatrix} 1 - c_1 & c_1 \\ c_2 & 1 - c_2 \end{pmatrix}, \quad c_1 + c_2 \neq 0,$$

$$P^s = \frac{1}{c_1 + c_2} \begin{pmatrix} c_2 & c_1 \\ c_2 & c_1 \end{pmatrix} + \frac{(1 - c_1 - c_2)^s}{c_1 + c_2} \begin{pmatrix} c_1 & -c_1 \\ -c_2 & c_2 \end{pmatrix}.$$

Вероятности всех событий, связанных с однородной марковской цепью, определяются набором начальных вероятностей $p_i = P\{X_0 = i\}$, $i = 1, \dots, l$ и матрицей переходных вероятностей P .

Однородную марковскую цепь называют **стационарной**, если вероятности $P\{X_{t+1} = i_1, \dots, X_{t+k} = i_k\}$ не зависят от t при любом наборе состояний i_1, \dots, i_k и любом $k = 1, 2, \dots$

Если вектор вероятностей начальных состояний $p^T = (p_1, p_2, \dots, p_l, \dots)$, $\sum_i p_i = 1$ удовлетворяет соотношению

$$p^T P = p^T \left(\sum_i p_i p_{ij} = p_j, j \in \mathcal{E} \right),$$

то соответствующая p, P марковская цепь X_t является стационарной, т. е.

$P\{X_t = i\} = p_i$ для любого $t \in \{0, 1, 2, \dots\}$.

Замечание. Для некоторых марковских цепей со счетным множеством состояний одинаковую матрицу P могут иметь различные стационарные марковские цепи.

Конечный набор состояний $i, i_1, \dots, i_{s-1}, j$ называют **путем** длины s из состояния i в состояние j , если $p_{i, i_1} p_{i_1, i_2} \dots p_{i_{s-1}, j} > 0$. Состояние j называют **достижимым** из состояния i ($i \rightarrow j$), если существует путь из состояния i в состояние j . Два состояния i и j называют **сообщающимися**, если j достижимо из состояния i и состояние i достижимо из состояния j ($i \leftrightarrow j$). Замкнутый путь из состояния i в i называют **циклом** состояния i . Длину этого пути называют **длиной цикла**. Если $p_{i, i}(n) > 0$, то существует цикл длины n . **Периодом** состояния i называют общий наибольший делитель (о. н. д.) $d(i)$ тех чисел $n \geq 1$, для которых $p_{i, i}(n) > 0$.

Подмножество состояний $\mathcal{E}' \subseteq \mathcal{E}$ называют **блоком** (классом) **сообщающихся** состояний, если любые два состояния из \mathcal{E}' являются **сообщающимися**. Множество всех состояний \mathcal{E} разбивается на непересекающиеся блоки \mathcal{E}_l **сообщающихся** состояний, $\mathcal{E} = \mathcal{E}_1 + \dots + \mathcal{E}_r, r \leq \infty$. Марковскую цепь называют **неразложимой**, если все \mathcal{E} являются одним блоком. Блок \mathcal{E}_l называют **замкнутым**, если для любого состояния $i \in \mathcal{E}_l$ переходные вероятности $p_{i, j} > 0$ только для состояний $j \in \mathcal{E}_l$. Если замкнутый блок содержит только одно состояние, то это состояние называют **поглощающим**. Состояние i называют **существенным**, если из любого **сообщающегося** с i состояния j есть путь, ведущий из j в i . В противном случае состояние i называют **несущественным**.

Неразложимую марковскую цепь называют **непериодичной**, если период каждого состояния равен единице. Если период каждого состояния равен $d > 1$, то неразложимую марковскую цепь называют **периодичной** с периодом d . В общем случае все состояния в замкнутом блоке имеют один и тот же

период d , который называют периодом замкнутого блока. Блок называют **периодичным**, если $d > 1$. Все состояния \mathcal{E} неразложимой периодичной марковской цепи с периодом d разбиваются на непересекающиеся **циклические классы** C_h , $\mathcal{E} = C_0 + C_1 + \dots + C_{d-1}$. Эволюция такой цепи соответствует переходу из класса C_l в класс C_{l+1} , $l \neq d-1$ и переходу из C_{d-1} в класс C_0 .

Пример 3. Периодичная неразложимая цепь с множеством состояний $\mathcal{E} = \{1, 2, 3\}$ определяется матрицей переходных вероятностей

$$P = \begin{pmatrix} 0 & 1 & 0 \\ q & 0 & p \\ 0 & 1 & 0 \end{pmatrix}.$$

Граф этой цепи показан на рис. 2.

Состояния любого блока либо все существенные, либо все несущественные. Блок замкнут только тогда, когда все его состояния существенные. В конечной марковской цепи есть хотя бы один замкнутый блок. Если в конечной марковской цепи сначала последовательно нумеруются состояния первого замкнутого блока, второго замкнутого блока и т. д., а затем последовательно нумеруются состояния каждого из незамкнутых блоков, то соответствующая такой нумерации матрица переходных вероятностей принимает блочный вид, называемый **каноническим**.

Конечную марковскую цепь называют **поглощающей**, если все замкнутые блоки содержат только по одной точке. Для такой цепи множество всех состояний $\mathcal{E} = A + B$, где A — множество несущественных состояний, а B — множество поглощающих состояний. Канонический вид матрицы переходных вероятностей

$$P = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix}, \quad \text{где } I = \begin{pmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1 \end{pmatrix} —$$

единичная матрица, соответствующая поглощающим состояниям, R образована $p_{i,j}$, $i \in A, j \in B$, Q образована $p_{i,j}$, $i, j \in A$, 0 — нулевая матрица.

Пусть для конечной поглощающей марковской цепи: τ_i — момент ухода из множества несущественных состояний A , если начальное состояние i ,

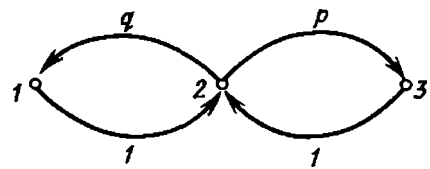


Рис. 2. Граф периодической неразложимой цепи с тремя состояниями

$N_{i,j}$ — число попаданий в несущественное состояние j , когда начальное состояние i . Тогда для некоторого $0 < d < 1, P\{\tau_i > t\} < Cd^t, M\tau^n < \infty$ для любого $n, m_{i,j} = MN_{i,j} < \infty$ для любых $i, j \in A$. Матрицу $M = (m_{i,j})$ находят из соотношения

$$M = (I - Q)^{-1}. \quad (4)$$

Пусть $b_{i,j}$ — вероятность, выйдя из $i \in A$, поглотится в $j \in B, V = (b_{i,j})$. Тогда

$$V = MR. \quad (5)$$

Пусть $u_{i,j}$ — вероятность, выйдя из $i \in A$, хотя бы один раз попадет в $j \in A, U = (u_{i,j})$. Тогда

$$U = (M - I)(M_{dg})^{-1}, \quad (6)$$

где $M_{dg} = (m_{ij}\delta_{ij}), \delta_{ij} = 0, i \neq j, \delta_{ii} = 1$.

Пусть $v_{ij} = \text{var}(N_{i,j}) = MN_{ij}^2 - (MN_{ij})^2$ — дисперсии $N_{ij}, V = (v_{ij}), M_{sq} = (m_{ij}^2)$. Тогда

$$V = 2MM_{dg} - M - M_{sq}. \quad (7)$$

Конечную непериодичную марковскую цепь, у которой все состояния образуют один замкнутый блок, называют **регулярной**.

Если у конечной марковской цепи все состояния сообщающиеся, а у одного из состояний общий наибольший делитель всех длин циклов равен единице, то эта марковская цепь регулярная.

У регулярной конечной марковской цепи для любых i, j существуют пре-

делы $\lim_{t \rightarrow \infty} p_{ij}(t) = a_j > 0, \sum_{j=1}^l a_j = 1$.

Пусть $A = \begin{pmatrix} a_1 & \dots & a_l \\ \dots & \dots & \dots \\ a_1 & \dots & a_l \end{pmatrix}$ — матрица размера $l \times l$, у которой все l строк одинаковы и равны $a^T = (a_1, \dots, a_l)$,

$|\mathcal{E}| = l$: a_j^{-1} равно среднему числу шагов до первого возвращения в j , если $X_0 = j$

$$AP = PA = A, A^m P^n = A \text{ для } m, n = 1, 2, \quad (8)$$

Если $a^T = (a_1, \dots, a_l)$ взять в качестве начального распределения, то регулярная марковская цепь X_t стационарна. Вектор a^T является единственным решением уравнений

$$a^T P = a^T, a^T 1 = \sum_{i=1}^l a_i = 1. \quad (9)$$

Если P соответствует регулярной конечной марковской цепи, то существует матрица

$$Z = (I + (P - A))^{-1}. \quad (10)$$

Пусть $f(i)$ — функция на множестве состояний E . Последовательность $\{f(X_t)\}$, $t = 0, 1, 2, \dots$ называют последовательностью случайных величин, связанных с марковской цепью X_t . Последовательность сумм $Y_t = \sum_{s=1}^t f(X_s)$ называют процессом накопления на марковской цепи, соответствующим функции $f(i)$, $i \in \mathcal{E}$.

Пусть $f^T = (f(1), \dots, f(l))$, $l < \infty$.

Для регулярной конечной марковской цепи

$$MY_t = (Af) t + (Z - I) f + 0(1), t \rightarrow \infty; \quad (11)$$

$$\text{var } Y_t = MY_t^2 - (MY_t)^2 = (f^T C f) t + 0(t), t \rightarrow \infty, \quad (12)$$

где матрица $C = (c_{ij}) = A_{dg} Z + Z^T A_{dg} - A_{dg} - a a^T$, $f^T C f = \sum_{i,j=1}^l f(i) c_{ij} f(j)$,

$a a^T = (a_i a_j)$ — матрица размера $l \times l$.

Теория марковских цепей со счетным множеством состояний \mathcal{E} существенно сложнее. Полезна следующая классификация состояний $i \in \mathcal{E}$. Пусть $u_{i,i}$ — вероятность того, что марковская цепь X_t побывает хотя бы при одном значении $t \geq 1$ в состоянии i , если $X_0 = i$.

Состояние i называют возвратным, если $u_{i,i} = 1$.

Состояние i называют невозвратным, если $u_{i,i} < 1$.

Состояние i возвратное тогда и только тогда, когда

$$\sum_{t=1}^{\infty} p_{i,i}(t) < \infty. \quad (13)$$

Пусть $\tau_{i,j}$ — число шагов до первого попадания в состояние j , когда i — начальное состояние, т. е. $\tau_{i,j} = \min \{n \mid X_n = j, \text{ если } X_0 = i\}$. Тогда $u_{i,i} = P \{\tau_{i,i} < \infty\}$. Среднее время до первого возвращения в состояние i равно $\bar{m}_{i,i} = M\tau_{i,i}$.

Состояние i называют положительным возвратным, если $u_{i,i} = 1$, $\bar{m}_{i,i} < \infty$. Состояние i называют нулевым возвратным, если $u_{i,i} = 1$, но $\bar{m}_{i,i} = \infty$.

В неразложимой марковской цепи либо все состояния положительные возвратные, либо все состояния нулевые возвратные, либо все состояния невозвратные.

Последовательность $\mu(i) \geq 0, i \in \mathcal{E}$ определяет меру $\mu(A) = \sum_{i \in A} \mu_i$,

$A \subseteq \mathcal{E}$. Мера μ — конечная, если $\mu(\mathcal{E}) < \infty$. Пусть $\mathfrak{B}(\mathcal{E}) = \{\mu : \mu(\mathcal{E}) < \infty\}$ — множество конечных мер, а $\mathfrak{F}(\mathcal{E}) = \left\{ f \mid \sup_{i \in \mathcal{E}} |f(i)| < \infty \right\}$ —

множество ограниченных функций. Степени матриц переходных вероятностей $P^t, t = 1, 2, \dots$ можно рассматривать как полугруппы операторов, действующих в $\mathfrak{B}(\mathcal{E})$ и $\mathfrak{F}(\mathcal{E})$:

$$f_t = P^t f, \mu_t^T = \mu^T P^t, \quad (14)$$

где f, f_t — векторы-столбцы, $f^T = (f(1), \dots, f(i), \dots)$, а μ^T — векторы-строки, $\mu^T = (\mu(1), \dots, \mu(i), \dots)$.

Функцию f_0 и меру μ_0 называют гармоническими, если

$$P f_0 = f_0, \mu_0^T = \mu_0^T P. \quad (15)$$

Исследование полугрупп, аналогичных (14), дает возможность получать важные результаты в теории марковских процессов.

Марковский процесс X_t с непрерывным временем $t \geq 0$ и не более чем счетным множеством состояний $\mathcal{E} = \{1, 2, \dots\}$ характеризуется свой-

ством зависимости условных вероятностей состояний в моменты $t > s$ лишь от последнего состояния i_s в момент s , т. е.

$$P\{X_t = j | X_u = i_u, u \leq s\} = P\{X_t = j | X_s = i_s\}.$$

Переходные вероятности $P_{ij}(s, t) = P\{X_t = j | X_s = i\}$, $s < t$, удовлетворяют уравнениям Колмогорова—Чепмена

$$P_{ik}(s, u) = \sum_{j \in \mathcal{E}} P_{ij}(s, t) P_{jk}(t, u),$$

$s < t < u$.

Если при $h \rightarrow 0$ равномерно по i $P_{jk}(t, t+h)/h \rightarrow c_j(t) p_{jk}(t)$, $[1 - P_{kk}(t, t+h)]/h \rightarrow c_k(t)$,

где $p_{jk}(t)$ непрерывны по t , $p_{jj} = 0$, $\sum_{k \in \mathcal{E}} p_{jk}(t) = 1$, то система прямых

уравнений имеет вид

$$\frac{\partial P_{ik}(s, t)}{\partial t} = -c_k(t) P_{ik}(s, t) + \sum_{j \in \mathcal{E}} P_{ij}(s, t) c_j(t) p_{jk}(t), \quad i, k \in \mathcal{E}.$$

Числа $p_{jk}(t)$ являются условными вероятностями скачка в момент t из состояния j в состояние k , $c_k(t) dt$ трактуют как вероятность скачка в бесконечно малом интервале времени $(t, t+dt)$.

Для расчета вероятностей $Q_{ij}(s, t)$ первого вхождения используется система обратных уравнений

$$\frac{\partial Q_{ik}(s, t)}{\partial s} = c_i(s) Q_{ik}(s, t) - c_i(s) \times \sum_{j \in \mathcal{E}} p_{ij}(s) Q_{js}(s, t).$$

Марковский процесс X_t с непрерывным временем $t \geq 0$ называют однородным, если $P_{ij}(s, t) = P_{ij}(t-s)$. Для однородного марковского процесса $c_j(t) = c_j$, $p_{jk}(t) = p_{jk}$, $p_{jj} = 0$.

Если все $p_{jk} = 0$, $k \neq j+1$, то однородный марковский процесс X_t называют процессом размножения и гибели.

Для однородного марковского процесса X_t переходные вероятности $P_{ik}(t) = P\{X_t = k | X_0 = i\}$ являются решениями системы прямых дифференциальных уравнений с постоянными коэффициентами

$$\frac{dP_{ik}(t)}{dt} = -c_k P_{ik}(t) + \sum_{j \in \mathcal{E}} P_{ij}(t) c_j p_{jk}, \quad i, k \in \mathcal{E}$$

с начальными условиями $P_{ii}(0) = 1$, $P_{ik}(0) = 0$, $k \neq i$.

Вероятности $Q_{ik}(t)$ попадания траектории марковского однородного процесса в состояние k за время $(0, t)$, когда i — начальное состояние, находятся как решение системы обратных дифференциальных уравнений с постоянными коэффициентами

$$\frac{dQ_{ik}(t)}{dt} = -c_i Q_{ik}(t) + c_i \sum_{j \in \mathcal{E}} p_{ij} Q_{jk}(t)$$

с начальными условиями $Q_{kk}(0) = 1$, $Q_{ik}(0) = 0$, $i \neq k$ и граничным условием $Q_{kk}(t) \equiv 1$.

Если марковский процесс принимает бесконечное число состояний, то бесконечная система прямых (обратных) уравнений может иметь много разных решений. Этим решениям соответствуют различающиеся друг от друга марковские процессы.

3. ПРОЦЕССЫ ВОССТАНОВЛЕНИЯ. СЛУЧАЙНЫЕ ТОЧЕЧНЫЕ ПРОЦЕССЫ

Пусть X_1, X_2, \dots — последовательность неотрицательных взаимно независимых случайных величин с функциями распределения (ф. р.)

$$G(s) = P\{X_1 \leq s\}, \quad F(s) = P\{X_n \leq s\}, \quad n = 2, 3,$$

Рассмотрим на полупрямой $R_+^1 = [0, \infty)$ последовательность случайных точек с координатами $s_1 = X_1$, $s_2 = X_1 + X_2$, $s_n = X_1 + \dots + X_n$ (рис. 3).

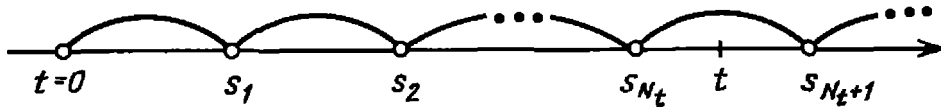


Рис. 3. Процесс восстановления

Случайную последовательность $\{s_n\}$ называют (**общим**) **процессом восстановления**, а точки с координатами $\{s_n\}$ называют **моментами (временами) восстановления**. Случайную величину N_t , равную наибольшему числу n , при котором $s_n \leq t$, называют **числом восстановлений** на интервале времени $[0, t]$,

$$N_t = \max \{n: s_n \leq t\}.$$

Пример 1. Допустим, что имеется неограниченное число исправных однотипных элементов $\{O_n\}$. В момент $t=0$ начинает использоваться элемент O_1 . Через время X_1 элемент O_1 отказывает и мгновенно заменяется элементом O_2 , который через время X_2 , т. е. в момент $s_2 = X_1 + X_2$, мгновенно заменяется элементом O_3 и т. д. Если считать продолжительности безотказной работы $\{X_n\}$ элементов $\{O_n\}$ взаимно независимыми случайными величинами, то моменты замен образуют процесс восстановления.

Если $G(t) = F(t)$, то математическое ожидание числа восстановления $H(t) = MN_t$,

рассматриваемое как функция t , называется **функцией восстановления**. Функция восстановления является единственным ограниченным решением интегрального уравнения

$$H(t) = F(t) + \int_0^t H(t-s) dF(s). \quad (16)$$

В общем случае, когда $G(s) \neq F(s)$, среднее число восстановлений $H_G(t) = MN_t$ будет

$$H_G(t) = G(t) + \int_0^t G(t-s) dH(s). \quad (17)$$

Функции $H(t)$ и $H_G(t)$ полностью определяют процесс восстановления. Преобразования Лапласа — Стильтеса

$$\begin{aligned} \tilde{F}(x) &= \int_0^\infty e^{-xs} dF(s), \quad \tilde{G}(x) = \\ &= \int_0^\infty e^{-xs} dG(s), \end{aligned}$$

$$\begin{aligned} \tilde{H}(x) &= \int_0^\infty e^{-xs} dH(s), \quad \tilde{H}_G(x) = \\ &= \int_0^\infty e^{-sx} dH_G(s) \end{aligned}$$

удовлетворяют соотношениям

$$\tilde{H}(x) = \frac{\tilde{F}(x)}{1 - \tilde{F}(x)}; \quad \tilde{H}_G(x) = \frac{\tilde{G}(x)}{1 - \tilde{F}(x)}. \quad (18)$$

Используя обращение преобразования Лапласа — Стильтеса, можно получить выражения для $H(t)$, $H_G(t)$.

Случайную величину $\alpha_t = s_{N_t+1} - t$ называют **остаточным временем** (до восстановления). Функция распределения $A_t(s) = P\{\alpha_t \leq s\}$ имеет вид

$$\begin{aligned} A_t(s) &= G(t+s) - \int_0^t \tilde{F}(t+s-u) \times \\ &\times dH_G(u). \end{aligned} \quad (19)$$

Пусть $m_G = MX_1 < \infty$, $m_F = MX_n < \infty$, $n \geq 1$. Тогда математическое ожидание остаточного времени

$$M\alpha_t = m_G + m_F H_G(t) - t. \quad (20)$$

Пример 2. Пусть $G(s) = F(s) = 1 - \exp\{-\lambda s\}$. Тогда процесс восстановления является пуассоновским точечным процессом. В этом случае $H(t) = \lambda t$, $A_t(s) = 1 - \exp(-\lambda s)$

Пусть ниже $F(s)$ — нерешетчатая ф. р. Тогда существует предел

$$\lim_{t \rightarrow \infty} M[N_{t+h} - N_t] = \frac{h}{m_F}.$$

(Этот результат называют теоремой Блекуэлла.)

Если $Q(s)$ — интегрируемая на $[0, \infty)$ функция ограниченной вариации, то существует предел

$$\begin{aligned} \lim_{t \rightarrow \infty} \int_0^t Q(t-s) dH_G(s) &= \\ &= \frac{1}{m_F} \int_0^\infty Q(s) ds. \end{aligned} \quad (21)$$

(Этот результат называют узловой теоремой.)

Если второй момент $m_{F,2} = \int_0^\infty s^2 dF(s) < \infty$, то при $t \rightarrow \infty$

$$H(t) = \frac{t}{m_F} + \frac{m_{F,2}}{2m_F^2} - 1 + o(1). \quad (22)$$

Если $G(s) = A(s)$, где $A(s) = \frac{1}{m_F} \int_0^s \bar{F}(u) du$, то процесс восстановлений стационарен на $[0, \infty)$.

В этом случае $A_t(s) = A(s)$, $t \geq 0$, $H_A(t) = t/m_F$, дисперсия числа восстановлений за время t

$$\text{Var}(N_t) = \frac{\sigma_F^2}{m_F^3} t + o(t), \quad t \rightarrow \infty, \quad (23)$$

где $\sigma_F^2 = m_{F,2} - m_F^2$ — дисперсия X_n , $n \geq 2$.

Под случайным точечным процессом понимают последовательность случайных точек на прямой (на плоскости, в пространстве и т. п.). Точечному процессу на прямой соответствует последовательность случайных величин $\dots t_{-1} < t_0 \leq 0 < t_1 < t_2 < \dots$. Каждой случайной величине t_i ставится в соответствие случайное целое число $\Phi(t_i)$, называемое кратностью точки t_i . Случайный точечный процесс называют

простым, если все $\Phi(t_i) = 1$. Пусть $\Lambda(B)$ — математическое ожидание числа точек точечного процесса, попавших в подмножество B . Для непересекающихся подмножеств

$$\{B_n\} \quad \Lambda\left(\bigcup_n B_n\right) = \sum_{n=1}^\infty \Lambda(B_n), \quad \text{т. е.}$$

$\Lambda(B)$ можно рассматривать как меру. Эту меру называют моментной мерой случайного точечного процесса.

Случайный точечный процесс на прямой (на плоскости, в пространстве и т. п.) называют пуассоновским точечным процессом с мерой интенсивностей Λ , если: а) для любого набора непересекающихся подмножеств B_n , $n = 1, 2, \dots$ числа точек $\Psi(B_n)$, попавших в B_n , являются взаимно независимыми случайными величинами;

б) $P\{\Psi(B_n) = k\} = \frac{[\Lambda(B_n)]^k}{k!} \times \exp\{-\Lambda(B_n)\}$, $k = 0, 1, \dots$ Для этого процесса $M\Psi(B_n) = \Lambda(B_n)$.

Пример 3. Пусть n есть значение целочисленной случайной величины, имеющей пуассоновское распределение $P\{n = k\} = \frac{(\lambda T)^k}{k!} \exp(-\lambda T)$. На интервал $[0, T]$ «бросим» независимо n точек, распределение координат которых является равномерным на $[0, T]$. Пусть t_1, \dots, t_n — значение координат этих точек. Тогда $\{t_1, \dots, t_n\}$ соответствует реализации пуассоновского точечного процесса с моментной мерой, пропорциональной мере Лебега, с коэффициентом пропорциональности λ .

4. МАРТИНГАЛЫ

Пусть $\{X_t, Y_t, t \in T\}$ — вещественные случайные процессы, определенные на одном вероятностном пространстве $\{\Omega, \mathfrak{F}, P\}$; $\{\mathfrak{F}_t\}$ — поток событий, порожденный случайным процессом $X_s, s \leq t, T = \{0, 1, 2, \dots\}$ или $T = R_+^1 = [0, \infty)$.

Вещественный случайный процесс Y_t называют мартингалом относительно потока событий $\{\mathfrak{F}_t\}$, если:

- а) Y_t согласован (см. параграф 1) с потоком \mathfrak{F}_t ;
- б) для любого t $M|Y_t| < \infty$;

в) условные математические ожидания (см. гл. 5) $M(Y_{t+1} | X_s, s \leq t) = Y_t$ с вероятностью 1.

Приращения $Y_{t+1} - Y_t = Z_t$ называют **мартингалом-разностью** ($T = \{0, 1, 2, \dots\}$).

Если выполнены свойства (i), (ii) и (iii') $M(Y_{t+1} | X_s, s \leq t) \geq Y_t$, то процесс Y_t называют **субмартингалом** относительно потока $\{\mathfrak{F}_t\}$.

Если $\{X_t\}$ — взаимно независимые случайные величины $M|X_t| < \infty$,

$MX_t = 0$, то $Y_t = \sum_{s=1}^t X_s$ есть мар-

тингал относительно собственного потока событий $\{\mathfrak{F}_t\}$, где \mathfrak{F}_t порождено значениями $Y_s, s \leq t, t \in T = \{0, 1, 2, \dots\}$.

Если f — гармоническая функция на множестве состояний \mathcal{E} марковской цепи X_t с матрицей переходных вероятностей P , удовлетворяющая уравнению (15), то $Y_t = f(X_s)$ есть мартингал относительно потока $\{\mathfrak{F}_t\}$, порожденного $X_s, s \leq t$.

Пусть Y_t — мартингал относительно потока $\{\mathfrak{F}_t\}$, а $Z_t = Y_{t+1} - Y_t$ — его мартингал-разность и $MY_t^2 < \infty$. Тогда для сумм $S_{m, t} =$

$= Y_t - Y_{m-1} = \sum_{s=m}^t Z_s$, выполнены соотношения

$$MS_{m, t}^2 = \sum_{s=m}^t MZ_s^2, \quad (24)$$

$$P \left\{ \max_{m \leq s \leq n} |S_{m, s}| > u \right\} \leq \frac{\sum_{s=m}^n MZ_s^2}{u^2}. \quad (25)$$

Пусть Y_t — мартингал относительно потока событий $\{\mathfrak{F}_t\}$, у которого равномерно ограничены вторые моменты, $M|Y_t|^2 \leq C < \infty, t = 1, 2$. Тогда с вероятностью 1 и в среднем квадратичном существует предел

$$\lim_{t \rightarrow \infty} Y_t = Y_\infty, M|Y_\infty|^2 \leq C.$$

Пусть Y_n — число потоков в n -м поколении ветвящегося процесса с производящей функцией $G(s)$, с конечным средним $m_G = G'(1) > 1$

и дисперсией $\sigma_G^2 = G''(1) - m_G + m_G^2$.

Тогда Y_n/m_G^n является мартингалом относительно собственного потока событий \mathfrak{F}_t , порожденного значениями

$Y_s, s \leq t$, причем $M \left| \frac{Y_n}{m_G^n} \right|^2 \leq \sigma_G^2 \times$

$$\times \frac{m_G^2 - m_G + 1}{m_G^2 - m_G}, \quad t = 1, 2, \dots$$

Существует предел отношения $Y_n/m_G^n, n \rightarrow \infty$.

Пусть Y_t — мартингал относительно потока событий $\{\mathfrak{F}_t\}$, а τ — момент остановки (см. параграф 1), причем

- а) $M\tau < \infty$;
- б) $M(|Y_{t+1} - Y_t| | Y_t, s \leq t) \leq C < \infty$ для всех $t = 1, 2, \dots$ и некоторого C . Тогда $M|Y_\tau| < \infty$ и $MY_\tau = MY_1$.

Термин мартингал связывают со следующей азартной игрой, носящей название «мартингал». Два игрока (разбойника) ставят на выигрыш равные суммы v . Первый игрок называет сторону, на которую упадет монета. Второй бросает симметричную монету. Если угадал первый игрок, он выигрывает сумму v . При втором бросании монеты ставки удваиваются и т. д. При n -м бросании монеты ставки равны $2^n v$. Суммарный выигрыш второго разбойника после t -го бросания

$$Y_t = vX_1 + 2vX_2 + \dots + 2^t vX_t,$$

где $X_n = 1$, если первый разбойник не угадал результат бросания, и $X_n = -1$, если угадал. $\{X_n\}$ — последовательность взаимно независимых случайных величин $P\{X_n = 1\} = P\{X_n = -1\} = \frac{1}{2}$. Y_t является

мартингалом относительно потока $\{\mathfrak{F}_t\}$, где \mathfrak{F}_t порождено значениями $X_1, \dots, X_t, MY_t = 0$ при любом t , т. е. игра справедлива. Допустим, что первый разбойник решает продолжить игру неограниченно, а второй тайно решает выйти из игры при первом же выигрыше. Момент τ прекращения игры (вторым разбойником) является моментом остановки, $MY_\tau = v$. Несовпадение $MY_t = 0$ и $MY_\tau = v$ происходит потому, что нарушено условие б).

**5. ПРОЦЕССЫ
С ОРТОГОНАЛЬНЫМИ
ПРИРАЩЕНИЯМИ.
СТОХАСТИЧЕСКИЕ ИНТЕГРАЛЫ.
ВИНЕРОВСКИЙ ПРОЦЕСС**

Случайный процесс $\zeta(t)$ с действительными или комплексными значениями, заданный на некотором подмножестве T числовой оси, называют процессом с ортогональными приращениями, если $M\zeta(t) = 0$ и для любых точек t_1, t_2, t_3, t_4 из T , таких, что $t_1 \leq t_2 \leq t_3 \leq t_4$, случайные величины $\zeta(t_2) - \zeta(t_1)$ и $\zeta(t_4) - \zeta(t_3)$ некоррелированы, т. е. $M(\zeta(t_2) - \zeta(t_1))(\zeta(t_4) - \zeta(t_3)) = 0$.

Пример 1. Пусть $X_s, s = 1, 2, \dots$ — последовательность некоррелированных случайных величин, тогда процесс

$$\zeta(t) = \sum_{s=1}^t X_s$$

является процессом с ортогональными приращениями, $T = \{1, 2, \dots\}$.

Пусть $\zeta(t)$ — процесс с ортогональными приращениями, заданный на некотором отрезке $[0, T]$, и $f(t)$ — непрерывная на $[0, T]$ функция. Пусть $0 = t_0 < t_1 < \dots < t_n = T$ — разбиение отрезка $[0, T]$, такое, что $\max_{i=1, \dots, n} |t_i - t_{i-1}| \rightarrow 0$ при $n \rightarrow \infty$.

Тогда последовательность сумм

$$\sum_{k=1}^n f(t_{k-1})(\zeta(t_k) - \zeta(t_{k-1}))$$

сходится в среднеквадратичном смысле к некоторой случайной величине, которую называют **стохастическим интегралом** от функции f по процессу ζ и обозначают как

$$\int_0^T f(t) d\zeta(t).$$

Такое название объясняется тем, что для стохастического интеграла выполнены все основные свойства обычного интеграла, такие, как линейность, средняя квадратическая непрерывность по пределам интегрирования и т. д.

Пример 2. Стандартным винеровским процессом, или броуновским движением, называют случайный процесс

$W(t), t \geq 0$, такой, что $W(0) = 0$, приращения $W(t) - W(s)$ для любых $s, t, 0 \leq s \leq t$ являются гауссовскими случайными величинами с нулевым средним и дисперсией $t - s$ и для любых непересекающихся интервалов $(s_k, t_k), k = 1, \dots, n$, приращения $W(t_k) - W(s_k), k = 1, \dots, n$, являются независимыми в совокупности случайными величинами. Таким образом, процесс $W(t)$ является не только процессом с ортогональными приращениями, но и процессом с независимыми приращениями.

Известно, что существует такой стандартный винеровский процесс, все реализации (см. параграф 1) которого непрерывны.

Винеровский процесс является предельным для серии случайных блужданий

$$\xi_n(t) = \sum_{k=1}^n \xi_{k,n}$$

где $\xi_{k,n}$ —

смещение на $\pm \Delta x$ с равными вероятностями за время $\Delta t, n = t/\Delta t$ и $(\Delta x)^2/\Delta t = 1$. Поскольку такое случайное блуждание является моделью диффузии, то процесс $W(t)$ называют также **броуновским движением**. Винеровский процесс может быть определен так же, как гауссовский случайный процесс на полупрямой $t \geq 0$ с нулевым математическим ожиданием и ковариационной функцией $R(t, s) = \min(t, s)$.

Пусть $W(t)$ — винеровский процесс. Тогда процесс $W_1(t) = 0$ при $t = 0$ и $W_1(t) = tW(1/t)$ при $t > 0$ также является стандартным винеровским процессом. Винеровским процессом является также процесс $W_2(t) = \sigma W(t/\sigma^2)$ для любого $\sigma > 0$.

Каноническое представление. Если $W(t)$ — винеровский процесс на отрезке $[0, 1]$, то $W(t) = \sum_{k=0}^{\infty} \zeta_k \varphi_k(t)$,

$$W(t) = \sum_{k=0}^{\infty} \zeta_k \varphi_k(t)$$

где ζ_k — независимые гауссовские величины с нулевыми средними и дисперсиями:

$$\text{Var}(\zeta_k) = \left(\frac{\pi}{2}(2k+1)\right)^{-1}, k = 0, 1, \dots, \varphi_k(t) = \sin\left(\frac{\pi}{2} \times (2k+1)t\right).$$

Распределение момента первого достижения. Пусть τ_a — момент первого достижения винеровским процессом точки a , $a \geq 0$. Тогда

$$P(\tau_a \leq t) = 2P(W(t) \geq a) = \frac{a}{\sqrt{2\pi}} \int_0^t s^{-\frac{3}{2}} e^{-\frac{a^2}{2s}} ds. \quad (26)$$

Распределение максимума

$$P\left(\max_{0 \leq s \leq t} W(s) \geq a\right) = P(\tau_a \leq t). \quad (27)$$

Процесс

$$W_a(t) = \begin{cases} W(t) & \text{при } t \leq \tau_a \\ 2a - W(t) & \text{при } t \geq \tau_a \end{cases}$$

также является стандартным винеровским процессом.

Пусть $I(x) = 1$ при $x \geq 0$ и $I(x) = 0$ при $x < 0$. Тогда интеграл

$$\tau_{\text{пр}} = \int_0^t I(W(s)) ds,$$

где интегрирование ведется по каждой реализации в обычном смысле, является временем пребывания процесса $W(t)$ на положительной полуоси. Его функция распределения

$$P(\tau_{\text{пр}} \leq s) = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{t}}. \quad (28)$$

Такое же распределение имеет время τ_M достижения абсолютного максимума реализацией винеровского процесса на отрезке $[0, t]$:

$$P(\tau_M \leq s) = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{t}}. \quad (29)$$

Последнее соотношение носит название **закона арксинуса**.

Закон повторного логарифма. С вероятностью единица имеет место соотношение

$$\overline{\lim}_{h \rightarrow 0} \frac{W(h)}{\sqrt{2h \ln \ln 1/h}} = 1. \quad (30)$$

Взяв теперь винеровский процесс $tW\left(\frac{1}{t}\right)$, получаем, что

$$\overline{\lim}_{t \rightarrow \infty} \frac{W(t)}{\sqrt{2t \ln \ln t}} = 1. \quad (31)$$

С помощью операции стохастического интегрирования по винеровскому процессу можно построить широкий класс так называемых диффузионных процессов. Такими процессами являются, в частности, процессы вида

$$X(t) = X(0) + \int_0^t \sigma(s) dW(s) + \int_0^t a(s) ds,$$

где $\sigma(s)$ — непрерывная неслучайная функция, называемая **коэффициентом диффузии**, а $a(s)$ — интегрируемая функция, **коэффициент переноса**.

6. СТАЦИОНАРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ПРОЦЕССЫ

Случайный процесс $X(t)$ с комплексными или действительными значениями называют **стационарным в узком смысле** или **строго стационарным**, если для любых $n, t_1, t_2, \dots, t_n, t$ распределение вероятностей вектора $(X(t_1), \dots, X(t_n))$ такое же, как и «сдвинутого» вектора $(X(t_1 + t), \dots, X(t_n + t))$. Время t может быть непрерывным или дискретным (целочисленным). В последнем случае случайный процесс называют **случайной последовательностью**.

Процесс $X(t)$ с конечным вторым моментом $MX(t)^2 < \infty$ называют **стационарным (точнее — стационарным в широком смысле)**, если его среднее значение $MX(t)$ не зависит от времени t , а **ковариационная функция** $R(t, s) = MX(t)X(s)$ зависит лишь от разности аргументов: $R(t, s) = R(t - s)$.

Класс строго стационарных процессов с конечным вторым моментом образует подкласс в классе стационарных процессов.

Для любого стационарного процесса существует единственная непрерывная справа и неубывающая функция $F(\lambda)$, такая, что

$$R(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} dF(\lambda), \quad F(-\infty) = 0. \quad (32)$$

в случае непрерывного времени и

$$R(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dF(\lambda), \quad F(-\pi) = 0 \quad (33)$$

в случае целочисленного времени. Интегралы понимаются как интегралы Лебега—Стилтьеса. Функцию $F(\lambda)$ называют спектральной функцией процесса, порожденную ею меру $F(\lambda, \mu) = F(\mu) - F(\lambda)$ называют спектральной мерой. Если спектральная мера имеет плотность, т. е. $dF(\lambda) = f(\lambda) d\lambda$, то функцию f называют спектральной плотностью процесса $X(t)$.

Пример 1. Пусть ξ и η — некоррелированные случайные величины с одинаковыми дисперсиями σ^2 и нулевыми средними. Тогда процесс

$$X(t) = \xi \cos t + \eta \sin t$$

является стационарным. Его ковариационная функция $R(t) = \sigma^2 \cos t$. Сумма независимых стационарных процессов также является стационарным процессом, поэтому случайный процесс

$$X(t) = \sum_{k=1}^n \xi_k \cos s_k t + \eta_k \sin s_k t,$$

где $\xi_i, \eta_i, i = 1, \dots, n$ — некоррелированные случайные величины с нулевыми средними и попарно равными дисперсиями $\sigma_k^2 = M\xi_k^2 = M\eta_k^2, k = 1, \dots, n, s_k, k = 1, \dots, n$ — некоторые попарно различные числа, также является стационарным. Его ковариационная функция $R(t) = \sum_{k=1}^n \sigma_k^2 \cos s_k t$,

а спектральная функция является ступенчатой со скачками величины $\frac{1}{2} \sigma_k^2$ в точках $\pm s_k$. Спектральной плотности у процесса $X(t)$ не существует.

Пример 2. Пусть U_1, U_2, \dots, U_n , — последовательность некоррелированных случайных величин с нулевыми средними и одинаковыми дисперсиями σ^2 . Пусть a_0, a_1, \dots, a_k — некоторые действительные числа. Обра-

зуем процесс скользящего среднего k -го порядка:

$$X(t) = a_0 U_t + a_1 U_{t-1} + \dots + a_k U_{t-k}, \quad t = 0, \pm 1,$$

Этот процесс является стационарным. Нетрудно подсчитать его спектральную плотность:

$$f(\lambda) = |a_0 + a_1 e^{i\lambda} + \dots + a_k e^{ik\lambda}|^2, \quad -\pi \leq \lambda \leq \pi.$$

Рассмотрим случайный процесс, удовлетворяющий другому уравнению:

$$b_0 X(t) + b_1 X(t-1) + \dots + b_k X(t-k) = U_t, \quad t = 0, \pm 1,$$

Если существует стационарный процесс $X(t)$, удовлетворяющий этому уравнению, то его называют процессом авторегрессии k -го порядка. Его спектральная плотность

$$f(\lambda) = |b_0 + b_1 e^{-i\lambda} + \dots + b_k e^{-ik\lambda}|^{-2}, \quad -\pi \leq \lambda \leq \pi.$$

Стационарный процесс авторегрессии существует тогда и только тогда, когда корни полинома $b_0 + \dots + b_k z^k$ лежат вне единичной окружности комплексной плоскости.

Процессы смешанного типа, у которых спектральная плотность имеет вид

$$f(\lambda) = \left| \frac{Q(e^{i\lambda})}{P(e^{i\lambda})} \right|^2 \quad -\pi \leq \lambda \leq \pi,$$

где Q и P — некоторые полиномы, называют процессами с рациональной спектральной плотностью или процессами авторегрессии — скользящего среднего.

Спектральную функцию процесса можно вычислить по следующим формулам. Если λ_1 и λ_2 — точки непрерывности функции $F(\lambda)$, то

$$F(\lambda_2) - F(\lambda_1) = \frac{1}{2\pi} \times \times \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it} R(t) dt \quad (34)$$

в случае непрерывного времени и

$$F(\lambda_2) - F(\lambda_1) = \frac{1}{2\pi} \times \times \sum_{-T}^T \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it} R(t) \quad (35)$$

в случае целочисленного t .

Спектральное разложение. В точках непрерывности λ_1, λ_2 спектральной функции $F(\lambda)$ имеет место сходимость в среднем квадратическом при $T \rightarrow \infty$ следующих выражений:

$$\frac{1}{2\pi} \int_{-T}^T \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it} X(t) dt \quad (36)$$

в случае непрерывного t и

$$\frac{1}{2\pi} \sum_{-T}^T \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it} X(t) \quad (37)$$

в случае целочисленного t . Этот предел удобно обозначить через $\zeta(\lambda_2) - \zeta(\lambda_1)$. Нетрудно проверить, что процесс $\zeta(\lambda)$ является процессом с ортогональными приращениями (см. параграф 5) и имеет место равенство $X(t) = \int e^{it\lambda} d\zeta(\lambda)$, где пределы интегрирования равны $-\infty, +\infty$ в случае непрерывного времени и $-\pi$ и π в случае целочисленного времени.

Если наблюдается процесс $X(t)$ на некотором достаточно большом отрезке $[-T, T]$, то, построив с помощью приведенных выше интегралов приближения для процесса $\zeta(\lambda)$, можно строить приближения (прогноз) для значений процесса $X(t)$ в точках вне отрезка $[-T, T]$.

Линейным преобразованием стационарного процесса $X(t)$ называют преобразование вида

$$Y(t) = \int e^{i\lambda t} \varphi(\lambda) d\zeta(\lambda), \quad (38)$$

где функция $|\varphi|^2$ интегрируема по $dF(\lambda)$ на всей прямой в случае непрерывного времени и на отрезке $[-\pi, \pi]$ — в случае целочисленного t . Функцию φ называют спектральной характеристикой данного преобразования. Случайный процесс $Y(t)$ также яв-

ляется стационарным со спектральной мерой

$$G(\Delta) = \int_{\Delta} |\varphi(\lambda)|^2 dF(\lambda). \quad (39)$$

В случае, когда процесс X имеет спектральную плотность f , процесс U также имеет плотность g , причем $g(\lambda) = |\varphi(\lambda)|^2 f(\lambda)$. Процесс $\eta(\lambda)$ с ортогональными приращениями, соответствующий процессу $Y(t)$, строится по следующей формуле:

$$\eta(\lambda_2) - \eta(\lambda_1) = \int_{\lambda_1}^{\lambda_2} \varphi(\lambda) d\zeta(\lambda). \quad (40)$$

Пример 3. Д и ф ф е р е н ц и р о в а н и е. Если спектральная мера F такова, что

$$\lambda_2 = \int_{-\infty}^{+\infty} \lambda^2 dF(\lambda) < \infty,$$

то функция $\varphi(\lambda) = i\lambda$ является спектральной характеристикой линейного преобразования $Y(t) = X'(t)$, где дифференцирование производится в среднем квадратическом смысле.

В ы н у ж д е н н ы е к о л е б а н и я м а я т н и к а. Линейное преобразование, задаваемое обыкновенным линейным дифференциальным уравнением

$$Y''(t) + 2hY'(t) + \omega_0^2 Y(t) = X(t),$$

имеет спектральную характеристику

$$\varphi(\lambda) = ((i\lambda)^2 + 2hi\lambda + \omega_0^2)^{-1}.$$

И н т е г р и р о в а н и е. Если выполнено условие

$$\lambda_{-2} = \int_{-\infty}^{+\infty} \lambda^{-2} dF(\lambda) < \infty,$$

то функция $\varphi(\lambda) = \frac{1}{i\lambda}$ является спектральной характеристикой преобразования

$$Y'(t) = X(t)$$

и дифференцирование понимается в среднем квадратическом.

Скользящее суммирование. Пусть $C(t)$ — интегрируемая функция, тогда функция

$$\varphi(\lambda) = \int_{-\infty}^{+\infty} e^{-i\lambda t} C(t) dt$$

является спектральной характеристикой линейного преобразования

$$Y(t) = \int_{-\infty}^{+\infty} C(t-s) X(s) ds.$$

Важным следствием спектрального разложения (36) является теорема Котельникова для процессов с ограниченным спектром. Если спектральная функция $F(\lambda)$ сосредоточена на отрезке $[-C, C]$, $C < \infty$, т. е. $F(-\infty) = F(-C)$ и $F(+\infty) = F(C+0)$, то

$$X(t) = \sum_{n=-\infty}^{+\infty} \frac{\sin(Ct - \pi n)}{Ct - \pi n} \times X\left(\frac{\pi n}{C}\right), \quad (41)$$

т. е. все значения процесса $X(t)$ восстанавливаются по его значениям в дискретные моменты времени.

Величину $\lambda_k = \int |\lambda|^k dF(\lambda)$ называют k -м абсолютным спектральным моментом стационарного процесса со спектральной функцией $F(\lambda)$. Если $R(t)$ — ковариационная функция этого процесса и $\lambda_{2k} < \infty$, то

$$\lambda_{2k} = (-1)^k \left. \frac{d^{2k} R(t)}{dt^{2k}} \right|_{t=0} \quad (42)$$

7. ГАУССОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

Случайный процесс $X(t)$, все конечномерные распределения которого, т. е. распределения векторов $(X(t_1), \dots, X(t_n))$, являются гауссовскими многомерными распределениями, называют гауссовским (или нормальным) случайным процессом.

Поскольку гауссовское многомерное распределение полностью характеризуется вектором средних значений и

ковариационной матрицей, то распределение произвольного функционала, порожденного гауссовским процессом, может быть выражено в терминах двух функций: функции математического ожидания $m(t) = MX(t)$ и функции ковариации $R(t, s) = M(X(t) - m(t))(X(s) - m(s))$.

Пример 1. Последовательность $X(t)$, $t = 0, \pm 1$, независимых нормальных случайных величин с нулевыми средними и дисперсиями σ^2 является гауссовским случайным процессом. Для этого процесса $m(t) \equiv 0$ и $R(t, s) = \sigma^2 \delta_{t,s}$, где $\delta_{t,s}$ — символ Кронеккера. Этот процесс называют белым гауссовским шумом с дискретным временем.

Последовательность

$$Y(t) = \sum_{k=0}^{\infty} a_k X(t-k),$$

$$t = 0, \pm 1, \dots,$$

где a_k — неслучайные числа, такие,

что $\sum_0^{\infty} a_k^2 < \infty$ также является гауссовским случайным процессом.

Для этого процесса

$$m(t) \equiv 0, \quad R(t, s) =$$

$$= \sum_{k=0}^{\infty} a_k a_{k+|t-s|}$$

Случайный процесс, полученный с помощью произвольного линейного преобразования гауссовского процесса, также является гауссовским случайным процессом.

Этот факт является простым следствием того, что произвольное линейное преобразование гауссовского случайного вектора также имеет гауссовское распределение.

Гауссовский стационарный процесс в широком смысле является стационарным в узком смысле.

Это следствие того, что произвольное конечномерное распределение гауссовского процесса является гауссовским.

Случайный процесс примера 1, параграф 6, где ξ и η — гауссовские независимые случайные величины с нулевыми средними и дисперсиями σ^2 ,

является простейшим примером гауссовского случайного процесса с непрерывным временем. Этот процесс называют косинус-процессом.

Стандартный винеровский процесс $W(t)$ (пример 2, параграф 5) является примером гауссовского нестационарного процесса. Для него $m(t) = 0$, $R(t, s) = \min(t, s)$, $t \geq 0, s \geq 0$.

Пусть имеется произвольный случайный процесс. Тогда существует гауссовский случайный процесс с теми же математическим ожиданием и ковариационной функцией. Иными словами, для любой функции $m(t)$ и функции $R(t, s)$, являющейся ковариационной, существует гауссовский случайный процесс с функцией математического ожидания $m(t)$ и функцией ковариаций $R(t, s)$.

Пример 2. Гауссовский стационарный случайный процесс с нулевым средним и ковариационной функцией $R(t) = \sigma^2 e^{-\alpha |t|}$ называют **процессом Орнштейна—Уленбека**. Этот процесс обычно применяют при описании шероховатых поверхностей. Его траектории с вероятностью единица недифференцируемы. Другим процессом, применяемым для описания шероховатости, является гауссовский процесс $X(t)$ с нулевым средним и с «треугольной» ковариационной функцией $R(t) = \sigma^2 \max(1 - a|t|, 0)$, $a > 0$. Имеет место соотношение

$$X(t) = \sigma (W_{(t+1)\sqrt{a}} - W_t \sqrt{a}),$$

где W_t — винеровский стандартный процесс.

Поскольку ковариационная функция стационарного случайного процесса однозначно связана с его спектральной функцией, то распределения гауссовского стационарного процесса можно полностью охарактеризовать его математическим ожиданием m и его спектральной функцией (или спектральной плотностью, если она существует).

Спектральной плотностью процесса Орнштейна—Уленбека является функция $f(\lambda) = \frac{\alpha \sigma^2}{\alpha^2 + \lambda^2}$. Спектральной плотностью белого шума с дискретным временем является константа: $f(\lambda) = \frac{\sigma^2}{2\pi}$, $-\pi \leq \lambda \leq \pi$. Используя выражения для спектральной плотности

линейного преобразования, можно получить спектры других приведенных выше гауссовских стационарных процессов.

Вопрос о локальных свойствах регулярности (непрерывность, дифференцируемость и т. п.) траекторий гауссовского процесса может быть эффективно решен в терминах его характеристик $m(t)$, $R(t, s)$.

Пусть имеется основное вероятностное пространство $(\Omega, \mathfrak{B}, P)$ и гауссовский случайный процесс $X(t) = X(t, \omega)$, $\omega \in \Omega$.

Закон 0 или 1. Реализации гауссовского случайного процесса либо с вероятностью единица непрерывны [почти наверное непрерывны (п. н.)], либо с вероятностью единица разрывны. Другими словами, вероятность события $\{\omega \mid X(t, \omega) \text{ непрерывна по } t\}$ равна 0 или 1.

Если гауссовский процесс является стационарным, то имеет место следующая альтернатива **Беляева**: либо его траектории с вероятностью единица непрерывны, либо они с вероятностью единица неограниченны на любом сколь угодно малом интервале (т. е. любая точка является точкой разрыва второго рода).

Для построения эффективных критериев непрерывности траекторий удобно ввести функцию стандартного отклонения: $d^2(t, s) = \text{Var}(X(t) - X(s))$.

Достаточное условие п. н. непрерывности. Если у гауссовского случайного процесса $X(t)$, заданного на некотором интервале $[0, T]$, математическое ожидание $m(t)$ является непрерывной функцией, а $d(t, s) \leq C |\ln |t - s||^{1/2 + \varepsilon}$ для некоторых $C < \infty$ и $\varepsilon > 0$, то траектории процесса $X(t)$ п. н. непрерывны. Если же для некоторого $\delta > 0$ $d(t, s) \geq \delta |\ln |t - s||^{1/2}$, то его траектории п. н. разрывны.

В случае, когда параметр t является векторным, т. е. имеется **гауссовское случайное поле**, все приведенные выше утверждения о локальных свойствах также имеют место с учетом того, что

$$|t| = \sqrt{\sum t_i^2} \text{ — евклидова норма.}$$

Гауссовский случайный процесс $X(t)$ k раз дифференцируем в среднем

квадратическом, если $m(t)$ k раз дифференцируема и функция $\frac{\partial^{2k}}{\partial t^k \partial s^k} R(t, s)$ непрерывна. Если его k -я средняя квадратическая производная удовлетворяет приведенному выше достаточному условию п. н. непрерывности, то траектории этого процесса k раз непрерывно дифференцируемы с вероятностью единица.

Неравенство Ферника. Пусть T — некоторое подмножество евклидова пространства R^n и $X(t)$ — гауссовское случайное поле, заданное на T . Введем функцию

$$\varphi(h) = \sup_{t \in T, s \in T, |t-s| \leq h} d(t, s),$$

$$0 \leq h \leq 1.$$

Если интеграл $\int_1^\infty \varphi(e^{-x^2}) dx$ конечен, то траектории поля $X(t)$ (или процесса в случае $n = 1$) непрерывны с вероятностью единица. Кроме того, для любого целого $p \geq 2$ и любого $x \geq \sqrt{1 + 4n \log_2 p}$ справедлива оценка

$$P \left(\sup_{t \in T} X(t) \geq x (\sup_{t \in T} \sqrt{\text{Var } X(t)} + (2 + \sqrt{2}) \int_1^\infty \varphi(p^{-u^2}) du) \right) \leq \leq \frac{5}{2} p^{2n} \int_x^\infty e^{-\frac{u^2}{2}} du. \tag{43}$$

Точная асимптотика для распределения максимума. Если ковариационная функция $R(t)$ имеет в точке ноль «угол»

$$R(t) = \sigma^2 (1 - \lambda |t| + o(|t|)), \quad t \rightarrow 0, \tag{44}$$

(как в примерах из параграфа 7) или является гладкой

$$R(t) = \sigma^2 \left(1 - \frac{1}{2} \lambda_2 t^2 + o(t^2) \right), \quad t \rightarrow 0, \tag{45}$$

как в примере 1, параграф 6), то мож-

но выписать точную асимптотику для распределения максимума или максимума модуля. Приведем соответствующую формулу для гауссовского однородного поля $X(t)$ с комбинированным поведением в нуле ковариационной функции. Пусть $t = (t_1, \dots, t_n)$, $MX(t) = m$, $R(t)$ — его ковариационная функция. Пусть множество A имеет ненулевой n -мерный объем, причем для любых различных точек t и s из A $R(t-s) < \sigma^2$, а при $|t| \rightarrow 0$

$$R(t) = \sigma^2 \left(1 - \left(\frac{1}{2} \sum_{i=1}^k \lambda_{2i} |t_i|^2 - \sum_{i=k+1}^n \lambda_i |t_i| \right) \right) (1 + o(1)),$$

где все числа λ_{2j} , λ_i , σ положительны. Тогда при $u \rightarrow \infty$

$$P \left(\max_{t \in A} \frac{X(t) - m}{\sigma} > u \right) = \frac{\prod_{i=1}^k \sqrt{\lambda_{2i}} \prod_{i=k+1}^n \lambda_i}{\sqrt{2} \pi^{\frac{k+1}{2}}} u^{2n-k-1} \times \times e^{-\frac{u^2}{2}} (1 + o(1)). \tag{46}$$

Для вероятности уклонения максимума модуля $\left| \frac{X(t) - m}{\sigma} \right|$ правая часть умножается на 2. При $n = 1$, $k = 0$ получаем формулу для процессов с ковариационной функцией вида (44); при $n = k = 1$ — для процессов с ковариационной функцией вида (45).

Предельное распределение максимума и максимума модуля. Если ковариационная функция $R(t)$ гауссовского стационарного процесса $X(t)$, $-\infty < t < +\infty$ достаточно быстро убывает на бесконечности: для некоторого $a > 0$ $\int_{-\infty}^{+\infty} |R(t)|^a dt < \infty$ (для непрерывного времени) или $\sum |R(t)|^a < \infty$ (для целочисленного времени), то имеет место предельное

соотношение, равномерное по всем x ,

$$\lim_{T \rightarrow \infty} P \left(l_T \left(\max_{0 \leq t \leq \alpha T} \frac{X(t) - m}{\sigma} - l_T \right) < x \right) = \exp \times \times \left(-e^{-\frac{x^2}{4l_T^2}} \right) = 0, \quad (47)$$

где $\alpha = 1$ в случае целочисленного времени; $\alpha = 1/\lambda$ в случае «угла» (44); $\alpha = 1/\sqrt{\lambda_2}$ в гладком случае (45); l_T — решение уравнения:

$$\frac{T}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx = 1, \quad l_T = \sqrt{2 \ln T} - \frac{\ln \sqrt{4\pi \ln T}}{\sqrt{2 \ln T}} (1 + o(1)), \quad T \rightarrow \infty$$

в случае целочисленного времени,

$$\frac{Tl}{\sqrt{2\pi}} e^{-\frac{l^2}{2}} = 1, \quad l_T = \sqrt{2 \ln T} + \frac{1}{2} \frac{\ln \ln T - \ln \pi}{\sqrt{2 \ln T}} (1 + o(1)), \quad T \rightarrow \infty$$

в случае «угла»,

$$l_T = \sqrt{2 \ln(T/2\pi)}$$

в гладком случае.

Слагаемое $-\frac{x^2}{4l_T^2}$ под двойной экспонентой в соотношении (47) улучшает приближение в случаях целочисленного времени и гладкой функции $R(t)$.

Пересечения уровня. Пусть $X(t)$ — гауссовский процесс, реализации которого — дифференцируемые функции. Тогда точкой **выхода** процесса $X(t)$ **за уровень** u называют точку τ , такую, что $X(\tau) = u$, $X'(\tau) > 0$. Точкой **входа под уровень** u называют точку: $X(\tau) = u$, $X'(\tau) < 0$. Для гауссовского процесса с необращающейся в нуль дисперсией нет точек касания $X(\tau) = u$, $X'(\tau) = 0$ с вероятностью единица. Входы и выходы называют пересечениями уровня. Приведем формулы для моментов числа входов $C(u, T)$ и пересечений $N(u, T)$ уровня u

гауссовским стационарным процессом с нулевым средним и ковариационной функцией, удовлетворяющей условию (45):

$$MN(u, T) = \frac{T}{\pi} \frac{\sqrt{\lambda_2}}{\sigma} e^{-u^2/2\sigma^2}, \quad (48)$$

$$MC(u, T) = \frac{1}{2} MN(u, T). \quad (49)$$

Пусть $M_k(T) = MC(u, T) (C(u, T) - 1) (C(u, T) - k + 1) = k$ -й факториальный момент числа $C(u, T)$, тогда

$$M_k(T) = \int_0^T \int_0^T dt_1 \dots dt_k \int_0^{\infty} y_1 \dots y_k \rho_t(u, u, y_1, y_k) dy_1 \dots dy_k, \quad (50)$$

где $\rho_t(x_1, \dots, x_k, y_1, \dots, y_k)$ — совместная плотность вероятностей случайных величин $X(t_1), \dots, X(t_k), X'(t_1), \dots, X'(t_k)$.

СПИСОК ЛИТЕРАТУРЫ

1. Гихман И. И., Скороход А. В. Введение в теорию случайных процессов. 2-е изд. М.: Наука, 1977. 568 с.
2. Ибрагимов И. А. Розанов Ю. А. Гауссовские случайные процессы. М.: Наука, 1970. 384 с.
3. Кемени Д. Д., Снелл Д. Л. Конечные цепи Маркова/Пер. с англ. М.: Наука, 1970. 271 с.
4. Кокс Д. Р., Смит В. Л. Теория восстановления/Пер. с англ. М.: Сов. радио, 1967. 299 с.
5. Крамер Г., Лидбеттер М. Стационарные случайные процессы/Пер. с англ. М.: Мир, 1969. 398 с.
6. Очереди и точечные процессы/П. Франкен, Д. Кёниг, У. Арндт и др. Пер. с англ. Киев: Наукова думка, 1984. 284 с.
7. Феллер В. Введение в теорию вероятностей и ее приложения/Пер. с англ., т. 1, М.: Мир, 1984. 528 с.
8. Чжун Кай-лай. Однородные цепи Маркова/Пер. с англ. М.: Наука, 1964. 425 с.

В математической статистике разрабатываются методы планирования получения данных, характеризуемых случайной изменчивостью, и методы анализа и интерпретации таких данных. Основными объектами исследования в математической статистике являются статистические модели.

1. ФУНКЦИЯ ПРАВДОПОДОБИЯ

Совокупность данных, подлежащих статистическому анализу, всегда определяется системой правил сбора данных или системой правил проведения эксперимента. Эту систему правил называют **планом получения данных**. Данные x рассматривают как результат случайных явлений, т. е. предполагают, что в \mathfrak{X} задан набор (σ -алгебра) подмножеств \mathfrak{B}_Π и семейство вероятностных распределений (мер) $\mathcal{P}_\Pi = \{P\}$. Множество \mathfrak{X} всех возможных случайных данных x при заданном плане Π получения данных называют **генеральной совокупностью**. То вероятностное распределение, в соответствии с которым получены данные x , называют **порождающим**. Порождающее распределение неизвестно. Обычно либо предполагается, что порождающее распределение входит в семейство \mathcal{P}_Π , либо ставится вопрос о принадлежности порождающего распределения семейству \mathcal{P}_Π .

Если все распределения семейства \mathcal{P}_Π можно однозначно определить набором конечного числа параметров, то семейство \mathcal{P}_Π называют **параметрическим**. В противном случае семейство распределений \mathcal{P}_Π называют **непараметрическим**.

Семейство вероятностных распределений \mathcal{P}_Π называют **доминированным** мерой μ , если для любой $P \in \mathcal{P}_\Pi$ и

любого $A \in \mathfrak{B}_\Pi$, такого, что $\mu(A) = 0$, следует, что $P(A) = 0$. Если семейство \mathcal{P}_Π доминируется σ -конечной мерой μ , то существуют плотности распределений $p(x) = \frac{dP}{d\mu}(x)$.

Если семейство распределений \mathcal{P}_Π параметрическое, то существует $\Theta = \{\theta\}$ — множество значений параметров θ , однозначно определяющих распределения $P_\theta \in \mathcal{P}_\Pi$. Если семейство $\mathcal{P}_\Pi = \{P_\theta : \theta \in \Theta\}$ доминируется мерой μ , то его плотности распределений

$p(x, \theta) = \frac{dP_\theta}{d\mu}(x)$ также однозначно определяются значением параметра θ для всех x .

Пусть $\mathcal{P}_\Pi = \{P_\theta : \theta \in \Theta\}$ — параметрическое доминированное семейство. При любом фиксированном x функцию параметра θ

$$L(\theta | x) = p(x, \theta)$$

называют **функцией правдоподобия**, а ее значение для любого фиксированного θ^* называют **правдоподобием θ^*** .

Если θ и x фиксированы, то функция правдоподобия численно равна плотности распределения с теми же значениями параметра θ и данных x . Во многих задачах часто удобнее рассматривать логарифм (натуральный) функции правдоподобия

$$l(\theta, x) = \ln L(\theta | x).$$

Пример 1. Выборочному контролю подвергают партию \mathfrak{B} , состоящую из N изделий, среди которых D изделий дефектные, $K = N - D$ — годные. Из партии берут случайную выборку (см. гл. 14) из n изделий, т. е. план контроля Π состоит в последовательном отборе без возвращения на контроль n изделий. Вероятность отбора одинако-

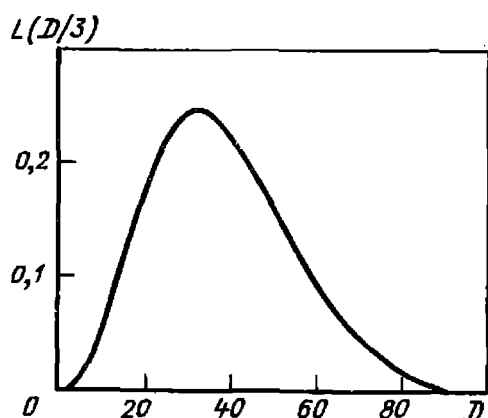


Рис. 1. Функция правдоподобия для семейства гипергеометрических распределений:

$$N = 300; n = 30; d = 3$$

ва для всех еще неотобранных изделий. Результатом контроля являются данные $x = d$ — число обнаруженных дефектных изделий. Пространство данных $\mathfrak{X}_\Pi = \{0, 1, \dots, n\}$. Вероятность получения данных $x = d$ будет

$$p(x, d) = h_{N, D}^{n, d} = \binom{D}{d} \binom{K}{k} / \binom{N}{n},$$

$$k + d = n.$$

Вероятности $p(x, D)$ определяют семейство вероятностных распределений (семейство гипергеометрических распределений). В качестве параметра θ взято общее число дефектных изделий D , т. е. $\theta = D$, $\Theta = \{0, 1, \dots, N\}$. Меняя значения D при фиксированном значении d , получим функцию правдоподобия $L(D | d) = p(d, D)$. На рис. 1 показан график функции правдоподобия для $N = 300$, $n = 30$, $D = 3$.

Пример 2. План Π получения данных при проведении испытаний на надежность состоит в постановке на испытания N изделий, регистрации всех отказов и окончания испытаний в момент наступления r -го отказа. Данные x состоят из последовательности моментов r отказов $x = (t_1, \dots, t_r)$, $0 \leq t_1 \leq \dots \leq t_r$. Пусть продолжительности безотказной работы испытуемых элементов — взаимно независимые одинаково распределенные (н. о. р.) случайные величины с ф. р. $F(t) = 1 - \exp(-\lambda t)$. Семейство вероятностных распределений \mathcal{P}_Π на $\mathfrak{X}_\Pi = \{x = (t_1, \dots, t_r): 0 \leq t_1 \leq \dots \leq t_r\} \subset R^k$ является параметрическим, $\theta =$

$= \lambda \in R_+^1 = [0, \infty = \Theta$. Плотности распределений из \mathcal{P}_Π имеют вид

$$p(x, \lambda) = N^{(r)} \lambda^r \exp\{-\lambda S(t_{(r)})\},$$

где $S(t_{(r)}) = N t_{(1)} + (N-1)(t_{(2)} - t_{(1)}) + \dots + (N-r+1)(t_{(r)} - t_{(r-1)})$ — суммарная наработка, $N^{(r)} = N(N-1)\dots(N-r+1)$.

Логарифм функции правдоподобия

$$l(\lambda, x) = \ln N^{(r)} + r \ln \lambda - \lambda S(t_{(r)}).$$

Семейство распределений \mathcal{P}_Π на некотором пространстве данных \mathfrak{X}_Π называют семейством экспоненциального типа, если оно доминируется мерой μ , является параметрическим с множеством значений параметров $\Theta = \{(\theta_1, \dots, \theta_m)\} \in R^m$, а его плотности относительно μ имеют вид

$$p(x, \theta) = \exp\left\{\sum_{j=1}^n v_j(\theta) T_j(x) + b(x) + c(\theta)\right\}.$$

Здесь $\{T_1(x), \dots, T_k(x)\}$ — линейно независимые вещественные функции данных x ; $b(x), c(\theta), v_1(\theta), \dots, v_k(\theta)$ — вещественные (измеримые) функции своих аргументов.

Если $v_j(\theta) = \theta_j$, $j = 1, \dots, k$, $k = m$, то Θ называют естественным параметрическим пространством, а семейство \mathcal{P}_Π называют k -параметрическим семейством экспоненциального типа.

Замечание. Семейство распределений в примере 1 не является семейством экспоненциального типа. Семейство распределений в примере 2 является семейством экспоненциального типа.

2. СТАТИСТИКИ

Статистиками называют такие (измеримые) функции $T(x)$ данных $x \in \mathfrak{X}_\Pi$, принимающие значения в некотором множестве \mathcal{U} , значения которых не зависят от неизвестных параметров или других неизвестных характеристик распределения, порождающего данные.

Пример 1. Пусть x_1, \dots, x_n — последовательность значений n вещественных н. о. р. случайных величин. Ста-

статистиками являются: среднее $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, эмпирическая диспер-

сия $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$, медиана

$x_{0.5} = x_{(m)}$, если $n = 2m + 1$, $x_{0.5} = \frac{1}{2} [x_{(m)} + x_{(m+1)}]$, если $n = 2m$, где $x_{(i)}$ — i -е по величине значение в последовательности x_1, \dots, x_n , размах $R = \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i = x_{(n)} - x_{(1)}$.

Упорядочив несовпадающие значения x_1, \dots, x_n по величине, получим последовательность $x_{(1)} < x_{(2)} < \dots < x_{(n)}$. Пусть $x_i = x_{(r_i)}$, т. е. i -й элемент в исходной последовательности является r_i -м по своей величине. r_i называют рангом i -го значения. Последовательность (r_1, \dots, r_n) является (ранговой) статистикой.

Статистику T называют достаточной для семейства распределений \mathcal{P}_Π , если условные распределения $P(A | T(x) = t)$ одинаковы для всех распределений $P \in \mathcal{P}_\Pi$.

Критерий факторизации. Пусть \mathcal{P}_Π — параметрическое, доминированное σ -конечной мерой μ , семейство распределений $\mathcal{P}_\Pi = \{P_\theta\}$ с плотностями $p(x, \theta)$. Статистика $T(x) \in \mathcal{U}$ достаточна для семейства \mathcal{P}_Π тогда, когда существуют неотрицательные (измеримые) функции h на \mathcal{X} и g_θ на \mathcal{U} , такие, что для любого θ почти наверное

$$p(x, \theta) = h(x) g_\theta(T(x)).$$

Множество значений в пространстве данных $A_t^T = \{x : x \in \mathcal{X}_\Pi, T(x) = t\}$, на котором статистика T принимает постоянное значение, называют атомом статистики T .

Атомы статистики T образуют разбиение множества данных на непересекающиеся подмножества.

Пусть $T^{-1}(\mathcal{B}_\Pi) = \sigma\{T^{-1}(B) : B \in \mathcal{B}_T\}$, где \mathcal{B}_T есть σ -алгебра множеств в пространстве \mathcal{U} значений статистики T , $T^{-1}(\mathcal{B}_\Pi) \subseteq \mathcal{B}_\Pi$, $T^{-1}(B) = \{x : x \in \mathcal{X}_\Pi, T(x) \in B\}$.

Две статистики называют эквивалентными, если определяемые их разби-

ениями пространства данных минимальные σ -алгебры совпадают.

Достаточную статистику T_1 называют минимальной, если для любой другой достаточной статистики T_2 : $T_2^{-1}(\mathcal{B}_\Pi) \subseteq T_1^{-1}(\mathcal{B}_\Pi)$.

Образно говоря, атомы статистики T_1 крупнее или совпадают с атомами любой другой достаточной статистики T_2 .

Если параметрическое семейство распределений \mathcal{P}_Π доминируется σ -конечной мерой μ , то найдется такая последовательность распределений $P_{\theta_i} \in \mathcal{P}_\Pi$, что \mathcal{P}_Π доминируется мерой

$$\lambda(A) = \sum_{i=1}^{\infty} c_i P_{\theta_i}(A), \quad c_i \geq 0, \quad \sum_{i=1}^{\infty} c_i = 1.$$

Пусть $\tilde{p}(x, \theta)$ — плотность относительно λ , тогда

$$\tilde{p}(x, \theta) = \frac{p(x, \theta)}{\sum_{i=1}^{\infty} c_i p(x, \theta_i)}.$$

Рассмотрим функцию правдоподобия

$$\tilde{L}(\theta | x) = \tilde{p}(x, \theta).$$

Функция правдоподобия \tilde{L} является минимальной достаточной статистикой.

Семейство распределений \mathcal{P}_Π называют полным (ограниченно полным), когда для любой (ограниченной) вещественной функции $\varphi(x)$, $x \in \mathcal{X}_\Pi$, имеющей для любой $P \in \mathcal{P}_\Pi$ конечное математическое ожидание, из тождества

$$M_P(\varphi) = \int_{x \in \mathcal{X}} \varphi(x) P(dx) = 0, \quad \text{верного}$$

для любой $P \in \mathcal{P}_\Pi$, следует, что $\varphi(x) = 0$ п. н. для любой $P \in \mathcal{P}_\Pi$.

Статистику T называют полной (ограниченно полной), если соответствующее ей семейство распределений \mathcal{P}^T является полным (ограниченно полным), где вероятность $P^T(A) = P(T(x) \in A)$ определяет распределение семейства \mathcal{P}^T в множестве \mathcal{U} значений статистики T .

Ограниченно полная достаточная статистика является минимальной достаточной статистикой.

3. ОЦЕНКИ

Если значение некоторой статистики рассматривают как приближение к определенной характеристике порождающего распределения, то такую статистику называют (точечной) **оценкой** этой характеристики.

Для параметрической модели $(\mathfrak{X}, \mathcal{P})$, $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ типичной является задача оценивания заданной функции от параметра $g(\theta)$.

Оценку $T(x)$ называют **несмещенной оценкой** для $g(\theta)$, если для всех значений $\theta \in \Theta$

$$M_\theta T(x) = g(\theta).$$

Замечание. Если функция $g(\theta)$ является векторной, т. е. $g(\theta) = (g_1(\theta), \dots, g_h(\theta))$, то в качестве оценок рассматривают векторные функции $T(x) = (T_1(x), \dots, T_h(x))$.

Пример 1. Пусть выполнены условия примера 1 параграфа 2 и $x_i \sim N(\mu, \sigma^2)$, тогда статистика $T(x_1, \dots, x_n) = (\bar{x}, s^2)$ является несмещенной оценкой векторного параметра $\theta = (\mu, \sigma^2)$.

Пример 2. Пусть $\mathfrak{B} = \{O_{11}, \dots, O_N\}$ — конечная совокупность N объектов (изделий), каждый из которых характеризуется скалярной величиной. Объекту O_i соответствует величина Y_i , $i = 1, \dots, N$. Из \mathfrak{B} берется случайная выборка, т. е. задается случайная последовательность чисел $\gamma_1, \dots, \gamma_N$, в которой $\gamma_i = 1$ соответствует включению объекта O_i в выборку, а $\gamma_i = 0$ — не включению объекта O_i в выборку. Пусть $\pi_i = P\{\gamma_i = 1\}$ — вероятность включения в выборку, а $s = \{i : \gamma_i = 1\}$ — номера объектов, составляющих выборку. Статистика

$$e(s) = \sum_{i \in s} \frac{Y_i}{\pi_i}$$

является несмещенной оценкой суммы $Y_1 + \dots + Y_N$. Оценки такого вида называют **оценками Горвица—Томпсона**.

Пример 3. Пусть $y = (y_1, \dots, y_n)$, x_i — н. о. р. случайные величины с ф. р. $F(x) = P\{y_i \leq x\}$. Положим $I(y_i \leq x) = 1$, если $y_i \leq x$, и $I(y_i \leq$

$\leq x) = 0$ в противном случае. Величины $I(y_i \leq x)$, $i = 1, \dots, n$ являются статистиками. Статистика

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(y_i \leq x)$$

является несмещенной оценкой значения ф. р. $F(x)$ в точке x . Если $y_{(1)} \leq \dots \leq y_{(n)}$ — порядковая статистика, то статистику

$$\hat{z}_\gamma = \begin{cases} y_{[\gamma n]+1}, & \text{если } \gamma n \text{ — не целое} \\ & \text{число,} \\ y_{\gamma n}, & \text{если } \gamma n \text{ — целое число} \end{cases}$$

можно рассматривать как оценку γ -квантили ф. р. $F(x)$. Статистика $\hat{m}_k =$

$$= \frac{1}{n} \sum_{i=1}^n y_i^k$$
 является несмещенной

оценкой k -го момента $m_k = \int_{-\infty}^{+\infty} x^k dF(x)$.

Замечание. Иногда вместо терминов «оценка квантили», «оценка момента» и т. д. используются термины «эмпирическая квантиль», «эмпирический момент».

Для получения оценок $g(\theta_0)$, где θ_0 — истинное значение параметра, определяющего порождающее распределение параметрической модели $(\mathfrak{X}, \mathcal{P})$, $\mathcal{P} = \{P_\theta, \theta \in \Theta \subseteq R^m, m < \infty\}$, используются общие методы.

Пусть $L(\theta | x)$ — функция правдоподобия параметрической модели, $C_g(u) = \{\theta : g(\theta) = u\}$ — атом функции $g(\theta)$, соответствующий уровню u . Статистику $\hat{g}(x)$ называют **оценкой максимального правдоподобия** (о. м. п.), если

$$\max_{\theta \in C_g(\hat{g}(x))} L(\theta | x) = \max_{\theta \in \Theta} L(\theta | x).$$

Если функция правдоподобия имеет непрерывные частные производные по θ_i , то о. м. п. $\theta = (\theta_1, \dots, \theta_m)$ находятся среди решений системы уравнений

$$\frac{\partial l(\theta | x)}{\partial \theta_i} = 0, \quad (1)$$

где $l(\theta | x) = \ln L(\theta | x)$.

Система уравнений (1) может быть решена методом последовательных приближений Ньютона — Рафсона. Если $\theta^{(0)}$ — исходное приближение, то

$$\theta_i^{(r+1)} = \theta_i^{(r)} - \sum_{j=1}^m I^{ij}(\theta^{(r)}) \times \frac{\partial l(\theta^{(r)} | x)}{\partial \theta_j},$$

где $I^{ij}(\theta^{(r)})$ — элемент матрицы $I^{-1}(\theta^{(r)})$ (см. определение информационной матрицы Фишера ниже в параграфе 4).

Замечание. О. м. п. могут не существовать. Для некоторых статистических моделей решение уравнений (1) может быть неоднозначным.

Пример 4. Пусть $t_1 < \dots < t_{(r)}$ — моменты первых r отказов при испытаниях по плану $[N, B, r]$ элементов с экспоненциальной ф. р. $F(t) = 1 - e^{-\lambda t}$, $\theta \doteq \lambda$. Функция правдоподобия $L(\lambda | x) = N^{(r)} \lambda^{(r)} e^{-\lambda S_r}$, где $x = (t_{(1)}, \dots, t_{(r)})$, $S_r = t_{(1)} + \dots + t_{(r-1)} + (N - r + 1) t_{(r)}$ — суммарная наработка. В соответствии с (1) о. м. п. параметра

$$\hat{\lambda} = r/S_r.$$

Эта оценка является смещенной. Смещение устраняется, если рассмотреть близкую оценку

$$\frac{r-1}{r} \hat{\lambda} = \frac{r-1}{S_r}.$$

Если для математического ожидания статистики $T(x)$ выполнено тождество $M_{\theta} T(x) \equiv h(\theta)$,

а разброс значений $T(x)$ относительно $h(\theta)$ невелик, то $T(x)$ можно использовать для получения оценки $\hat{\theta}$ для неизвестного параметра θ_0 . Оценку $\hat{\theta}$ находят как решение уравнения

$$T(x) = h(\hat{\theta}). \tag{2}$$

Замечание. В случае векторного параметра $\theta = (\theta_1, \dots, \theta_m)$ используют векторную статистику $T(x) = (T_1(x), \dots, T_m(x))$ и векторную функцию $h(\theta) = (h_1(\theta), \dots, h_m(\theta))$.

Иногда метод получения оценок на основе уравнения (2) называют **методом моментов**.

Пример 5. Пусть $x = (y_1, \dots, y_n)$, y_i — значения н. о. р. случайные величины с ф. р. $y_i \sim \Gamma(\alpha, \beta)$, п. р. соответствующего Γ -распределения $f(x, \theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$, $\theta = (\beta, \alpha)$, $x \geq 0$.

Пусть $T(x) = \left(\frac{1}{n} \sum_{i=1}^n y_i, \frac{1}{n} \times \sum_{i=1}^n y_i^2 \right)$, $h(\theta) = \left(\frac{\alpha}{\beta}, \frac{\alpha(\alpha-1)}{\beta^2} \right)$.

Уравнения (2) имеют вид $\hat{m}_1 = \hat{\alpha}/\hat{\beta}$, $\hat{m}_2 = \hat{\alpha}(\hat{\alpha}-1)/\hat{\beta}^2$,

где $m_1 = \frac{1}{n} \sum_{i=1}^n y_i$, $\hat{m}_2 = \frac{1}{n} \sum_{i=1}^n y_i^2$.

Решения этих уравнений

$$\hat{\beta} = \frac{\hat{m}_1}{\hat{m}_2 - (\hat{m}_1)^2}; \quad \hat{\alpha} = \frac{(\hat{m}_1)^2}{\hat{m}_2 - (\hat{m}_1)^2}.$$

Пусть система уравнений $F(x_{\gamma_i}, \tilde{\theta}) = \gamma_i$, $i = 1, \dots, m$ однозначно определяет возможные значения параметра $\tilde{\theta}(\tilde{\theta}_1, \dots, \tilde{\theta}_m)$, а соответствие между $\gamma = (\gamma_1, \dots, \gamma_m)$ и $\tilde{\theta}$ является взаимно однозначным и непрерывным. Если \hat{x}_i — оценки γ_i -квантили ф. р. $F(x, \theta_0)$, то оценку $\hat{\theta}$ параметра θ_0 можно получить как решение системы уравнений

$$F(\hat{x}_i, \hat{\theta}) = \gamma_i, \quad i = 1, \dots, m. \tag{3}$$

Метод получения оценки $\hat{\theta}$ на основе решения системы уравнений (3) называют **методом квантилей**.

Пусть испытания N изделий с ф. р. $F(t, \theta) = F_0\left(\frac{g(t) - \delta}{\sigma}\right)$, $\theta = (\delta, \sigma)$ проводятся до r -го отказа, т. е. по плану $[N, B, r]$.

Данные испытаний $x = (t_1, \dots, t_{(r)})$, $t_{(i)}$ — момент i -го отказа. Распределения случайных величин (не статистик!) $g(h) = \frac{g(t_{(h)}) - \delta}{\sigma}$ не зависят от δ и σ .

Пусть $a_k = Mg(k)$ и $b_{k,l} = Mg_{(k)}g_{(l)}$. Если искомая функция $g(\theta) = l_1\delta + l_2\sigma$, где l_1, l_2 — известные константы, то для ее оценивания можно использовать несмещенную оценку

$$T(x) = \sum_{k=1}^r c_k g(t_{(k)}),$$

где

$$\sum_{k=1}^r c_k = l_1, \quad \sum_{k=1}^r a_k c_k = l_2.$$

Дисперсия этой оценки равна

$$\sum_{k,l=1}^r b_{kl} c_k c_l. \text{ Для получения оценки с наименьшей дисперсией нужно найти набор } c_1, \dots, c_r, \text{ для которого минимально значение } \sum_{k,l=1}^r b_{kl} c_k c_l \text{ при ограничении } \sum_{k=1}^r c_k = l_1.$$

4. ПОКАЗАТЕЛИ КАЧЕСТВА ОЦЕНОК

Качество оценок \hat{g} может быть определено на основе задания (выбора) функции риска $w(g', g'') \geq 0$, $w(g', g') = 0$. Число $w(g', g'')$ характеризует «потери» в случае, когда истинное значение оцениваемой величины равно g'' , а в качестве оценки взято значение g' . Тогда показателем качества оценки $\hat{g} = \hat{g}(x)$ является средний риск

$$R(\hat{g}, \theta) = M_{\theta} w(\hat{g}, \hat{g}(\theta)).$$

Если $g(\theta)$ и \hat{g} — скалярные величины, а $w(g', g'') = (g' - g'')^2$, то $R(\hat{g}, \theta) = M_{\theta} (\hat{g}(x) - g(\theta))^2$ — среднее квадратическое отклонение. Если к тому же \hat{g} — несмещенная оценка для $g(\theta)$, то среднее квадратическое отклонение совпадает с дисперсией оценки \hat{g} .

Пусть $\theta = (\theta_1, \dots, \theta_m)$, $\hat{g}(x) = (g_1(x), \dots, g_m(x))$ — несмещенная оценка для $g(\theta) = (g_1(\theta), \dots, g_m(\theta))$. Матрица $B(\hat{g}) = (B_{ij}(\theta))$ — ковариационная матрица компонент оценки для $\hat{g}(x)$, т. е. $B_{ij}(\theta) = M_{\theta} (\hat{g}_i(x) - g_i(\theta)) (\hat{g}_j(x) - g_j(\theta))$. Допустим, что $B(\hat{g})$ — невырожденная, т. е.

существует обратная матрица $B^{-1}(\hat{g}) = (B^{ij}(\theta))$. Эллипсоид

$$\mathcal{E}(\hat{g}) = \{\theta: \sum (\hat{g}_i(x) - g_i(\theta)) B^{ij}(\theta_0) \times (\hat{g}_j(x) - g_j(\theta)) \leq k + 2\} \subset \Theta$$

называют эллипсоидом рассеяния.

Качество оценок можно сравнивать на основе их эллипсоидов рассеяния. Например, если эллипсоид рассеяния оценки $\hat{g}^{(1)}$ лежит внутри эллипсоида рассеяния оценки $\hat{g}^{(2)}$, т. е. $\mathcal{E}(\hat{g}^{(1)}) \subseteq \mathcal{E}(\hat{g}^{(2)})$, то оценку $\hat{g}^{(1)}$ считают предпочтительней оценки $\hat{g}^{(2)}$.

Эллипсоиды рассеяния особенно часто используют при графической иллюстрации полученных результатов в случае $m = 2$. При этом значение $B^{ij}(\theta_0)$ заменяют оценкой $B^{ij}(\hat{\theta})$, где $\hat{\theta}$ находят одним из способов, например, на основе метода моментов: $\hat{g}(x) = g(\hat{\theta})$.

Замечание. Если рассмотреть случайный вектор, равномерно распределенный внутри эллипсоида рассеяния $\mathcal{E}(\hat{g})$, то ковариационная матрица компонент такого вектора совпадает с матрицей $B(\hat{\theta})$.

Пусть: (i) $\hat{g}(x)$ — несмещенная оценка для $g(\theta)$; (ii) статистическая параметрическая модель $(\mathcal{X}, \mathcal{P})$, $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$ допускает достаточную статистику $T = T(x)$; (iii) функция риска $W(g', g'')$ строго выпуклая (вниз) функция по первому аргументу. Тогда условное математическое ожидание

$$\tilde{g}(T) = M_{\theta} (\hat{g}(x) | T(x)) \quad (4)$$

также является несмещенной оценкой для $g(\theta)$ и среднее значение риска

$$R(\tilde{g}, \theta) \leq R(\hat{g}, \theta).$$

Соотношение (4) дает возможность, исходя из простых несмещенных оценок, получать более точные несмещенные оценки.

Если T — минимальная (ограниченно) полная достаточная статистика, то в классе несмещенных оценок для функции $g(\theta)$ оценка \tilde{g} , найденная по формуле (4), является единственной оптимальной несмещенной оценкой, т. е. единственной несмещенной оценкой, обладающей наименьшим значением

среднего риска при любых возможных значениях истинного параметра θ_0 .

Пример 1. Пусть $0 < t_{i1} < < t_{id_i} < T_i$ — точки пуассоновских случайных процессов с параметром λ , $i = 1, \dots, m$, $\theta = \lambda \in R_+^1$; $g(\lambda) = e^{-\lambda T_0}$ — вероятность безотказной работы в течение заданного времени $T_0 < T_i$; $T = T(d_1, \dots, d_m) = \sum_{i=1}^m d_i$

— минимальная ограниченно полная достаточная статистика. Использование формулы (4) дает оптимальную в смысле минимума дисперсии оценку

$$\hat{g}(T) = \left(1 - \frac{T_0}{\sum_{i=1}^m T_i} \right)^T,$$

являющуюся несмещенной оценкой для $g(\lambda) = e^{-\lambda T_0}$. В качестве исходной оценки можно взять, например, индикаторную статистику $g = I(t_{i1} > T_0)$.

Оценка $\hat{g}^{(1)}$ предпочтительней оценки $\hat{g}^{(2)}$ при заданной функции риска $W(g', g'')$, если для средних рисков выполнено неравенство

$$R(\hat{g}^{(1)}, \theta) \leq R(\hat{g}^{(2)}, \theta), \theta \in \Theta. \quad (5)$$

Причем существует такое значение θ , при котором неравенство (5) является строгим.

Оценку \hat{g} называют **допустимой** в некотором классе оценок \mathcal{H} , если в этом классе нет оценок предпочтительней оценки \hat{g} .

Пример 2. Если $x_n = (y_1, \dots, y_n)$, где y_i — значения н. о. р. случайных величин $y_i \sim N(\mu, \sigma^2)$, функция риска $W(\mu', \mu'') = (\mu' - \mu'')^2$, то оценка $x =$

$$= \frac{1}{n} \sum_{i=1}^n x_i \text{ является допустимой в}$$

классе несмещенных оценок среднего μ .

Замечание. Если $F(x) = P\{y_i \leq x\} \neq \Phi\left(\frac{x-\mu}{\sigma}\right)$, т. е. не совпадает с нормальным распределением, то в условиях примера \bar{x} не является допустимой оценкой для μ . Таким образом, требование допустимости является сильным ограничением.

Пример 3. Если в условиях примера 2, параграфа 3 функция риска такова, что средний риск несмещенной оценки совпадает с ее дисперсией, то в классе

несмещенных оценок для $g(\theta) = \sum_{i=1}^N Y_i$

оценка Горвица — Томпсона является допустимой оценкой.

Пусть существуют частные производные логарифма функции правдоподобия $\frac{\partial l(\theta | x)}{\partial \theta_i}$, $i = 1, \dots, m$ и

$$I_{ij}(\theta) = M_{\theta} \frac{\partial l(\theta | x)}{\partial \theta_i} \frac{\partial l(\theta | x)}{\partial \theta_j},$$

$$|I_{ij}(\theta)| < \infty.$$

Вектор-столбец

$$\lambda(\theta | x) = \left(\frac{\partial l(\theta | x)}{\partial \theta_1}, \dots, \frac{\partial l(\theta | x)}{\partial \theta_m} \right)^T$$

называют **вкладом**.

Матрицу $I(\theta) = (I_{ij}(\theta))$, $i, j = 1, \dots, m$ называют **информационной матрицей Фишера**.

Пусть $\mathcal{U}(g)$ — класс несмещенных оценок $\hat{g}(x) = (\hat{g}_1(x), \dots, \hat{g}_k(x))$ векторной функции $g(\theta) = (g_1(\theta), \dots, g_k(\theta))$, $k \leq m$, $\theta \in \Theta \subseteq R^m$.

Условия регулярности (Крамера—Рао). U_1). У логарифма функции правдоподобия существуют все первые и

вторые частные производные $\frac{\partial l(\theta | x)}{\partial \theta_i}$, $\frac{\partial^2 l(\theta | x)}{\partial \theta_i \partial \theta_j}$, $i, j = 1, \dots, m$. Возможно

двукратное дифференцирование тождества $M_{\theta} l \equiv 1$ по θ_i и θ_j .

$$U_2). M_{\theta} \lambda(\theta | x) = 0, \theta \in \Theta.$$

U_3). Информационная матрица Фишера $I(\theta)$ строго невырожденная, т. е. $\det I(\theta) > 0$, $\theta \in \Theta$.

U_4). Матрица $J(\theta) = \left(\frac{\partial g_i(\theta)}{\partial \theta_j} \right)$, $i = 1, \dots, k$, $j = 1, \dots, m$ определена для всех $\theta \in \Theta$.

U_5). Класс несмещенных оценок $\hat{g} \in \mathcal{U}(g)$ функции $g(\theta)$ не пуст, тождество $M_{\theta} \hat{g} = g(\theta)$ можно дифференцировать по θ_i , $i = 1, \dots, m$.

Если выполнены условия $Y_1 \dots Y_n$, то для несмещенной оценки $\hat{g} \in \mathcal{U}(g)$ и функции риска

$$W(\hat{g}, g) = \|\hat{g}(x) - g(\theta)\|^2 = \sum_{i=1}^k (\hat{g}_i(x) - g_i(\theta))^2$$

выполнено неравенство (Крамера — Рао)

$$M_\theta \|\hat{g}(x) - g(\theta)\|^2 \geq M_\theta \|a(\theta | x)\|^2, \quad (6)$$

где $a(\theta | x) = J(\theta) I^{-1}(\theta) \lambda(\theta | x)$.

Правая часть неравенства (6) не зависит от выбора оценки $\hat{g} \in \mathcal{U}(g)$ и называется **границей Крамера—Рао**.

Пусть $k=1$, т. е. оцениваемая функция является скалярной величиной, тогда при выполнении условий регулярности дисперсия несмещенной оценки \hat{g} для $g(\theta)$ удовлетворяет неравенству

$$\text{Var}(\hat{g}) \geq \sum_{i,j=1}^m \frac{\partial g(\theta)}{\partial \theta_i} I^{ij}(\theta) \frac{\partial g(\theta)}{\partial \theta_j}, \quad (7)$$

где $I^{ij}(\theta)$ — элементы матрицы $I^{-1}(\theta)$, обратной информационной матрице Фишера $I(\theta)$.

Замечание. Неравенство (7) показывает, что дисперсия несмещенной оценки не может быть сделана сколь угодно малой при ограниченном объеме статистических данных.

Равенство в неравенствах (6)—(7) достигается только в том случае, когда несмещенная оценка для всех $\theta \in \Theta$ имеет вид

$$\hat{g}(x) = g(\theta) + a(\theta | x). \quad (8)$$

Пример 4. Пусть $x = (y_1, \dots, y_n)$, где y_i — н. о. р. случайные величины $P(t_i > t) = e^{-t/\theta}$, θ — среднее время безотказной работы. Здесь $k=1$, $m=1$, существует несмещенная оценка

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i,$$

$$\lambda(\theta | x) = \left(\sum_{i=1}^n y_i / \theta^2 \right) - n/\theta, \quad I(\theta) = M_\theta \lambda(x, \theta)^2 = \frac{n}{\theta^2}, \quad \tau(\theta) = 1.$$

Нижняя граница Крамера — Рао, т. е. правая часть неравенства (7)

$I^{-1}(\theta) = \theta^2/n$. Для $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i$ выполнено равенство (8), поэтому дисперсия оценки $\hat{\theta}$ достигает минимально возможного значения при каждом $\theta \in \Theta$.

Пусть (X^n, \mathcal{F}^n) — последовательность статистических моделей, соответствующих процессу накопления данных при $n \rightarrow \infty$. Последовательность оценок $\hat{g}_n = \hat{g}_n(x_n)$, $x_n \in X^n$, называют **состоятельной оценкой** для $g(\theta)$, если при $n \rightarrow \infty$ $\hat{g}_n \rightarrow g(\theta_0)$ по вероятности при любом истинном значении параметра $\theta_0 \in \Theta$. Если сходимость $\hat{g}_n \rightarrow g(\theta_0)$ почти наверное при любом $\theta_0 \in \Theta$, то \hat{g}_n называют **строго состоятельной оценкой**.

Пример 5. В условиях примера 3, параграфа 3 эмпирическая ф. р. $F_n(x)$ является строго состоятельной оценкой значения ф. р. $F(x)$ в любой точке x . При $n \rightarrow \infty \sup_x |F_n(x) - F(x)| \rightarrow 0$ почти наверное (теорема Гливенко).

Пусть $x_n = (y_1, \dots, y_n)$, y_i — н. о. р. случайные величины, имеющие п. р. $f(y, \theta)$, $\theta \in \Theta$. Пусть выполняются следующие условия:

A_1) Множество Θ является замыканием ограниченного открытого подмножества в R^m .

A_2) п. р. $f(x, \theta)$ — непрерывная функция по $\theta \in \Theta$.

A_3) Для любого $\theta \in \Theta$ и достаточно малого $\varepsilon = \varepsilon(\theta) > 0$

$$M_\theta \left\{ \sup_{\|\hat{\theta} - \theta\| < \varepsilon} |\ln f(y_i, \theta)| \right\} < \infty.$$

При выполнении условий A_1 — A_3 оценка максимального правдоподобия $\hat{\theta}_n$ является строго состоятельной оценкой параметра θ .

Пусть выполнены все условия $A_1 - A_3$ и условия регулярности Крамера — Рао $\mathcal{Y}_1 - \mathcal{Y}_5$, а также условие, что п. р. $f(x, \theta) > 0$ на множестве значений x , которое не зависит от параметра θ . Наблюдению y_i сопоставим информационную матрицу Фишера

$$\tilde{I}_m = (\tilde{I}_{ij}(\theta)), \tilde{I}_{ij}(\theta) = M_\theta \frac{\partial \ln f(y, \theta)}{\partial \theta_i} \frac{\partial \ln f(y, \theta)}{\partial \theta_j} f(y, \theta) dy,$$

$i, j = 1, \dots, m$. При $n \rightarrow \infty$ асимптотическое распределение о. м. п. $\hat{\theta}_n$ является нормальным порядка $1/\sqrt{n}$, т. е.

$$\sqrt{n}(\hat{\theta}_n - \theta) \sim \mathcal{N}_m(0, \Sigma), n \rightarrow \infty,$$

причем $\Sigma = \tilde{I}^{-1}(\theta)$. Ковариационная матрица уклонений $\sqrt{n} M_\theta(\hat{\theta}_i - \theta_i) \times (\hat{\theta}_j - \theta_j)$ компонент оценки сближается с матрицей, обратной информационной матрице Фишера для одного наблюдения, соответствующей значению параметра θ , порождающего распределения.

5. ДОВЕРИТЕЛЬНОЕ ОЦЕНИВАНИЕ

Статистику $T(x)$ называют нижней γ -доверительной границей функции $g(\theta)$, если

$$\inf_{\theta \in \Theta} P_\theta \{T(x) \leq g(\theta)\} \geq \gamma. \quad (9)$$

Статистику $T(x)$ называют верхней γ -доверительной границей функции $g(\theta)$, если

$$\inf_{\theta \in \Theta} P_\theta \{g(\theta) \leq T(x)\} \geq \gamma. \quad (10)$$

Если для двух статистик $T_1(x) \leq T_2(x)$

$$\inf_{\theta \in \Theta} P_\theta \{T_1(x) \leq g(\theta) \leq T_2(x)\} \geq \gamma,$$

(11)

то случайный интервал $[T_1(x), T_2(x)]$ называют γ -доверительным интервалом для $g(\theta)$. Если в (9)–(11) достигается минимальное значение, то число γ называют коэффициентом доверия. В про-

тивном случае говорят, что коэффициент доверия не менее γ .

Пример 1. Пусть наблюдаются $t_1 < \dots < t_{(r)}$ первые r значений N случайных величин, имеющих экспоненциальную ф. р. $F(t) = 1 - e^{-\lambda t}$ Статистики $T_1 = \frac{\chi_{\varepsilon_1}^2(2r)}{2s_r}$, $T_2 = \frac{\chi_{1-\varepsilon_2}^2(2r)}{2s_r}$, где $\chi_p^2(l)$ соответствует p -квантили χ^2 -распределения с l степенями свободы, $\varepsilon_1 + \varepsilon_2 = 1 - \gamma$, $S_r = t_{(1)} + \dots + t_{(r-1)} + (N - r)t_r$ — величина суммарной наработки в момент наблюдения $t_{(r)}$. Случайный интервал $[T_1, T_2]$ является γ -доверительным интервалом. Более того, при любом значении λ

$$P_\lambda \{T_1 \leq \lambda \leq T_2\} \equiv \gamma.$$

Пусть $\varepsilon_1, \dots, \varepsilon_N$ — последовательность испытаний Бернулли, $p = P\{\varepsilon_i = 1\}$, $P\{\varepsilon_i = 0\} = 1 - p$, $d = \sum_{i=1}^N \varepsilon_i$. Нижнюю γ -доверительную границу для вероятности p находят как решение уравнения

$$\sum_{c=0}^d \binom{N}{c} (1 - p_*)^c p_*^{N-c} = 1 - \gamma.$$

При малых значениях отношения d/N приближенное значение нижней γ -доверительной границы находят по формуле

$$p_* = \frac{4N - 2d - \chi_\gamma^2(2d + 2)}{4N - 2d + \chi_\gamma^2(2d + 2)}.$$

Каждому значению данных $x \in X$ статистической параметрической модели $(\mathcal{X}, \mathcal{P})$, $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ сопоставим подмножество $C_x \subseteq \Theta$. Если зафиксировать θ и рассмотреть подмножество $B_\theta = \{x : \theta \in C_x\}$, то события $\{\theta \in C_x\}$ и $\{x \in B_\theta\}$ являются эквивалентными. Если для любого $\theta \in \Theta$

$$P_\theta \{\theta \in C_x\} = P_\theta \{x \in B_\theta\} \geq \gamma, \quad (12)$$

то систему подмножеств $\{C_x\}$ называют системой γ -доверительных множеств. Если в соотношении (12) для некоторого θ достигается равенство, то γ называют коэффициентом доверия системы γ -доверительных множеств.

Если $\{C_x\}$ — система γ -доверительных множеств, то верхнюю и нижнюю γ -доверительные границы для функции $g(\theta)$ находят по формулам

$$g_*(x) = \inf_{\theta \in C_x} g(\theta), \quad g^*(x) = \sup_{\theta \in C_x} g(\theta). \quad (13)$$

Построение γ -доверительных границ для скалярной функции $g(\theta)$ можно основывать исходя из вещественной статистики $T(x)$. Пусть $F(t, \theta)$ — ф. р. этой статистики; положим

$$u_\gamma(\theta) = \inf \{t: F(t, \theta) \geq \gamma\}, \quad v_\gamma(\theta) = \sup \{t: 1 - F(t, \theta) \geq \gamma\}.$$

Определим две функции

$$u_\gamma(g) = \sup_{\theta \in \Gamma_g} u_\gamma(\theta), \quad v_\gamma(g) = \inf_{\theta \in \Gamma_g} v_\gamma(\theta),$$

где $\Gamma_g = \{\theta: g(\theta) = g\}$ — атом функции $g(\theta)$ уровня g .

Если $u_\gamma(g)$ и $v_\gamma(g)$ являются неубывающими по g функциями, то нижнюю $g_* = g_*(T)$ и верхнюю $g^* = g^*(T)$ γ -доверительные границы находят как решения уравнений

$$T(x) = u_\gamma(g_*), \quad T(x) = v_\gamma(g^*). \quad (14)$$

Если же $u_\gamma(g)$ и $v_\gamma(g)$ являются невозрастающими функциями аргумента g , то соответствующие γ -доверительные границы находят как решения уравнений

$$T(x) = u_\gamma(g^*), \quad T(x) = v_\gamma(g_*). \quad (15)$$

Пусть $(\mathcal{X}^n, \mathcal{P}^n)$, $\mathcal{P}^n = \{P^n, \theta \in \Theta\}$ — последовательность статистических моделей, соответствующих процессу накопления данных при $n \rightarrow \infty$, а Θ^n — неубывающая последовательность подмножеств $\Theta^n \subseteq \Theta^{n+1} \bigcup_{n=1}^{\infty} \Theta^n = \Theta$. Систему подмножеств $\{C_{x^n}\}$ множества Θ , $x^n \in X^n$ называют асимптотической системой γ -доверительных множеств, если

$$\lim_{n \rightarrow \infty} \inf_{\theta \in \Theta} P_\theta^n \{\theta \in C_{x^n}\} \geq \gamma. \quad (16)$$

Последовательность статистики $\{T^n(x)\}$ называют асимптотической нижней (верхней) γ -доверительной границей для функции $g(\theta)$, если соответственно выполняются следующие соотношения:

$$\lim_{n \rightarrow \infty} \inf_{\theta \in \Theta^n} P_\theta \{T^n(x) \leq g(x)\} \geq \gamma; \quad (17)$$

$$\lim_{n \rightarrow \infty} \inf_{\theta \in \Theta^n} P_\theta \{g(x) \leq T^n(x)\} \geq \gamma. \quad (18)$$

Аналогично определяют асимптотически γ -доверительный интервал $[T_1^n(x), T_2^n(x)]$:

$$\lim_{n \rightarrow \infty} \inf_{\theta \in \Theta^n} P_\theta \{T_1^n(x) \leq g(\theta) \leq T_2^n(x)\} \geq \gamma. \quad (19)$$

Если найдется такая последовательность значений параметров $\{\theta^n\}$, что при $\theta = \theta^n$ в соотношениях (16)–(19) достигается в пределе значение γ , то значение γ называют асимптотическим коэффициентом доверия.

6. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Статистической гипотезой называют утверждение о том, что распределение, порождающее данные x , принадлежит заданному подмножеству распределений. Для параметрических семейств гипотезы определяются подмножествами параметров. В параметрическом семействе $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ выделяют исходную (основную) гипотезу H_0 о том, что $\theta \in \Theta_0 \subset \Theta$, и конкурирующую (альтернативную) гипотезу H_1 , соответствующую значениям $\theta \in \Theta_1 \subset \Theta$, $\Theta_1 \cap \Theta_0 = \emptyset$. Гипотезы, которым соответствует только одно вероятностное распределение, называют простыми, в противном случае гипотезы называют сложными.

Если Θ — подмножество r -мерного пространства, $\theta = (\theta_1, \dots, \theta_r)$, а гипотезы Θ_i выражаются в терминах только части координат θ_i с номерами i_1, \dots, i_k , $k < r$, то оставшиеся значения координат носят название мешающих параметров.

Пример 1. Пусть семейство вероятностных распределений \mathcal{P} для данных $x = (x_1, \dots, x_n)$, где x_i — n н. о. р. — случайные величины, имеющие плотности вида

$$p(x, \theta) = \begin{cases} 0, & x \leq t, \\ \lambda \exp(-\lambda(x-t)), & x > t, \theta = (t, \lambda), \end{cases}$$

тогда при проверке сложных гипотез о значении абсолютно гарантированного времени безотказной работы t $\Theta_0 = \{\theta \mid t \leq \tau_0\}$, $\Theta_1 = \{\theta \mid t > \tau_0\}$ параметр λ является мешающим.

Пример 2. Пусть \mathcal{P} — семейство распределений Вейбулла — Гнеденко. Тогда гипотеза о том, что $F(t) = 1 - e^{-0,1t}$ $t > 0$, является простой, а гипотеза, соответствующая $\Theta_0 = \{(\lambda, p) \mid p = 1, \lambda > 0\}$, т. е. гипотеза об экспоненциальности распределения, является сложной.

Критерием называют правило, следуя которому по результатам наблюдений x делают утверждение, что данные либо согласуются с гипотезой H_0 , а конкурирующая гипотеза H_1 отвергается, либо данные противоречат гипотезе H_0 и согласуются с конкурирующей гипотезой.

Выделяется класс критериев значимости, в которых конкурирующие гипотезы не выделяются, а задается лишь исходная гипотеза H_0 , которую по наблюдаемым данным надо принять или отвергнуть.

Критерии проверки статистических гипотез могут быть **рандомизированными**, т. е. при данных x с вероятностью $\varphi(x)$ отвергается H_0 , а с вероятностью $(1 - \varphi(x))$ — принимается. Функция $\varphi(x)$ характеризует критерий и называется **критической**. Если $\varphi(x)$ принимает только значения 0 и 1, то критерий называют **нерандомизированным**, а множество $R = \{x \mid \varphi(x) = 1\}$ — **критическим**. R соответствует данным, при наблюдении которых H_0 отвергается. Критерий, характеризующий критической функцией φ , кратко обозначается как φ -критерий.

При использовании φ -критерия вероятность $\beta_\varphi(\theta)$ отвержения гипотезы H_0 , рассматриваемую как функцию

параметра $\theta \in \Theta$, называют **функцией мощности φ -критерия**,

$$\beta_\varphi(\theta) = E_\theta \varphi = \int_{\mathcal{X}} \varphi(x) P(dx).$$

Для нерандомизированного φ -критерия

$$\beta_\varphi(\theta) = P_\theta(R),$$

т. е. функция мощности равна вероятности того, что данные x принадлежат критической области.

Если $\theta \in \Theta_0$, то $\beta_\varphi(\theta)$ равна вероятности ошибочного отвержения правильной гипотезы H_0 . Максимальное значение вероятностей таких ошибок называют **размером критерия**, т. е. размер φ -критерия равен

$$\alpha = \sup_{\theta \in \Theta} \beta_\varphi(\theta).$$

Если φ -критерий — нерандомизированный, то α называют **размером критической области**. Значения $\beta_\varphi(\theta)$, когда $\theta \in \Theta_0$, называют **вероятностями ошибок 1-го рода**; когда же $\theta \in \Theta_1$, тогда $1 - \beta_\varphi(\theta)$ называют **вероятностями ошибок 2-го рода**.

Во многих критериях решение о принятии или отвержении гипотез H_0, H_1 основывается на значениях специальным образом подобранных статистик, называемых **тестовыми**. Пусть T — тестовая статистика, $T(x)$ — ее значение, соответствующее данным x . Если гипотеза H_0 отвергается при наблюдении больших значений $T(x)$, то вероятность $P_\theta(T > T(x)) = \alpha(x)$ называют **фактическим уровнем значимости критерия**.

φ -критерий называют **равномерно наиболее мощным (РНМ-критерием)** в классе критериев с уровнем значимости α , если $\beta_\varphi(\theta)$ достигает максимума в этом классе для всех значений $\theta \in \Theta$.

РНМ-критерии существуют для небольшого числа математических моделей, как правило связанных с экспоненциальными семействами распределений.

φ -критерий называют **несмещенным**, если вероятность ошибочного отвержения гипотезы H_0 , когда она верна, не больше правильного отвержения H_0 , когда верна гипотеза H_1 . В терминах

функции мощности это означает, что

$$\sup_{\theta \in \Theta_0} \beta_{\Phi}(\theta) \leq \inf_{\theta \in \Theta_1} \beta_{\Phi}(\theta).$$

РНМ-критерий в классе несмещенных критериев размера α называют **равномерно наиболее мощным несмещенным критерием** (РНМН-критерием) размера α .

Два критерия с критическими функциями φ_1 и φ_2 называют **эквивалентными**, если $\beta_{\varphi_1}(\theta) \equiv \beta_{\varphi_2}(\theta)$ для всех $\theta \in \Theta$.

Если семейство вероятностных распределений \mathcal{P} рассматриваемой математической модели имеет достаточную статистику S , то условное математическое ожидание критической функции φ относительно S снова дает критическую функцию

$$\varphi^*(S) = E_{\theta}(\varphi/S).$$

φ^* -критерий эквивалентен исходному критерию φ .

Пусть гипотезы: $H_0 = \{\theta \in \Theta_0\}$, $H_1 = \{\theta \in \Theta_1\}$, где $\Theta_i = \{\theta = (\theta_1 - \theta_0)\} \in R^m$. Граничную точку $\theta_0 \in \Theta$ называют **регулярной** в направлении $d = (d_1, \dots, d_m)$, если для $0 < t < \varepsilon$ $\theta_0 + dt \in \Theta_1$. Рассмотрим два критерия k_i , определяемые тестовыми статистиками $T_n^{(i)}$, $i = 1, 2$, где $n = 1, 2$, соответствует накоплению данных.

Предположим, что для $\theta_n(d) = \theta_0 + \frac{dt}{\sqrt{n}}$ статистики $T_n^{(i)}$ являются асимптотически нормальными

$$\sqrt{n} \frac{T_n^{(i)}(\theta_n(d)) - \mu_n^{(i)}(\theta_n(d))}{\sigma_n^{(i)}(\theta_n(d))} \sim$$

$$\sim N(0, 1),$$

причем $\frac{\sigma_n(\theta_n(d))}{\sigma_n(\theta_0)} \rightarrow 1$, а непрерывно дифференцируемые по θ функции $\mu_n^{(i)}(\theta_n(d)) \rightarrow \mu_0^{(i)}(\theta_0)$ и $D_d \mu_n^{(i)}(\theta_0) \rightarrow D_d \mu_0^{(i)}(\theta_0)$, где $D_d \mu_n^{(i)}(\theta) = \lim_{h \rightarrow 0} \frac{1}{h} \times \mu_n^{(i)}(\theta_0 + dh) - \mu_n^{(i)}(\theta_0)$ — производная по направлению d , $i = 1, 2$.

Тогда **асимптотической эффективностью по Питмену** критерия K_1 относительно критерия K_2 в направлении d для регулярной граничной точки θ_0 называют

$$e(T_1; T_2) = \frac{\left[\frac{D_d \mu_0^{(1)}(\theta_0)}{\sigma^{(2)}(\theta_0)} \right]^2}{\left[\frac{D_d \mu_0^{(2)}(\theta_0)}{\sigma^{(1)}(\theta_0)} \right]^2}.$$

Асимптотическая эффективность по Питмену соответствует отношению объемов выборок, дающих при различении сближающихся вдоль направления d гипотез $\theta \approx \theta_0$ и $\theta = \theta_0 \rightarrow \theta_0$ одинаковые значения ошибок 1-го и 2-го рода.

7. ИНВАРИАНТНОСТЬ

Если статистическая модель обладает свойствами симметрии, то могут быть использованы методы теории групп. Пусть $G = \{g\}$ — группа преобразований множества данных \mathfrak{X} . Для любых данных $x \in \mathfrak{X}$, $gx \in \mathfrak{X}$. Положим $gA = \{x : g^{-1}x \in A\}$, где $A \in \mathfrak{X}$. Предполагается, что из $A \in \mathfrak{B}$ следует $gA \in \mathfrak{B}$. Группу преобразований $\tilde{G} = \{\tilde{g}\}$ множества параметров Θ , удовлетворяющую при любом $A \in \mathfrak{B}$, $\theta \in \Theta$ и $g \in G$ соотношению

$$P_{\tilde{g}\theta} \{gA\} = P_{\theta} \{A\}, \quad (20)$$

где $\tilde{g} \in \tilde{G}$, называют **индуцированной**. Между преобразованиями g и \tilde{g} существует взаимно однозначное соответствие.

Пример 1. Пусть $x = (x_1, \dots, x_n)$ — значения n о. р. случайных величин с ф. р. $F(s) = 1 - \exp(-\lambda s)$, $s > 0$, $\Theta = \{\lambda, \lambda > 0\}$. Группа $G = \{g_c, c > 0\}$ определяет преобразование данных

$$g_c x = (cx_1, \dots, cx_n).$$

Из соотношения (20) находим соответствующее индуцированное g_c преобразование \tilde{g}_c . Это преобразование определяется соотношением

$$\tilde{g}_c \lambda = \lambda/c, \quad \lambda \in \Theta.$$

Статистику $t(x)$ называют **инвариантной** статистикой, если ее значение не

меняется при преобразованных данных, т. е. $t(gx) = t(x)$. Орбитой G_x , соответствующей $x \in \mathfrak{X}$, называют множество

$$G_x = \{gx: g \in G\},$$

т. е. множество, получаемое из x путем его преобразований элементами $g \in G$. Статистику называют **максимальным инвариантом**, если ее атомы совпадают с орбитами.

Пример 2. В модели примера 1, параграфа 7 максимальными инвариантами будут статистики

$$t_1(x_n) = \left(\frac{x_1}{\bar{x}}, \quad \frac{x_n}{\bar{x}} \right),$$

$$t_2(x_n) = \left(\frac{x_2}{x_1}, \quad \frac{x_n}{x_1} \right).$$

Пусть Θ_0 и Θ_1 — подмножества параметров, соответствующих гипотезам H_0 и H_1 . Гипотезы H_0 , H_1 называют **инвариантными относительно группы преобразований \tilde{G}** , если для любого $\tilde{g} \in \tilde{G}$, $\tilde{g}\Theta_0 = \Theta_0$, $\tilde{g}\Theta_1 = \Theta_1$, где $\tilde{g}\theta_i = \{g\theta: \theta \in \Theta_i\}$.

Пусть гипотезы H_0 , H_1 инвариантны относительно \tilde{G} . Тогда φ -критерий называют инвариантным, если его критическая функция φ инвариантна относительно группы преобразований \tilde{G} , т. е. $\varphi(x) = \varphi(gx)$ для любого $x \in \mathfrak{X}$, $g \in G$.

Принцип инвариантности состоит в том, что наиболее мощный критерий ищется в классе инвариантных критериев размера α . В ряде математических моделей удается найти равномерно наиболее мощный инвариантный критерий (РНИ-критерий) размера α .

СПИСОК ЛИТЕРАТУРЫ

1. **Беляев Ю. К.** Множительные оценки вероятности безотказной работы. — Изв. АН СССР. Техническая кибернетика, 1985, № 4, с. 45—59.
2. **Беляев Ю. К., Чепурин Е. В.** Основы математической статистики. М.: Изд-во МГУ Ч. 1, 1982, 100 с., Ч. 2 и 3, 1983. 145 с.
3. **Бикел П., Доксам К.** Математическая статистика/Пер. с англ. М.: Финансы и статистика. Вып. 1, 1983, 278, с., вып. 2, 1983, 254 с.
4. **Вопросы математической теории надежности/Е. Ю. Барзилович, Ю. К. Беляев, В. А. Каштанов и др.** М.: Радио и связь, 1983. 376 с.
5. **Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д.** Математические методы в теории надежности. М.: Наука, 1965. 524 с.
6. **Ивченко Г. И., Медведев Ю. И.** Математическая статистика. М.: Высшая школа, 1984. 248 с.
7. **Крамер Г.** Математические методы статистики/Пер. с англ. М.: Мир, 1975. 648 с.

1. ЭЛЕМЕНТЫ ВЫПУКЛОГО АНАЛИЗА

1.1. ВЫПУКЛЫЕ МНОЖЕСТВА

Пусть $x, y \in E_n$, $x^T = (x_1, \dots, x_n)$ — элементы евклидова пространства E_n ,

$$\|x\| = (x, x) = \sum_{i=1}^n x_i^2.$$

Будем говорить, что последовательность $\{x_k\}$ точек из E_n сходится к точке x при $k \rightarrow \infty$, т. е. $\lim_{k \rightarrow \infty} x_k = x$, если

$$\lim_{k \rightarrow \infty} \|x_k - x\| = 0.$$

ε -окрестностью точки $x \in U(x, \varepsilon)$ будем называть множество $U(x, \varepsilon) = \{y \in E_n \mid \|y - x\| \leq \varepsilon\}$ ($\varepsilon \in R^+$).

Множество $X \subseteq E_n$ называют **замкнутым**, если оно содержит все свои предельные точки (т. е. такие точки, что для любой окрестности каждой из них принадлежит бесконечно много точек из X).

Точку $x \in X$ называют **внутренней точкой** множества X , если существует такая его окрестность, все точки которой принадлежат X .

Точку $x \in X$ называют **граничной точкой** множества X , если любая ее окрестность содержит как точки, принадлежащие X , так и точки, не принадлежащие X . Множество состоящее из всех граничных точек множества X , называют **границей** множества X .

Множество $X \subseteq E_n$ называют **выпуклым**, если вместе с любыми двумя точками $x, y \in X$ ему принадлежит и соединяющий их отрезок $[x, y]$.

Выпуклость множества X означает, что из $x, y \in X$ следует, что $z = \alpha x + (1 - \alpha)y \in X$ для всех $\alpha \in [0, 1]$.

Примеры: 1. Выпуклые множества на плоскости (E_2): отрезок, полупря-

мая, прямая, круг, треугольник, полуокружность, вся плоскость.

2. Пусть $X = \{x \in E_n \mid Ax \geq a, Bx = b\}$, где A, B — матрицы, тогда оно выпукло и замкнуто.

Проекцией точки $v \in E_n$ на выпуклое множество X называют такую точку $p \in X$, что

$$\|p - v\| = \inf_{x \in X} \|x - v\| = \delta,$$

где δ — расстояние точки v до множества X .

Теорема. Для любого выпуклого замкнутого множества X и любой точки $v \in E_n$ существует единственная точка $p \in X$, являющаяся проекцией v на X .

Теорема. Для того чтобы точка $p \in X$ была проекцией точки v на замкнутое выпуклое множество X , необходимо и достаточно, чтобы для всех $x \in X$ выполнялось неравенство $(x - p, v - p) \leq 0$.

Гиперплоскостью Γ в E_n называют множество вида $\Gamma = \{x \mid (c, x) = \lambda\}$, где $c \in E_n$, $c \neq 0$.

В пространстве E_n гиперплоскость Γ определяет два полупространства $\{x \mid (c, x) \leq \lambda\}$ и $\{x \mid (c, x) > \lambda\}$.

Теорема. Пусть $X \subseteq E_n$ — выпуклое множество, тогда для любой точки $v \in E_n$, внешней относительно замыкания \bar{X} множества X , существует такая гиперплоскость $\Gamma = \{x \mid (c, x) = \lambda\}$, что $(c, v) = \lambda$ и для всех $x \in X$ $(c, x) < \lambda$.

Замечание. Геометрический смысл теоремы: существует проходящая через точку v гиперплоскость Γ , такая, что X лежит в одном из полупространств.

Теорема. В любой граничной точке x^Γ выпуклого множества X существуют опорная гиперплоскость (т. е.

существуют $c \neq 0$ и λ , такие, что $\Gamma = \{x \mid (c, x) = \lambda\}$, $\lambda = (c, x^\Gamma)$ и для всех $x \in X$ $(c, x) \leq \lambda$.

Теорема. Пусть X_0 — непустое множество внутренних точек выпуклого множества X , Y — некоторое выпуклое множество и $X_0 \cap Y = \emptyset$. Тогда для множеств X и Y существует разделяющая гиперплоскость.

Точку $x \in X$ называют **крайней** (угловой), если в X не существует таких точек x' и x'' , $x' \neq x''$, что $x = \alpha x' + (1 - \alpha)x''$ при $\alpha \in (0, 1)$.

Примеры крайних точек: 1) для круга любая точка ограничивающей окружности является угловой; 2) вершины выпуклого многогранника — крайние точки.

Точку x называют **выпуклой комбинацией** точек x_1, \dots, x_N , если существуют такие числа $\alpha_i \geq 0$, $i = 1, \dots, N$, что

$$x = \sum_{i=1}^N \alpha_i x_i \text{ и } \sum_{i=1}^N \alpha_i = 1.$$

Примеры: 1. Любая внутренняя точка x круга является выпуклой комбинацией концов хорды, проходящей через x . 2. Любая точка треугольника является выпуклой комбинацией его вершин.

Теорема. Любая точка $x \in X$, где X — выпуклое, замкнутое, ограниченное множество, может быть представлена в виде выпуклой комбинации конечного числа крайних точек этого множества.

Теорема Фаркаша. Если для произвольной матрицы B существует такой вектор v , что для всех x , удовлетворяющих неравенству $Bx \leq 0$, будет $(v, x) \leq 0$, то всегда найдется вектор $u \geq 0$, такой, что $v = B^T u$.

Множество $K \subseteq E_n$ называют **конусом**, если из $x \in K$ следует $\lambda x \in K$ для всех $\lambda \geq 0$.

Теорема (о замкнутости конуса). Множество $Y = \{y : y = Ax, x \geq 0\}$ — замкнуто.

1.2. ВЫПУКЛЫЕ И ВОГНУТЫЕ ФУНКЦИИ

Скалярную функцию $\varphi(x)$, $x \in X$ называют **выпуклой** на выпуклом множестве X , если для любого $x, y \in X$

и $\alpha \in [0, 1]$ выполняется неравенство

$$\varphi(\alpha x + (1 - \alpha)y) \leq \alpha \varphi(x) + (1 - \alpha) \varphi(y).$$

Если

$$\varphi(\alpha x + (1 - \alpha)y) \geq \alpha \varphi(x) + (1 - \alpha) \varphi(y),$$

то $\varphi(x)$ называют **вогнутой**.

Если для любого $\alpha \in (0, 1)$ и

$$\varphi(\alpha x + (1 - \alpha)y) < \alpha \varphi(x) + (1 - \alpha) \varphi(y),$$

то $\varphi(x)$ называют **строго выпуклой** (вогнутой).

Примеры: 1. Пусть X — выпуклое множество, $\varphi(x)$ — непрерывна и

$$\varphi\left(\frac{x+y}{2}\right) \leq \frac{1}{2} \varphi(x) + \frac{1}{2} \varphi(y),$$

то $\varphi(x)$ — выпуклая функция. 2. Сумма выпуклых (вогнутых) функций является выпуклой (вогнутой) функцией. 3. Пусть A — положительно определенная симметрическая $n \times n$ матрица, $x \in E_n$, то $x^T A x$ — выпуклая функция.

Выпуклая функция $\varphi(x)$, $x \in X$ — выпуклое множество, имеет в любой внутренней точке $x \in X$ производную по любому направлению S ($\|S\| = 1$)

$$\frac{\partial \varphi}{\partial S}(x) = \lim_{\lambda \rightarrow 0} \frac{\varphi(x + \lambda S) - \varphi(x)}{\lambda}.$$

Выпуклая функция $\varphi(x)$ непрерывна в каждой внутренней точке $x \in X$, однако это неверно для граничных точек. Например, $x \in [0, 1]$, функция

$$\varphi(x) = \begin{cases} x, & x \in [0, 1) \\ 2, & x = 1. \end{cases}$$

— выпукла, но в точке $x = 1$ терпит разрыв первого рода.

Функция $\varphi(x)$, дифференцируемая на выпуклом замкнутом множестве X , выпукла тогда и только тогда, если для любых $x, y \in X$ будет

$$\varphi'(x) \cdot (x - y) \leq \varphi(y) - \varphi(x).$$

Направление $S \neq 0$ в точке $x \in X$ выпуклого множества называют **возможным**, если существует такое число $\lambda_0 > 0$, что для всех $\lambda \in [0, \lambda_0]$ $x + \lambda S \in X$.

Пример. Если $X = \{x \mid x \geq 0\}$, то в точке $x = 0$ любой вектор $S \geq 0$, $S \neq 0$ задает возможное направление, причем в точке $x^T = (x_1 = 0, x_2 > 0, \dots, x_n > 0)$ возможным направлением является любое $S^T = (s_1 \geq 0, s_2, \dots, s_n)$, где s_2, \dots, s_n — произвольные числа, $S \neq 0$.

Теорема. Для того чтобы точка x^* выпуклого множества X была точкой минимума выпуклой дифференцируемой на X функции $\varphi(x)$, необходимо и достаточно, чтобы для любого $x \in X$ выполнялось неравенство

$$(\varphi'(x^*), x - x^*) \geq 0.$$

Теорема. Если $\varphi(x)$ — строго выпукла на выпуклом множестве X и точка $x^* \in X$ такая, что $\varphi(x^*) = \min_{x \in X} \varphi(x)$, то для всех $x \in X$ и $x \neq x^*$ $\varphi(x) > \varphi(x^*)$.

Теорема. Для того чтобы точка x^* выпуклого множества X являлась точкой минимума выпуклой дифференцируемой на X функции $\varphi(x)$, необходимо и достаточно, чтобы

$$x^* = p(v^*), \text{ где } v^* = x^* - \lambda \varphi'(x^*),$$

$p(v^*)$ — проекция точки v^* на X .

Функцию $\varphi(x)$, определенную на выпуклом множестве X , называют **сильно выпуклой**, если существует константа $\rho > 0$, такая, что для любых $x, y \in X$ и $\alpha \in [0, 1]$ выполняется неравенство

$$\varphi(\alpha x + (1 - \alpha)y) \leq \alpha \varphi(x) + (1 - \alpha)\varphi(y) - \alpha(1 - \alpha)\rho \|x - y\|^2,$$

а ρ называют **параметром сильной выпуклости**.

Пример. Рассмотрим

$$\varphi(x) = (x, Bx) + (p, x),$$

где B — положительно определенная матрица; $\varphi(x)$ — сильно выпуклая функция.

Свойства сильно выпуклых функций $\varphi(x)$: 1) $\varphi(x)$ — один раз дифферен-

цируема на X , а X — выпуклое множество, то для всех $x, y \in X$ справедливо неравенство

$$(\varphi'(x) - \varphi'(y), x - y) \geq \rho \|x - y\|^2;$$

2) если X — выпуклое замкнутое множество, то:

а) для любой точки x^0 множество $X_0 = \{x \in X: \varphi(x) \leq \varphi(x^0)\}$ ограничено;

б) существует единственная точка x^* , такая, что

$$\varphi(x^*) = \min_{x \in X} \varphi(x);$$

3) если X — замкнутое выпуклое множество, то для всех $x \in X$ справедливо неравенство

$$\|x - x^*\| \leq \frac{2}{\rho} [\varphi(x) - \varphi(x^*)].$$

Если, кроме того, $\varphi(x)$ — один раз дифференцируема, то

$$\|x - x^*\| \leq \frac{1}{\rho} \|\varphi'(x)\|;$$

$$0 \leq \varphi(x) - \varphi(x^*) \leq \frac{1}{\rho} \|\varphi'(x)\|^2;$$

4) если $\varphi(x)$ — один раз дифференцируема на замкнутом выпуклом множестве X и существует такая константа $L > 0$, что для всех $x, y \in X$

$$\|\varphi'(x) - \varphi'(y)\| \leq L \|x - y\|,$$

то для любых $x, y \in X$, таких, что $\varphi(x) < \varphi(y)$, будет выполняться неравенство

$$\|\varphi'(x)\| \leq \sqrt{2 + 12 \frac{L^2}{\rho^2}} \|\varphi'(y)\|.$$

2. ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ

Постановка задачи выпуклого программирования. Рассмотрим множество $X = \{x \in J \mid f(x) \geq b\}$, где $f^T(x) = (f_1(x), \dots, f_m(x))$; $f_i(x)$, $i = 1, m$ — вогнутые непрерывные на J скалярные функции, а J — заданное выпуклое замкнутое множество (в частности, J может совпадать с E_n).

Множество X , когда $J = E_n$, выпукло, поэтому X — выпукло, когда и $J \subseteq E_n$. В силу непрерывности $f_i(x)$ и замкнутости J следует замкнутость множества X .

Пусть $\varphi(x)$ — выпукла, а X удовлетворяет указанным выше условиям. Рассмотрим основную задачу выпуклого программирования: $\min_{x \in X} \varphi(x)$, т. е.

требуется найти:

1) либо такую точку $x^* \in X$, что $\varphi(x^*) \in \min_{x \in X} \varphi(x)$;

x^* в дальнейшем будем называть оптимальной точкой, а множество X — допустимым множеством;

2) либо если не существует такого x^* , то найти

$$\varphi^* = \inf_{x \in X} \varphi(x);$$

3) либо убедиться, что $\varphi(x)$ не ограничена снизу на множестве X ;

4) либо убедиться в том, что $X = \emptyset$.

Условие регулярности (Слейтера).

Если для каждого $i = 1, \dots, m$ существует такая точка $x_i \in X$, что

$$f_i(x_i) > b_i,$$

то говорят, что множество X удовлетворяет условию регулярности.

Последнее эквивалентно условию Слейтера: существует такая точка $x \in$

$$\in X, \text{ что } \vec{f}(x) > \vec{b}.$$

Функция Лагранжа. Обозначим

$$h(x) = \vec{b} - \vec{f}(x). \text{ Функцию}$$

$$L(x, y) \stackrel{\Delta}{=} \varphi(x) + (y, h(x)),$$

где $x \in J, y \geq 0$, называют функцией Лагранжа для основной задачи выпуклого программирования.

Седловая точка. Пару (x^*, y^*) называют седловой точкой функции $L(x, y)$ на множестве $x \in J, y \geq 0$, если $x^* \in J, y^* \geq 0$ и

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*)$$

для всех $x \in J, y \geq 0$. Последние неравенства можно записать следующим образом:

$$L(x^*, y^*) = \min_{x \in J} \max_{y \geq 0} L(x, y) = \max_{y \geq 0} \min_{x \in J} L(x, y).$$

Теорема. Если пара (x^*, y^*) — седловая точка функции Лагранжа на множестве $x \in J, y \geq 0$, то x^* — оптимальная точка основной задачи выпуклого программирования.

Теорема Куна—Таккера. Пусть в основной задаче выпуклого программирования $\min_{x \in X} \varphi(x)$ множество $X =$

$= \{x \in J \mid \vec{f}(x) \geq \vec{b}\}$ обладает свойством регулярности. Тогда необходимым и достаточным условием оптимальности точки x^* является существование такого $y^* \geq 0$, чтобы пара (x^*, y^*) была седловой точкой функции Лагранжа $L(x, y)$ на множестве $x \in J, y \geq 0$.

Теорема. Если функции $\varphi(x)$ и $f_i(x), i = 1, \dots, m$ основной задачи выпуклого программирования непрерывно дифференцируемы на множестве $J \{x \mid x \geq 0\}$, то для того, чтобы пара (x^*, y^*) была седловой точкой функции Лагранжа в области $x \geq 0, y \geq 0$, необходимо и достаточно выполнение следующих условий:

$$\frac{\partial L}{\partial x}(x^*, y^*) \geq 0, \quad (x^*,$$

$$\frac{\partial L}{\partial x}(x^*, y^*) = 0, \quad x^* \geq 0;$$

$$\frac{\partial L}{\partial y}(x^*, y^*) \leq 0, \quad (y^*,$$

$$\frac{\partial L}{\partial y}(x^*, y^*) = 0, \quad y^* \geq 0.$$

$(-\varphi'(x) - f'_i(x))$ будем называть антиградиентами функций соответственно $\varphi(x)$ и $f_i(x)$.

Теорема. Пусть X — множество основной задачи выпуклого программирования — обладает свойством регулярности, а функции $\varphi(x), f_i(x)$ — непрерывно дифференцируемы на множестве $\Gamma = \{x \mid x \geq 0\}$. Тогда для оптимальности точки $x^* \in X$ необходимо и достаточно, чтобы антиградиент целевой функции можно было

представить в виде линейной комбинации:

$$-\varphi'(x^*) = - \sum_{i \in I(x^*)} y_i f'_i(x^*) - \sum_{j \in \mathfrak{J}(x^*)} v_j l_j,$$

где

$$I(x) = \{i: f_i(x) = b_i\}; \mathfrak{J}(x) = \{j: x_j = 0\}; -l_j - \text{внешняя нормаль относительно } X \text{ к граничной гиперплоскости } x_j = 0.$$

Пример. Задача с линейными ограничениями.

Пусть допустимое множество имеет вид

$$R_1 = \{x: Ax \geq b, x \geq 0\}.$$

Рассмотрим задачу: $\min_{x \in R_1} \varphi(x)$, где

$\varphi(x)$ — выпуклая непрерывно дифференцируемая функция.

Теорема. Для того чтобы у выпуклой непрерывно дифференцируемой функции $\varphi(x)$ в R_1 существовала оптимальная точка x^* , необходимо и достаточно существование такого $y^* \geq 0$, чтобы пара (x^*, y^*) была седловой точкой функции Лагранжа $L(x, y)$ в области $x \geq 0, y \geq 0$.

3. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ (ЛП)

3.1. ЗАДАЧА ЛП И ЕЕ СВОЙСТВА

Постановка задачи. Следующую задачу называют основной задачей линейного программирования (ОЗЛП)

$$\min_{x \in R_1} (c, x),$$

где $R_1 = \{x: Ax \geq b, x \geq 0\}$.

Двойственность. Двойственной задачей к ОЗЛП будем называть задачу

$$\max_{y \in Q_1} (b, y),$$

где $Q_1 = \{y: A^T y \leq c, y \geq 0\}$.

Две экстремальные задачи будем называть эквивалентными, если либо множества их решений совпадают, либо обе задачи не имеют решения.

Терминология. Матрицу A называют

матрицей условий ОЗЛП, а ее столбцы a_i — векторами условий ОЗЛП.

Вектор b называют вектором ограничения ОЗЛП. Точку $x \in R_1$ называют допустимой точкой или планом; крайнюю точку R_1 называют опорным планом. Решение задачи ЛП (оптимальную точку) называют оптимальным планом.

Обозначим через $L_1(x, y)$ функцию Лагранжа

$$L_1(x, y) = (c, x) + (y, b - Ax).$$

Теорема. Для того чтобы точка $x^* \in R_1$ была оптимальной для ОЗЛП, необходимо и достаточно, чтобы пара (x^*, y^*) была седловой точкой функции Лагранжа $L_1(x, y)$ в области $x \geq 0, y \geq 0$, т. е.

$$L_1(x^*, y) \leq L_1(x^*, y^*) \leq L_1(x, y^*).$$

Теорема (двойственности). Прямая и двойственная задачи либо обе имеют оптимальные точки x^* и y^* , причем $(c, x^*) = (b, y^*)$, либо обе их не имеют.

Теорема. Для любых $x \in R_1$ и $y \in Q_1$ справедливо неравенство $(c, x) \geq (b, y)$.

Если $x^* \in R_1$ и $y^* \in Q_1$, а $(c, x^*) = (b, y^*)$, то x^* и y^* — оптимальные точки для задач ОЗЛП и двойственной к ОЗЛП. Справедливо и обратное.

Если линейная форма (c, x) , $c \neq 0$ ограничена снизу на непустом R_1 , то существует точка $x^* \in R_1$, такая, что

$$(c, x^*) = \min_{x \in R_1} (c, x).$$

Теорема (существования). Для существования решения одной из двойственных задач необходимо и достаточно, чтобы $R_1 \neq \emptyset$ и $Q_1 \neq \emptyset$.

Если (c, x) ((b, y)) ограничена снизу (сверху) на R_1 (Q_1), то $Q_1 = \emptyset$ ($R_1 \neq \emptyset$).

Если $Q_1 = \emptyset$ ($R_1 = \emptyset$), а $R_1 \neq \emptyset$ ($Q_1 \neq \emptyset$), то R_1 (Q_1) не ограничено и (c, x) ((b, y)) не ограничено снизу (сверху) на R_1 (Q_1).

Теорема. Если $(c, x^*) = \min_{x \in R_1} (c, x)$

и существуют $x_i, i = \overline{1, M}$, такие, что

$x_i \in R_1$ и $x^* = \sum_{i=1}^M \alpha_i x_i$, где $\sum_{i=1}^M \alpha_i = 1$ и $\alpha_i > 0$, то $(c, x^*) = (c, x_1) = \dots = (c, x_m)$.

Теорема. Если ОЗЛП имеет решение x^* , то существует крайняя точка x , такая, что

$$(c, x) = \min_{x \in R_1} (c, x).$$

Для того чтобы точка $x \neq 0$ являлась крайней точкой множества R_1 , необходимо и достаточно, чтобы x удовлетворял неособой системе уравнений

$$(Ax)_i = b_i \quad i \in I(x);$$

$$x_j = 0 \quad j \in \mathfrak{F}(x),$$

где $I(x) = \{i \mid (Ax)_i = b_i\}$; $\mathfrak{F} = \{j \mid x_j = 0\}$.

Если точка $x > 0$ — крайняя, то систему линейно независимых векторов a_1, \dots, a_k в представлении

$$b = \sum_{i=1}^k x_i a_i \quad x_i > 0, \quad i = \overline{1, k}$$

называют базисом угловой точки, а матрицу $B = [a_1, \dots, a_k]$ — матрицей базиса угловой точки.

3.2. ДВОЙСТВЕННЫЕ ЗАДАЧИ ЛП СО СМЕШАННЫМИ ОГРАНИЧЕНИЯМИ

Задачи ЛП со смешанными ограничениями — это такие задачи ЛП, в которых допустимое множество задается системой равенств и неравенств, причем часть переменных может быть свободна от ограничений.

Введем обозначения:

$$I = \{i = \overline{1, m}\}, \quad \mathfrak{F} = \{j = \overline{1, n}\}, \quad I_1 \subseteq I,$$

$$I_2 = I \setminus I_1, \quad \mathfrak{F}_1 \subseteq \mathfrak{F}, \quad \mathfrak{F}_2 = \mathfrak{F} \setminus \mathfrak{F}_1.$$

Пару задач:

1) $\min (c, x)$	2) $\max (b, y)$
$(Ax)_i \geq b_i \quad i \in I_1$	$(A^T y)_j \leq c_j \quad j \in \mathfrak{F}_1$
$(Ax)_i = b_i \quad i \in I_2$	$(A^T y)_j = c_j \quad j \in \mathfrak{F}_2$
$x_j \geq 0 \quad j \in \mathfrak{F}_1$	$y_i \geq 0 \quad i \in I_1$

называют двойственными задачами со смешанными ограничениями.

Приведение задач со смешанными ограничениями к эквивалентным задачам. Обозначим через x_1 — вектор, составленный из компонент x_j вектора x , для которых $j \in \mathfrak{F}_1$, а x_2 составлен из компонент x_j вектора x , для которых $j \in \mathfrak{F}_2$. Им соответствуют векторы c_1 и c_2 ; аналогичным образом векторы $y_k, k = 1, 2$, составлены из компонент y_i вектора y , для которых $i \in I_k$, а b_k соответствует y_k ; для матрицы A введем обозначения клеток:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Тогда пара двойственных задач со смешанными ограничениями может быть переписана в виде:

$$1) \min [(c_1, x_1) + (c_2, x_2)]$$

$$A_{11}x_1 + A_{12}x_2 \geq b_1$$

$$A_{21}x_1 + A_{22}x_2 = b_2$$

$$x_1 \geq 0$$

$$2) \max [(b_1, y_1) + (b_2, y_2)],$$

$$A_{11}^T y_1 + A_{21}^T y_2 \leq c_1$$

$$A_{12}^T y_1 + A_{22}^T y_2 = c_2$$

$$y_1 \geq 0.$$

Каждое равенство можно заменить двумя неравенствами. Например, $\varphi = 0$ эквивалентно $\varphi \geq 0$ и $-\varphi \geq 0$ или $\varphi \leq 0$ и $-\varphi \leq 0$. Кроме того, вместо переменной z , на которую не наложены условия неотрицательности, введем неотрицательные переменные \bar{z} и \underline{z} , следующим образом: $z = \bar{z} - \underline{z}$, где $\bar{z} = \max(z, 0) \geq 0$; $\underline{z} = \max(-z, 0) \geq 0$.

Итак, сделаем замены:

$$x_2 = \bar{x}_2 - \underline{x}_2; \quad y_2 = \bar{y}_2 - \underline{y}_2,$$

где

$$\bar{x}_{j_2} = \max(x_{j_2}, 0), \quad j \in \mathfrak{F}_2;$$

$$\underline{x}_{j_2} = \max(-x_{j_2}, 0), \quad j \in \mathfrak{F}_2;$$

$$\bar{y}_{i_2} = \max(y_{i_2}, 0), \quad i \in I_2;$$

$$\underline{y}_{i_2} = \max(-y_{i_2}, 0), \quad i \in I_2.$$

Тогда пара двойственных задач будет иметь вид:

$$1) \min [(c_1, x_1) + (c_2, \bar{x}_2) - (c_2, \bar{x}_2)]$$

$$A_{11}x_1 + A_{12}\bar{x}_2 - A_{22}\bar{x}_2 \geq b_1$$

$$A_{21}x_1 + A_{22}\bar{x}_2 - A_{22}\bar{x}_2 \geq b_2$$

$$-A_{21}x_1 - A_{22}\bar{x}_2 + A_{22}\bar{x}_2 \geq -b_2$$

$$x_1 \geq 0, \bar{x}_2 \geq 0, \bar{x}_2 \geq 0.$$

$$2) \max [(b_1, y_1) + (b_2, \bar{y}_2) - (b_2, \bar{y}_2)]$$

$$A_{11}^T y_1 + A_{21}^T \bar{y}_2 - A_{21}^T \bar{y}_2 \leq c_1$$

$$A_{12}^T y_1 + A_{22}^T \bar{y}_2 - A_{22}^T \bar{y}_2 \leq c_2$$

$$-A_{12}^T y_1 - A_{22}^T \bar{y}_2 + A_{22}^T \bar{y}_2 \leq -c_2$$

$$y_1 \geq 0, \bar{y}_2 \geq 0, \bar{y}_2 \geq 0$$

Канонический вид задачи ЛП. Задачу

$$\min (c, x) \\ x \in R_0$$

$$R_0 = \{x: Ax = b, x \geq 0\}$$

называют задачей ЛП в каноническом виде.

Эта задача является частным случаем задачи со смешанными ограничениями ($I_1 = \emptyset, I_2 = \emptyset$).

Приведение основной задачи ЛП

$$\min (c, x)$$

$$Ax \geq b, x \geq 0$$

к каноническому виду производится путем введения дополнительных переменных:

$$\min (c, x)$$

$$Ax - u = b, x \geq 0, u \geq 0.$$

3.3. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ЛП

Симплекс-метод. Симплекс-метод представляет собой последовательный перебор крайних точек, при котором значение целевой функции убывает от итерации к итерации, т. е. от одной угловой точки к другой.

Рассмотрим задачу ЛП в каноническом виде

$$\min (c, x), R_0 = \{x: Ax = b, x \geq 0\}. \\ x \in R_0$$

Крайнюю точку множества R_0 называют **невырожденной**, если матрица ее базиса имеет размерность $m \times m$ (последнее означает, что у невырожденной угловой точки x число положительных компонент равно m).

Задачу ЛП называют **невырожденной**, если невырождена каждая угловая точка множества R_0 .

Положим, что наша задача ЛП не вырождена.

Итерационный шаг метода состоит в переходе от угловой точки x к угловой точке x' , при котором значение целевой функции убывает:

$$(c, x') < (c, x).$$

Пусть известна угловая точка x . Не ограничивая общности, можно сказать, что базис B этой угловой точки образуют первые m столбцов матрицы A , которую запишем в виде:

$$A = [BD],$$

где $B = [a_1, \dots, a_m]$; $D = [a_{m+1}, \dots, a_n]$.

Теперь запишем векторы c и x в виде:

$$x^T = (x_B^T, x_D^T);$$

$$x_B^T = (x_1, \dots, x_m);$$

$$c_B^T = (c_1, \dots, c_m);$$

$$c^T = (c_B^T, c_D^T);$$

$$x_D^T = (x_{m+1}, \dots, x_n);$$

$$c_D^T = (c_{m+1}, \dots, c_n).$$

Компоненты вектора x_B обычно называют **базисными**, а векторы x_D — **внебазисными**.

Если хотя бы одна компонента вектора x_B — нулевая, то крайняя точка x — вырожденная. Действительно, если, например, $x_m = 0$, то, вычеркивая из матрицы последнюю строку и столбец, видим, что x удовлетворяет неособой линейной системе уравнений $(m-1)$ -го порядка, а это означает, что угловая точка x — вырождена.

Таким образом, известно, что точка x угловая и

$$x_B > 0, x_D = 0, \det B \neq 0, Bx_B = b.$$

Обозначим

$$x_k = \begin{cases} x_B - \lambda B^{-1}a_k \\ 0 \\ \lambda \\ 0 \end{cases}, \quad k = \overline{m+1, n},$$

где λ — k -компонента вектора x_k , $\lambda > 0$.

Так как $x_B > 0$, $x_k \geq 0$ и

$$\begin{aligned} Ax_k &= \sum_{i=1}^m [x_i - \lambda (B^{-1}a_k)_i] a_i + \lambda a_k = \\ &= Bx_B = b, \end{aligned}$$

то $x_k \in R_0$ при малых $\lambda > 0$. Кроме того,

$$\begin{aligned} (c, x_k) &= (c_B, x_B) - \lambda (c_B, B^{-1}a_k) + \\ &+ \lambda c_k = (c, x) - \lambda [(c_B, B^{-1}a_k) - c_k]. \end{aligned}$$

Обозначим $\Delta_k = (c_B, B^{-1}a_k) - c_k$.

Ясно, что величина Δ_k определена для любого $k = \overline{1, n}$, причем при $k = \overline{1, m}$, $B^{-1}a_k = e_k$ (e_k — k -координатный вектор) и

$$\begin{aligned} \Delta_k &= (c_B, B^{-1}a_k) - c_k = \\ &= (c_B, e_k) - c_k = c_k - c_k = 0. \end{aligned}$$

Следовательно,

$$(c, x_k) = (c, x) - \lambda \Delta_k.$$

В зависимости от знаков Δ_k и $(B^{-1}a_k)$ возможны три случая:

1) если для любого $k = \overline{1, n}$ $\Delta_k \leq 0$, то точка x — оптимальная;

2) если найдется номер $k \geq m+1$, такой, что $\Delta_k > 0$ и $B^{-1}a_k \leq 0$, то множество R_0 — не ограничено и (c, x) не ограничена снизу на R_0 ;

3) если найдутся такие $k \geq m+1$ и $i \leq m$, что $\Delta_k > 0$ и $(B^{-1}a_k)_i > 0$, то в этом случае делается следующий итерационный шаг.

Обозначим

$$\tilde{I}_k = \{i \mid (B^{-1}a_k)_i > 0\}.$$

Выберем

$$\lambda = \lambda_0 = \min_{i \in \tilde{I}_k} \frac{(B^{-1}b)_i}{(B^{-1}a_k)_i} = \frac{(B^{-1}b)_s}{(B^{-1}a_k)_s}. \quad (1)$$

Очевидно, что $\lambda_0 \geq 0$, так как $(B^{-1}b)_i = x_i > 0$, $i = \overline{1, m}$. При таком выборе величины λ_0 точка x_k становится допустимой точкой множества R_0 . Ясно, что номер s , на котором достигается минимум в (1), единственен из-за невырожденности задачи ЛП.

Теперь надо проверить, что точка x_k — крайняя. Для этого достаточно показать, что система векторов $a_1, \dots, a_{s-1}, a_{s+1}, \dots, a_m, a_k$ линейно независима. Тогда

$$(c, x_k) = (c, x) - \lambda \Delta_k < (c, x).$$

Таким образом, итерационный шаг симплекс-метода состоит в таком переходе от базиса $a_1, \dots, a_{s-1}, a_s, a_{s+1}, \dots, a_m$ к базису $a_1, \dots, a_{s-1}, a_{s+1}, \dots, a_m, a_k$.

Замечание. Установим связь между параметрами последовательных итераций.

Обозначим:

$$\begin{aligned} a_0 &\stackrel{\Delta}{=} b, \quad z_{ik} \stackrel{\Delta}{=} (B^{-1}a_k)_i, \quad i = \overline{1, m}, \\ k &= \overline{0, n}, \\ z_{0j} &\stackrel{\Delta}{=} \Delta_j, \quad j = \overline{1, n}, \quad z_{00} \stackrel{\Delta}{=} (c, x). \end{aligned}$$

Параметрами, соответствующими угловой точке x , являются числа z_{ij} ($i = \overline{0, m}, j = \overline{0, n}$), а параметрами новой угловой точки $v = x_k$ будут числа v_{ij} . Связь между этими параметрами устанавливается формулами:

$$\begin{aligned} v_{ij} &= z_{ij} - z_{ik} z_{sj} / z_{sk} \quad (i = \overline{0, m}, \\ &i \neq s, j = \overline{0, n}); \\ v_{kj} &= z_{sj} / z_{sk} \quad (j = \overline{0, n}). \end{aligned} \quad (2)$$

Отыскание исходной крайней точки.

Метод искусственного базиса. Рассмотрим задачу отыскания угловой точки множества

$$R_0 = \{x \mid Ax = b, x \geq 0\}.$$

Без ограничения общности можно положить $b \geq 0$.

Рассмотрим вспомогательную задачу в пространстве E_{n+m} :

$$\left\{ \begin{array}{l} \min_{x, u \in W} \sum_{i=1}^m u_i \\ W = \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \mid Ax + u = b, x \geq 0, u \geq 0 \right\} \end{array} \right\} \quad (3)$$

с заранее известной крайней точкой $\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$ множества W . Применяя симплекс-метод, находят $\begin{pmatrix} x^* \\ u^* \end{pmatrix}$ решение задачи (3).

Обозначим $\mu \triangleq \sum_{i=1}^m u_i^*$

Теорема. Если $\mu = 0$, то x^* — угловая точка множества R_0 . Если $\mu > 0$, то $R_0 = \emptyset$.

M-метод. Рассмотрим задачу (M):

$$\min_{x, u \in W} \left[(c, x) + M \sum_{i=1}^m u_i \right]$$

$$W \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \mid Ax + u = b, x \geq 0, u \geq 0 \right\}.$$

Теорема. Если задача (M) решима, то найдется такое число M_0 , что для всех $M \geq M_0$ в любом решении $\begin{pmatrix} x^* \\ u^* \end{pmatrix}$ точка x^* будет оптимальной для задачи ЛП.

Замечания: 1. Очевидно, что при $b \geq 0$ точка $\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$ является угловой точкой множества W , поэтому ее можно выбирать в качестве исходной для алгоритма симплекс-метода.

2. На практике нет необходимости определять число M_0 . В качестве M можно выбрать

$$M > \max_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \{ |a_{ij}|, |c_i|, |b_j| \}.$$

Модифицированный симплекс-метод. В обычном симплекс-методе на каждой итерации все элементы матрицы A преобразуются по формулам (2). В модифицированном симплекс-методе на каждой итерации вычисляют матрицу

B^{-1} и затем, пользуясь ею, вычисляют

$$\bar{y} = (B^{-1})^T c_B, \quad \Delta_k = (\bar{y}_1 a_k) - c_k, \quad B^{-1}b$$

и т. д.

Затем с помощью рекуррентных формул переходят от матрицы

$$B^{-1} = [a_1, \dots, a_m]^{-1}$$

к матрице

$$\bar{B}^{-1} = [a_1, \dots, a_{s-1}, a_s, a_{s+1}, \dots, a_m]^{-1},$$

соответствующей новой угловой точке.

Приведем формулы перехода от B^{-1} к \bar{B}^{-1} .

Если $B^{-1} = (\hat{b}_{ij}^{-1})$, а $\bar{B}^{-1} = (\bar{b}_{ij}^{-1})$, то

$$\bar{b}_{ij}^{-1} = \hat{b}_{ij}^{-1} - z_{ik} \frac{\hat{b}_{sj}^{-1}}{z_{sk}}, \quad i \neq s;$$

$$\bar{b}_s^{-1} = \frac{\hat{b}_{sj}^{-1}}{z_{sk}}.$$

4. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

4.1. СУЩНОСТЬ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Постановка задачи. Пусть E_n — n -мерное евклидово пространство и $x \in E_n$, $x^T = (x_1, \dots, x_n)$, где x_i , $i = 1, n$, — координаты. Предположим, что в пространстве E_n задано множество $G = G_1 \times G_2 \times \dots \times G_n$, где G_i , $i = 1, n$, — множества на числовой оси, и скалярная функция $f(x)$; $x \in G$ вида

$$f(x) = \sum_{i=1}^n f_i(x_i),$$

где $f_i(x_i)$, $i = 1, n$ — скалярные функции, определенные на G_i .

Задача 1. Среди векторов $x \in G$ найти такой x^* , на котором функция $f(x)$ достигает минимума: $f(x^*) = \min_{x \in G} f(x)$.

Решение. Шаг 1. Рассмотрим семейство задач

$$\min_{\substack{x_i \in G_i \\ i=\overline{1, s}}} \sum_{i=1}^s f_i(x_i), \quad s = \overline{1, n}.$$

Очевидно, что среди этих задач находится и задача 1, которая получается при $s = n$, т. е. задача 1 погружена в семейство задач.

Введем функцию Беллмана

$$B_s \triangleq \min_{\substack{x_i \in G_i \\ i=\overline{1, s}}} \sum_{i=1}^s f_i(x_i).$$

Шаг 2. Составим уравнение, которому удовлетворяет функция Беллмана.

В данном случае

$$\min_{\substack{\alpha \in A \\ \beta \in B}} [f_1(\alpha) + f_2(\beta)] = \min_{\alpha \in A} f_1(\alpha) +$$

$$+ \min_{\beta \in B} f_2(\beta); \quad (A \cap B = \emptyset).$$

Из определения функции Беллмана и сделанного замечания имеем

$$\begin{aligned} B_{s+1} &= \min_{\substack{x_i \in G_i \\ i=\overline{1, s+1}}} \sum_{i=1}^{s+1} f_i(x_i) = \\ &= \min_{\substack{x_i \in G_i \\ i=\overline{1, s+1}}} \left[f_{s+1}(x_{s+1}) + \sum_{i=1}^s f_i(x_i) \right] = \\ &= \min_{\alpha \in G_{s+1}} \left[f_{s+1}(\alpha) + \right. \\ &\left. + \min_{\substack{x_i \in G_i \\ i=\overline{1, s}}} \sum_{i=1}^s f_i(x_i) \right] = \\ &= \min_{\alpha \in G_{s+1}} f_{s+1}(\alpha) + B_s. \end{aligned}$$

Таким образом, мы получили

$$B_{s+1} = B_s + \min_{\alpha \in G_{s+1}} f_{s+1}(\alpha),$$

$$s = \overline{1, n-1}.$$

Это уравнение (рекуррентно-функциональное) носит название прямого уравнения Беллмана.

Для решения этого уравнения надо задать начальное условие

$$B_1 = \min_{\alpha \in G_1} f_1(\alpha)$$

(последнее следует из определения функции Беллмана).

Замечания: 1. Задача 1 нами сведена к задачам минимизации функций $f_i(x_i)$.

2. Функцию Беллмана можно определить в данном случае так:

$$\bar{B}_s = \min_{\substack{x_i \in G_i \\ i=\overline{s, n}}} \sum_{i=s}^n f_i(x_i).$$

Тогда, аналогично решению задачи 1, можно установить, что \bar{B}_s удовлетворяет уравнению

$$\bar{B}_s = \bar{B}_{s+1} + \min_{\alpha \in G_s} f_s(\alpha),$$

$$\bar{B}_s|_{s=n} = \min_{\alpha \in G_n} f_n(\alpha),$$

которое обычно называют обратным уравнением Беллмана или просто уравнением Беллмана.

Задача минимизации с дополнительными ограничениями.

Пусть кроме функций $f_i(x_i)$ и множества G , определенных в задаче 1, задана функция $g(x) = \sum_{i=1}^n g_i(x_i)$, $x \in G$.

Задача 2. Среди элементов $x \in G$, удовлетворяющих условию $g(x) \leq a$, $a \in R^1$, найти такое x^* , на котором $f(x)$ достигает минимума:

$$f(x^*) = \min_{\substack{x \in G \\ g(x) \leq a}} f(x).$$

Решение. Задача 2 вложена в семейство задач

$$\min_{x \in G_i} \sum_{i=1}^n f_i(x_i).$$

$$\sum_{i=s}^n g_i(x_i) \leq y$$

$$s = \overline{n, 1}$$

Функцию Беллмана определим равенством

$$B_s(y) \triangleq \min_{x_i \in G_i} \sum_{i=s}^n f_i(x_i).$$

$$\sum_{i=s}^n g_i(x_i) \leq y$$

Заметим, что

$$\min_{\alpha \in A} \min_{\beta \in B(\alpha)} [f_1(\alpha) + f_2(\alpha, \beta)] =$$

$$= \min_{\alpha \in A} \left[f_1(\alpha) + \min_{\beta \in B(\alpha)} f_2(\alpha, \beta) \right]$$

Из определения функции Беллмана и сделанного замечания получим уравнение Беллмана

$$B_s(y) = \min_{\alpha \in G_s} [f_1(\alpha) + B_{s+1}(y - g_s(\alpha))]$$

$$B_s(y) |_{s=n} = \min_{\alpha \in G_n} f_n(\alpha).$$

$$g_n(\alpha) \leq y$$

Задачи 3. Пусть минимизируемая функция определена на множестве G (см. задачу 1) и имеет вид

$$f(x) = f(x_1, \dots, x_n) =$$

$$= \sum_{i=1}^{n-k} f_i(x_i, \dots, x_{i+k}), \quad k = \overline{0, n-1}.$$

Требуется найти такой вектор $x^* \in G$, что

$$f(x^*) = \min_{x \in G} f(x).$$

Решение. Рассмотрим семейство задач

$$F^s(y_1, \dots, y_k, x_{k+s}, \dots, x_n) =$$

$$= \sum_{i=s}^{n-k} f_i(x_i, \dots, x_{i+k}) \rightarrow \min_{x_i \in G_i}$$

$$i = \overline{k+s, n}$$

$$s = \overline{1, n-k}$$

где переменные x_s, \dots, x_{s+k-1} считаются фиксированными: $x_{i+s-1} = y_i, i = \overline{1, k}$.

Введем функцию Беллмана

$$B_s(y_1, \dots, y_k) =$$

$$= \min_{x_i \in G_i} F^s(y_1, \dots, y_k, x_{k+s}, \dots, x_n).$$

$$i = \overline{k+s, n}$$

Аналогично предыдущему, несложно получить уравнение Беллмана

$$B_s(y_1, \dots, y_k) =$$

$$= \min_{\alpha \in G_{s+k}} [f_s(y_1, \dots, y_k, \alpha) +$$

$$+ B_{s+1}(y_2, \dots, y_k, \alpha)];$$

$$B_{n-k}(y_1, \dots, y_k) = \min_{\alpha \in G_n} f_{n-k}(y_1, \dots,$$

$$y_k, \alpha).$$

4.2. ОПТИМИЗАЦИЯ ДИСКРЕТНЫХ ПРОЦЕССОВ

Основные определения. Постановка задачи. Пусть имеется некоторый объект, состояние которого меняется в дискретные моменты времени $t = t_0, t_0 + 1, t_0 + 2, \dots$, образуя некоторый процесс. Предположим, что в момент t состояние объекта полностью описывается n -мерным вектором $x^T(t) = (x_1(t), \dots, x_n(t))^T$. Будем считать, что процесс $\{x(t)\}_{t \geq t_0}$ определяется как собственной динамикой объекта, так и внешними воздействиями $u^T(t) = (u_1(t), \dots, u_2(t))^T$ с помощью рекуррентного соотношения

$$x(t+1) = f(x(t), u(t), t),$$

$$t \geq t_0, x(t_0) = x_0, \quad (4)$$

где t_0 — начальный момент времени; x_0 — начальное состояние объекта; $f(x, u, t)$ — n -мерная функция, опре-

деленная в области значений своих аргументов.

Функцию $u(t)$, $t \geq t_0$ будем называть управлением. Выбирая то или иное $u(t)$, получим согласно (4) соответствующее течение процесса $\{x(t)\}_{t > t_0}$.

Управление $u(t)$ называют допустимым, если в каждый момент времени $t \geq t_0$ оно принимает значение из заданного подмножества $U(t)$ r -мерного евклидова пространства:

$$u(t) \in U(t).$$

Процесс $\{x(t)\}_{t \geq t_0}$, соответствующий допустимому управлению, называют допустимым процессом.

Рассмотрим процесс $x(t)$, $t \in [t_0, t_1]$. Качество процесса $\{x(t)\}_{t > t_0}$ оценим, например, величиной

$$I(u) = \sum_{t=t_0}^{t_1} f_0(x(t), u(t), t),$$

которую называют критерием качества.

Простейшая задача оптимизации многошагового процесса $\{x(t)\}_{t > t_0}$ состоит в минимизации $I(u)$ критерия качества по всем допустимым управлениям, т. е.

$$I(u^0) = \min_{\substack{u(t) \in U(t) \\ t \in [t_0, t_1]}} I(u).$$

Траекторию процесса $\{x^0(t)\}_{t > t_0}$, соответствующую оптимальному управлению $\{u^0(t)\}_{t > t_0}$, называют оптимальной.

Принцип оптимальности Беллмана. Предположим, что управляемый процесс $x(t)$ удовлетворяет (4) $t \in [t_0, t_1 - 1]$. Основной особенностью задания процесса $\{x(t)\}$ является то, что для любого вектора y , числа $\tau \in [t_0, t_1 - 1]$ выбор начального условия $x(\tau) = y$ и функции $u(t)$, $t \in [\tau, t_1 - 1]$ однозначно определяет течение процесса $x(t)$ на промежутке $[\tau, t_1]$. Другая особенность состоит в том, что при любом $\tau \in [t_0 + 1, t_1 - 1]$ состояние $x(t)$, $t \in [t_0, \tau_1]$ не зависит от выбора управления на $[\tau + 1, t_1 - 1]$.

При любом $\tau \geq t_0$ рассматриваемый процесс $x(t)$, $t \geq \tau$ однозначно определяется состоянием $x(\tau)$ и управле-

нием $u(t)$, $t \geq \tau$, поэтому при оценке для этого процесса значения критерия качества будет функцией этих параметров: $I(x(\tau), u[\tau, t_1])$, где через $u[\tau, t_1 - 1]$ обозначено множество $\{u(\tau), \dots, u(t_1)\}$.

Пусть управление $u(t)$ процессом $x(t)$ определено на отрезке $[t_0, t_1]$. Разделим промежуток $[t_0, t_1]$ на два участка: первый — $[t_0, \tau - 1]$, второй — $[\tau, t_1]$, где $\tau \in [t_0 + 1, t_1 - 1]$.

Управление u на первом участке $[t_0, \tau - 1]$ приводит процесс в состояние $x(\tau)$, которое является начальным для управления u на втором участке $[\tau, t_1]$. Пусть критерий качества определен на любом участке $[\tau, t_1]$. Тогда качество управления на первом участке зависит от $(x(t_0))$ и управления $u[t_0, \tau - 1]$, а на втором — оценивается величиной $I(x(\tau), u[\tau, t_1])$.

Допустим, что между вычисленными значениями критерия существует связь

$$I(x(t_0), u[t_0, t_1]) = v(x(t_0), u[t_0, \tau - 1]), I(x(\tau), u[\tau, t_1]), \quad (5)$$

где $v(\alpha, \beta, \gamma)$ — функция, такая, что при любых $\alpha, \beta, \gamma, \bar{\gamma}$ $v(\alpha, \beta, \bar{\gamma}) < v(\alpha, \beta, \gamma)$ при $\bar{\gamma} < \gamma$.

Будем считать, что выполнены условия на управления и коэффициенты (5), обеспечивающие существование допустимых управлений и процесса.

Принцип Беллмана. 1. (Необходимое условие). Для оптимальности управления необходимо, чтобы при любой длительности участков управление на втором участке было оптимально относительно состояния, в котором оказался процесс после управления на первом участке.

2. (Достаточное условие). Если управление на втором участке оптимально относительно состояния, возникшего в результате управления на первом участке, причем управление на первом участке оптимально, то управление на всем промежутке будет оптимальным.

Уравнение Беллмана. Введем функцию Беллмана:

$$B(\tau, y) \stackrel{\Delta}{=} \min I(\tau, y, u[\tau, t_1]),$$

где минимум берется по всем допустимым $u[\tau, t_1]$.

Функция Беллмана определена при всех $\tau \in [t_0, t_1]$ и y .

Предположим, что выполнены приведенные выше условия. Тогда $B(\tau, y)$ удовлетворяет уравнению при $\tau \leq t_1$

$$B(\tau, y) = \min_u v(\tau, y, u,$$

$$B(\tau + 1, f(\tau, y, u)),$$

$$B(t_1, y) = \min_u I(t_1, y, u[t_1]),$$

причем минимум берется по всем u из множества допустимых управлений.

Пример. Линейные системы с квадратичным критерием качества.

Рассмотрим задачу

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)u(t), \\ x(t_0) &= x_0, t \in [t_0, t_1 - 1] \end{aligned}$$

$$I(u) = \sum_{t=t_0}^{t_1-1} [x^T(t)M(t)x(t) +$$

$$+ u^T(t)R(t)u(t)] \rightarrow \min,$$

где $M(t)$, $R(t)$ — симметричные положительно определенные матрицы-функции; $A(t)$, $B(t)$ — матрицы-функции размера соответственно $n \times n$ и $n \times r$; $x(t)$, $u(t)$ — n -мерные и r -мерные вектор-функции. Ограничений на управление u не накладывается.

Уравнение Беллмана в данном случае имеет вид

$$\begin{cases} B(\tau, y) = \min_v [y^T M(\tau) y + v^T R(\tau) v + \\ + B(\tau + 1, A(\tau) y + B(\tau) v)]; \\ B(t_1 - 1, y) = y^T M(t_1 - 1) y. \end{cases}$$

Решение рассматриваемого уравнения ищем в виде

$$B(\tau, y) = y^T L(\tau) y,$$

где $L(\tau)$ — некоторая неизвестная матрица.

Из граничного условия ясно, что $L(t_1 - 1) = M(t_1 - 1)$. Несложно убедиться в том, что $L(\tau)$ удовлетворяет уравнению:

$$\begin{aligned} L(\tau) &= M(\tau) + \{ [R(\tau) + \\ &+ B^T(\tau) L(\tau + 1) B(\tau)]^{-1} B^T(\tau) \times \\ &\times L(\tau + 1) A(\tau) \}^T R(\tau) \{ [R(\tau) + \\ &+ B^T(\tau) L(\tau + 1) B(\tau)]^{-1} B^T(\tau) \times \\ &\times L(\tau + 1) A(\tau) \} + \{ A(\tau) - \\ &- B(\tau) [R(\tau) + B^T(\tau) L(\tau + 1) \times \\ &\times B(\tau)]^{-1} B^T(\tau) L(\tau + 1) A(\tau) \}^T \times \\ &\times L(\tau + 1) \{ A(\tau) - B(\tau) [R(\tau) + \\ &+ B^T(\tau) L(\tau + 1) B(\tau)]^{-1} B^T(\tau) \times \\ &\times L(\tau + 1) A(\tau) \}, \end{aligned}$$

а оптимальное управление имеет вид

$$u(y, \tau) = - [R(\tau) + B^T(\tau) L(\tau + 1) \times \\ \times B(\tau)]^{-1} B^T(\tau) L(\tau + 1) A(\tau) y.$$

СПИСОК ЛИТЕРАТУРЫ

1. Болтянский В. Г. Оптимальное управление дискретными системами. М.: Наука, 1973. 446 с.
2. Васильев Ф. П. Лекции по методам решения экстремальных задач. М.: Изд-во МГУ, 1974. 374 с.
3. Гольдштейн Е. Г. Теория двойственности в математическом программировании. М.: Наука, 1971. 351 с.
4. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. М.: Наука, 1974. 479 с.
5. Карманов В. Г. Математическое программирование. М.: Наука, 1975. 272 с.
6. Рокафеллар Р. Г. Выпуклый анализ. М.: Мир, 1973. 469 с.

Часть II. МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ И ЭФФЕКТИВНОСТИ

Глава 9. Надежность элемента

1. НАДЕЖНОСТЬ НЕВОССТАНАВЛИВАЕМОГО ЭЛЕМЕНТА

Будем считать, что время безотказной работы элемента ξ есть случайная величина с функцией распределения $F(t) = P\{\xi < t\}$, которая предполагается абсолютно непрерывной, т. е. существует плотность $f(t) = F'(t)$ и

$$F(t) = \int_0^t f'(x) dx. \text{ Дополнительную}$$

вероятность $P\{\xi \geq t\} = 1 - \bar{F}(t) = 1 - F(t)$ будем называть вероятностью безотказной работы. Обозначаем через ξ_t остаточное время жизни элемента, т. е. ξ_t есть случайная величина $(\xi - t)$ при условии, что $\xi \geq t$. Эта величина имеет распределение

$$P\{\xi_t \geq x\} = \frac{\bar{F}(x+t)}{\bar{F}(t)}$$

В частности, вероятность отказа элемента на интервале $(t, t + \Delta t)$ при условии, что он дожил до момента t

$$\begin{aligned} P\{\xi_t < \Delta t\} &= \frac{F(t + \Delta t) - F(t)}{\bar{F}(t)} = \\ &= \frac{f(t)}{\bar{F}(t)} \Delta t + o(\Delta t). \end{aligned}$$

Функцию $\lambda(t) = \frac{f(t)}{\bar{F}(t)}$ называют интенсивностью отказа элемента. Ее можно трактовать как вероятность того, что элемент, доживший до момента t , откажет за последующую (малую) единицу времени. Нетрудно найти обратное выражение

$$\bar{F}(t) = e^{-\int_0^t \lambda(x) dx} \quad (1)$$

Из него следует, что

$$P(\xi_t \geq x) = e^{-\int_t^{t+x} \lambda(u) du} \quad (2)$$

Таким образом, интенсивность отказа $\lambda(t)$ является локальной характеристикой надежности в том смысле, что вероятность отказа элемента, дожившего до момента t , на очередном интервале $(t, t + x)$ зависит только от значений функции $\lambda(u)$ на этом интервале и не зависит от ее поведения вне этого интервала. Кроме функциональных характеристик $F(t)$, $f(t)$, $\bar{F}(t)$, $\lambda(t)$, надежность элемента часто характеризуют числовыми величинами. Наиболее важными являются сред-

нее время безотказной работы элемента и его дисперсия:

$$T = M\xi = \int_0^{\infty} t dF(t) = \int_0^{\infty} \bar{F}(t) dt; \quad (3)$$

$$D\xi = \sigma^2 = \int_0^{\infty} (t - T)^2 dF(t).$$

2. РАСПРЕДЕЛЕНИЯ, ИСПОЛЬЗУЕМЫЕ В ТЕОРИИ НАДЕЖНОСТИ

Экспоненциальный (показательный) закон. Так называют распределение $\bar{F}(t) = e^{-\lambda t}$, для которого

$$\lambda(t) = \lambda; \quad T = M\xi = \frac{1}{\lambda};$$

$$D\xi = \sigma^2 = \frac{1}{\lambda^2}. \quad (4)$$

Для показательного распределения остаточное время жизни ξ_t имеет распределение

$$P\{\xi_t \geq x\} = \frac{\exp[-\lambda(t+x)]}{\exp[-\lambda t]} = \exp(-\lambda x), \quad (5)$$

независящее от момента t . Это свойство показательного распределения является характеристическим свойством в том смысле, что из условия независимости ξ_t от t следует, что распределение времени жизни будет показательным. Таким образом, показательный закон можно использовать в тех случаях, когда элемент не стареет, т. е. его остаточное время жизни не зависит от того, сколько времени он проработал до данного момента.

Нормальный закон. Строго говоря, в теории надежности используют усеченный нормальный закон с плотностью

$$f(t) = \frac{c}{\sqrt{2\pi}\sigma} e^{-\frac{(t-T)^2}{2\sigma^2}}, \quad (6)$$

где постоянная c определяется из условия $\int_0^{\infty} f(t) dt = 1$. Обычно вели-

чина σ в несколько раз меньше T . В этом случае $c \approx 1$, а T и σ^2 равны соответственно среднему и дисперсии времени жизни. Нормальный закон часто используют в тех случаях, когда отказы носят постепенный характер, являются следствием направленных физико-химических изменений в элементе.

Закон Вейбулла—Гнеденко. В тех случаях, когда плотность распределения отказа имеет несимметричный вид, часто используют закон Вейбулла—Гнеденко

$$\bar{F}(t) = e^{-(\lambda t)^p}, \quad p > 0, \lambda > 0. \quad (7)$$

Для него интенсивность отказа $\lambda(t) = p\lambda^p t^{p-1}$, а среднее $T = \frac{1}{\lambda} \Gamma\left(1 + \frac{1}{p}\right)$. Так как показательное распределение входит в это семейство распределений при $p = 1$, ясно, что любое реальное распределение приближается распределением вида (7) лучше, чем показательным.

Гамма-распределение. Так называют распределение с плотностью

$$f(t) = \frac{\lambda^p t^{p-1}}{\Gamma(p)} e^{-\lambda t}, \quad p > 0, \lambda > 0. \quad (8)$$

Среднее и дисперсия имеют вид

$$T = \frac{p}{\lambda}; \quad D\xi = \sigma^2 = \frac{p}{\lambda^2}.$$

При $p = 1$ получаем показательное распределение. Поскольку преобразование Лапласа этого распределения имеет простой вид

$$\int_0^{\infty} e^{-zt} f(t) dt = \frac{\lambda^p}{(\lambda + z)^p}, \quad (9)$$

гамма-распределением удобно аппроксимировать реальное распределение в тех задачах, которые решаются в терминах преобразования Лапласа.

Смесь распределений. Предположим, что один и тот же элемент изготавливается на различных заводах, на которых разная технология, сырье и культура производства. В этом случае надежность элемента, изготовленного на разных заводах, будет разная. Пусть k -й завод изготавливает элемент с вероятностью безотказной работы $\bar{F}_k(t)$

и элементы, изготовленные на всех этих заводах, поступают на склад, где перемешиваются. Тогда наугад взятый со склада элемент будет иметь надежность

$$\bar{F}(t) = \sum_{k=1}^n \alpha_k \bar{F}_k(t), \quad (10)$$

где α_k — доля продукции, изготавливаемой на k -м заводе.

Распределение $F(t)$ называют смесью распределений $F_k(t)$. Если, в частности, $\bar{F}_k(t) = e^{-\lambda_k t}$ то

$$\bar{F}(t) = \sum_{k=1}^n \alpha_k e^{-\lambda_k t} \quad (11)$$

Можно показать, что для такого закона надежности интенсивность отказа $\lambda(t)$ монотонно убывает, т. е. с течением времени надежность элемента растет.

Стареющие распределения. Элемент называют стареющим, если его интенсивность отказа $\lambda(t)$ монотонно возрастает. В этом случае говорят также, что распределение $F(t)$ стареющее и время жизни — стареющее. Для характеристик надежности стареющего элемента можно получить большое число весьма полезных неравенств. Приведем некоторые из них:

а) если известно среднее время безотказной работы $M\xi = T$, то

$$\bar{F}(t) \geq e^{-\frac{t}{T}} \quad (12)$$

для всех $t \leq T$, что позволяет оценивать надежность снизу;

б) если мы, проведя испытания элементов в течение большого промежутка времени t_2 , нашли нижнюю оценку надежности $\bar{F}(t_2) \geq p$, то при любом $t_1 < t_2$

$$\bar{F}(t_1) \geq [\bar{F}(t_2)]^{\frac{t_1}{t_2}} \geq p^{\frac{t_1}{t_2}}, \quad (13)$$

что позволяет весьма точно оценивать надежность на начальном участке времени;

в) для стареющих распределений справедливо неравенство

$$D\xi = \sigma^2 \leq (M\xi)^2 = T^2, \quad (14)$$

причем равенство может быть только в том случае, когда распределение $F(t)$ показательное. Этот факт можно использовать для статистической проверки гипотезы о показательном распределении времени жизни, так как в классе стареющих распределений условие $\sigma = T$ необходимо и достаточно для того, чтобы распределение $F(t)$ было показательным.

Ниже приведены и другие неравенства для характеристик надежности, верные в классе стареющих распределений. Существуют и другие, более широкие классы стареющих распределений, которые удобно использовать при оценке надежности.

3. НАДЕЖНОСТЬ ВОССТАНАВЛИВАЕМОГО ЭЛЕМЕНТА

Случай мгновенного восстановления. Рассмотрим теперь элемент, который после каждого отказа восстанавливается и снова начинает работать. Мы будем различать два типа восстановления — замену и ремонт. Однако в том и другом случае предполагаем, что восстановление полное, т. е. после восстановления элемент имеет такую же надежность, что и в начальный момент. Рассмотрим сначала случай, когда время восстановления мало и им можно пренебречь. Пусть $0 < t_1 < t_2 < \dots < t_n \dots$ — последовательные моменты отказов (и восстановлений) элемента, $\xi_1 = t_1$, $\xi_2 = t_2 - t_1$, ..., $\xi_n = t_n - t_{n-1}$, — времена безотказной работы элемента до первого отказа, после первого восстановления, второго восстановления и т. д. Согласно сказанному выше предполагаем, что все величины ξ_k независимы и имеют одинаковое распределение

$$P\{\xi_k < t\} = F(t), \quad M\xi_k = T,$$

$$D\xi_k = \sigma^2.$$

Последовательность случайных моментов t_1, t_2, \dots, t_n называют процессом восстановления, а раздел теории вероятностей, который занимается изучением этого процесса, называют теорией восстановления. В нашем случае характеристики процесса восстановления будут характеристиками на-

дежности восстанавливаемого элемента. Перечислим основные из этих характеристик.

1. Число отказов до момента t $v(t)$ имеет распределение

$$P\{v(t) = n\} = F_n(t) - F_{n+1}(t), \quad (15)$$

где $F_n(t) = P\{t_n < t\}$ определяется рекуррентно

$$F_n(t) = \int_0^t F_{n-1}(t-x) dF(x),$$

$$F_1(t) = F(t).$$

2. Функцией восстановления $H(t)$ называют среднее число отказов до момента t

$$H(t) = Mv(t) = \sum_{k=1}^{\infty} F_k(t). \quad (16)$$

Отсюда среднее число отказов на интервале $(t, t+x)$ равно $H(t+x) - H(t)$.

3. Плотностью восстановления называют производную

$$h(t) = H'(t) = \sum_{k=1}^{\infty} f_k(t). \quad (17)$$

В теории надежности эту характеристику называют обычно интенсивностью отказов. Она имеет двойной смысл. С одной стороны, среднее число отказов на бесконечно малом интервале $(t, t + \Delta t)$ равно $h(t) \Delta t + o(\Delta t)$, т. е. $h(t)$ есть среднее число отказов за малую единицу времени, следующую за моментом t . С другой стороны, можно показать, что вероятность отказа элемента на интервале $(t, t + \Delta t)$ тоже равна $h(t) \Delta t + o(\Delta t)$, и, следовательно, $h(t)$ есть вероятность отказа за малую единицу времени. Не следует смешивать интенсивность отказа $\lambda(t)$ и интенсивность отказов $h(t)$.

4. Остаточным временем жизни ζ_t называют интервал от момента t до ближайшего справа отказа. Нетрудно найти точное распределение величины ζ_t :

$$P\{\zeta_t > x\} =$$

$$= \bar{F}(t+x) + \int_0^t \bar{F}(t+x-u) h(u) du. \quad (18)$$

Для случая, когда распределение $F(t)$ показательное, т. е. $\bar{F}(t) = e^{-\lambda t}$, процесс восстановления будет пуассоновским процессом и для него все определенные выше характеристики будут иметь более простой вид

$$P\{v(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t};$$

$$H(t) = \lambda t; \quad h(t) = \lambda;$$

$$P\{\zeta_t > x\} = e^{-\lambda x}. \quad (19)$$

Асимптотическое поведение процесса восстановления. Для многих реальных систем элемент функционирует в течение времени, которое во много раз больше среднего времени безотказной работы элемента. Естественно поэтому изучить поведение процесса восстановления при $t \rightarrow \infty$ и воспользоваться полученными асимптотическими результатами для оценки характеристик надежности восстанавливаемого элемента для большого промежутка времени t . Приведем некоторые результаты из теории восстановления.

1. Теорема Блэкуэлла.

Если распределение F нерешетчатое, то при любом $x > 0$

$$H(t+x) - H(t) \rightarrow \frac{x}{T}. \quad (20)$$

Распределение называют решетчатым, если случайная величина может принимать только значения вида an , $n = 0, 1, 2,$

2. Узловая теорема Смита.

Если распределение F нерешетчатое, а функция $Q(x)$ монотонна и интегрируема на $(0, \infty)$, то

$$\lim_{t \rightarrow \infty} \int_0^t Q(t-x) dH(x) = \frac{1}{T} \int_0^{\infty} Q(x) dx. \quad (21)$$

3. Если распределение F нерешетчатое, то

$$\lim_{t \rightarrow \infty} \left[H(t) - \frac{t}{T} \right] = \frac{\sigma^2 - T^2}{2T^2}. \quad (22)$$

4. В тех же предположениях

$$\lim_{t \rightarrow \infty} P \left\{ \frac{v(t) - \frac{t}{T}}{\sqrt{\frac{\sigma^2 t}{T^3}}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \Phi(x). \quad (23)$$

Все эти предложения можно использовать для оценки характеристик надежности восстанавливаемого элемента. Если элемент функционирует достаточно большое время t , то, как следует из формулы (20), среднее число отказов на интервале $(t, t+x)$ не зависит от t и равно x/T . Среднее число отказов на интервале $(0, t)$ при большом t приближенно равно, как следует из выражения (22),

$$H(t) \approx \frac{t}{T} + \frac{\sigma^2 - T^2}{2T^2}.$$

Утверждение (23) можно использовать для расчета ЗИПа. Если элемент, у которого восстановление состоит в замене, функционирует время t , то $v(t)$ есть число запасных элементов, необходимых для непрерывной работы элемента до момента t . Как следует из формулы (23), величина $v(t)$ при большом t имеет приближенно нормальное распределение. Число запасных элементов n_0 выбираем из условия $P\{v(t) \leq n_0\} = 1 - \varepsilon$, где ε — малое число. Используя нормальную аппроксимацию, получаем

$$n_0 = \frac{t}{T} + x_{1-\varepsilon} \sqrt{\frac{\sigma^2 t}{T^3}}, \quad (24)$$

где квантиль $x_{1-\varepsilon}$ берется по таблицам из условия

$$\Phi(x_{1-\varepsilon}) = 1 - \varepsilon.$$

Узловую теорему Смита (21) удобно использовать для доказательства различных предельных теорем в теории

восстановления. В частности, из нее и из формулы (18) легко получается распределение стационарного остаточного времени

$$\lim_{x \rightarrow \infty} P\{\zeta_t > x\} = P\{\zeta > x\} = \frac{1}{T} \int_x^\infty \bar{F}(u) du. \quad (25)$$

Отсюда среднее остаточное время

$$M\zeta = \frac{T}{2} + \frac{\sigma^2}{2T}. \quad (26)$$

Восстанавливаемый элемент с конечным временем восстановления. Предположим теперь, что время восстановления элемента конечно, и обозначим через $\xi_1, \xi_2, \xi_3, \dots$ последовательные интервалы безотказной работы элемента, а через $\eta_1, \eta_2, \eta_3, \dots$ последовательные участки восстановления. Как и выше, предполагаем, что все величины ξ_i и η_j независимы в совокупности

$$P\{\xi_i < t\} = F(t), \quad M\xi_i = T_1,$$

$$D\xi_i = \sigma_1^2;$$

$$P\{\eta_j < t\} = G(t), \quad M\eta_j = T_2,$$

$$D\eta_j = \sigma_2^2.$$

В этом случае моменты отказов и моменты восстановлений уже не совпадают. Обозначим $v_1(t)$ число отказов до момента t , $v_2(t)$ — число восстановлений до момента t ,

$$H_1(t) = Mv_1(t); \quad H_2(t) = Mv_2(t).$$

Для этих величин можно выписать формулы, аналогичные формулам (15) — (19). Моменты восстановлений образуют процесс восстановления, у которого интервал между соседними точками имеет функцию распределения

$$\Phi(t) = \int_0^t F(t-x) dG(x), \quad \text{а моменты}$$

отказов образуют расширенный процесс восстановления, у которого первый интервал имеет функцию распределения $F(t)$, а остальные $\Phi(t)$. Поэтому утверждения (20) и (21) остаются справедливыми, если вместо $H(t)$ подставить $H_1(t)$ и $H_2(t)$, а вместо T —

сумму $T_1 + T_2$. Утверждение (22) для функций $H_1(t)$ и $H_2(t)$ имеет вид

$$\lim_{t \rightarrow \infty} \left[H_2(t) - \frac{t}{T_1 + T_2} \right] = \frac{\sigma_1^2 + \sigma_2^2 - (T_1 + T_2)^2}{2(T_1 + T_2)^2}; \quad (27)$$

$$\lim_{t \rightarrow \infty} \left[H_1(t) - \frac{t}{T_1 + T_2} \right] = \frac{\sigma_1^2 + \sigma_2^2 - (T_1 + T_2)^2}{2(T_1 + T_2)^2} + \frac{T_2}{T_1 + T_2}. \quad (28)$$

Остаточное время ξ_t определяется здесь несколько иначе: $\xi_t = 0$, если момент t попал на участок восстановления; в противном случае ξ_t есть время до первого после момента t отказа. Можно найти точное распределение остаточного времени и, используя теорему Смита, найти распределение стационарного остаточного времени

$$P\{\xi > x\} = \frac{1}{T_1 + T_2} \int_x^\infty \bar{F}(u) du. \quad (29)$$

Отсюда, в частности, величина

$$P\{\xi > 0\} = \frac{T_1}{T_1 + T_2}, \quad (30)$$

называемая коэффициентом готовности, есть вероятность того, что в наугад взятый момент в стационарном режиме элемент будет исправен. Для элемента с конечным временем восстановления важную роль играет еще одна характеристика, которую обычно называют суммарной наработкой S_t , — это суммарное время работы элемента до момента t . Пусть h_x есть момент, в который суммарная наработка достигнет величины x . Тогда справедливы следующие формулы:

$$P\{S_t < x\} = P\{h_x > t\} = \sum_{n=0}^{\infty} [G_n(t-x) - G_{n+1}(t-x)] F_{n+1}(x), \quad (31)$$

где $F_n(x)$ определены выше, $G_0(t) \equiv 1$, а $G_n(t)$ определяются рекуррентно

$$G_{n+1}(t) = \int_0^t G_n(t-x) dG(x),$$

$$G_1(t) \equiv G(t).$$

Формула (31) малоэффективна для расчетов. Однако при большом t для приближенного определения распределения суммарной наработки можно использовать следующее утверждение: при $t \rightarrow \infty$ и $x \rightarrow \infty$ величины s_t и h_x асимптотически нормальны со средними

$$Ms_t \sim \frac{T_1}{T_1 + T_2} t, \quad Mh_x \sim \frac{T_1 + T_2}{T_1} x$$

и дисперсиями

$$Ds_t \sim \frac{T_1^2 T_2^2 \left(\frac{\sigma_1^2}{T_1^2} + \frac{\sigma_2^2}{T_2^2} \right)}{(T_1 + T_2)^3} t,$$

$$Dh_x \sim \frac{\left(\frac{\sigma_1^2}{T_1^2} + \frac{\sigma_2^2}{T_2^2} \right) T_2^2}{T_1} x.$$

Некоторые неравенства для характеристик надежности восстанавливаемого элемента. Если данный элемент входит в систему с большим числом элементов, то на интересующем нас участке времени его вероятность отказа, как правило, мала. Поэтому представляет интерес нахождение оценок характеристик надежности восстанавливаемого

элемента на начальном участке времени, когда $F(t) \ll 1$. Среднее число отказов $H(t)$ оценивается просто

$$F(t) \leq H(t) \leq \frac{F(t)}{1-F(t)}.$$

Для интенсивности отказов справедливо неравенство

$$f(t) \leq h(t) \leq f(t) + M(t) \frac{F(t)}{1-F(t)},$$

где $M(t) = \max_{x \leq t} f(x)$.

Приведем некоторые неравенства для случая, когда распределение $F(t)$ — стареющее. В этом случае функция $H(t)$ удовлетворяет неравенству

$$\frac{t}{T} - 1 \leq H(t) \leq \frac{t}{T}$$

при любом $t \geq 0$, причем левое неравенство справедливо всегда. Для

распределения $v(t)$ мгновенно восстанавливаемого элемента верно неравенство

$$P\{v(t) \geq n\} \leq 1 - \sum_{k=0}^{n-1} \frac{1}{k!} \left[\ln \frac{1}{\bar{F}(t)} \right]^k \bar{F}(t),$$

удобное тем, что его правая часть зависит только от значения F в момент t .

Распределение остаточного времени жизни ζ_t удовлетворяет неравенству

$$P\{\zeta_t > x\} \leq \bar{F}(x),$$

а для стационарного остаточного времени жизни мгновенно восстанавливаемого элемента справедливо неравенство

$$1 - \frac{t}{T} \leq P\{\zeta > t\} \leq e^{-\frac{t}{T}},$$

в котором левое неравенство справедливо всегда.

1. НАДЕЖНОСТЬ СИСТЕМЫ С НЕЗАВИСИМО ОТКАЗЫВАЮЩИМИ ЭЛЕМЕНТАМИ

Общий случай. Предположим что система состоит из n элементов. Время жизни i -го элемента ξ_i есть случайная величина с функцией распределения $F_i(t) = P\{\xi_i < t\}$. Предположим также, что:

- а) величины ξ_i независимы;
- б) состояние элементов системы (исправен-неисправен) однозначно определяет состояние всей системы;
- в) после отказа элементы не восстанавливаются. Состояние всех элементов системы мы будем задавать двоичным вектором

$$e(t) = [e_1(t), e_2(t), \dots, e_n(t)],$$

где $e_i(t) = 0$, если в момент t i -й элемент исправен, и $e_i(t) = 1$, если к этому моменту он уже отказал. Вследствие наложенных выше ограничений все множество $E = \{e\}$ из 2^n состояний разбивается на два подмножества $E = E_+ + E_-$, где E_+ — множество исправных состояний системы; E_- — множество неисправных состояний системы. Ниже множество E_+ мы будем определять задавая соединение элементов в системе. На это разбиение в теории надежности всегда накладывается так называемое условие монотонности. Чтобы сформулировать его, введем в множестве векторов упорядоченность — скажем, что $e = (e_1, e_2, \dots, e_n) < e' = (e'_1, e'_2, \dots, e'_n)$, если для любого i $e_i \leq e'_i$. Тогда условие монотонности записывается так: для любых векторов e и e' из условий $e < e'$ и $e \in E_-$ следует, что $e' \in E_-$.

Иными словами, если система находится в неисправном состоянии e , то дополнительные отказы элементов, при которых мы попадаем в новое состояние e' , не могут перевести систему в исправное состояние.

Пусть ξ — время безотказной работы системы и $F(t) = P\{\xi < t\}$ — его функция распределения. Нетрудно написать общее выражение для $F(t)$:

$$F(t) = \sum_{e \in E_-} p(e), \quad (1)$$

где

$$p(e) = \prod_{i=1}^n \bar{F}_i^{1-e_i}(t) F_i^{e_i}(t), \quad 0^0 = 1$$

есть вероятность того, что в момент t система находится в состоянии e .

Если множество E_- удовлетворяет сформулированному выше условию монотонности, то вероятность $F(t)$ будет монотонно возрастать по каждому аргументу $F_i(t)$. Поэтому, имея верхние или нижние оценки надежности элементов и подставляя их в (1), получим соответственно верхние и нижние оценки надежности системы.

Формула (1) при небольшом числе элементов в системе позволяет рассчитать надежность системы при любом соединении ее элементов (т. е. при любом множестве E_-). Если число n велико, то эта формула дает возможность эффективно рассчитать надежность лишь для самых простых соединений, когда множество E_- (или E_+) состоит из небольшого числа элементов. Для более сложных соединений число слагаемых в формуле (1) будет очень велико, что делает эту формулу неэффективной. Ниже рассмотрены некоторые методы, позволяющие эф-

эффективно находить или оценивать надежность системы.

Последовательное и параллельное соединения. Рассмотрим сначала простейшие и вместе с тем наиболее часто встречающиеся на практике соединения.

Скажем, что элементы в системе соединены последовательно, если отказ любого элемента вызывает отказ системы. В этом случае вероятность безотказной работы системы

$$\bar{F}(t) = \prod_{i=1}^n \bar{F}_i(t) \quad (2)$$

Если $\sum_{i=1}^n F_i(t) \ll 1$, то справедлива приближенная формула

$$F(t) \approx \sum_{i=1}^n F_i(t), \quad (3)$$

погрешность которой не превосходит

$$\frac{1}{2} \left(\sum_{i=1}^n F_i(t) \right)^2$$

Пусть $\lambda_i(t)$ — интенсивность отказа i -го элемента, а $\lambda(t)$ — интенсивность отказа системы. Из формулы (2) следует, что

$$\lambda(t) = \sum_{i=1}^n \lambda_i(t), \quad (4)$$

т. е. при последовательном соединении интенсивности отказа складываются. В частности, если время безотказной работы каждого элемента имеет показательное распределение $\bar{F}_i(t) = e^{-\lambda_i t}$, то и время безотказной работы системы тоже имеет показательное распределение

$$\bar{F}(t) = e^{-\lambda t}, \quad (5)$$

$$\text{где } \lambda = \sum_{i=1}^n \lambda_i.$$

Пусть $T_i = \int_0^{\infty} \bar{F}_i(t) dt$ — среднее время жизни i -го элемента, а

$T = \int_0^{\infty} \bar{F}(t) dt$ — среднее время жизни системы.

Для случая показательного распределения, как следует из (5),

$$T = \left(\sum_{i=1}^n \frac{1}{T_i} \right)^{-1} \quad (6)$$

Если все элементы, составляющие систему, стареющие, то, как видно из (4), сама система также будет стареющей. В этом случае справедливо неравенство

$$T \geq \left(\sum_{i=1}^n \frac{1}{T_i} \right)^{-1} \quad (7)$$

В общем случае оценка среднего времени T представляет довольно трудную проблему.

Рассмотрим теперь другой весьма распространенный тип соединения элементов. Скажем, что элементы соединены параллельно, если отказ системы наступает только тогда, когда отказывают все элементы. Для этого случая

$$F(t) = \prod_{i=1}^n F_i(t). \quad (8)$$

Параллельное соединение возникает обычно тогда, когда все элементы выполняют одну и ту же функцию. Для ее выполнения достаточно одного элемента, остальные играют роль резервных. Такой тип резервирования называют горячим или нагруженным резервом. В такой ситуации элементы, как правило, бывают одинаковыми и поэтому имеют равную надежность $F_i(t) = F_0(t)$. Тогда

$$F(t) = F_0^n(t). \quad (9)$$

Среднее время жизни такой системы

$$T = \int_0^{\infty} [1 - F_0^n(t)] dt. \quad (10)$$

В некоторых случаях этот интеграл вычисляется. Так, если распределение $F_0(t)$ — показательное $\bar{F}_0(t) = e^{-\lambda t}$,

то

$$T = \frac{1}{\lambda} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} \right). \quad (11)$$

В общем случае можно дать оценку для среднего. Если распределение $F_0(t)$ стареющее, то справедливо неравенство

$$\bar{F}_0(T) \geq \frac{0,56146}{n+1}, \quad (12)$$

которое дает верхнюю оценку $T \leq \leq T_2$, $\bar{F}_0(T_2) = \frac{0,56146}{n+1}$. Если же

касательная к графику $\bar{F}_0(t)$ в точке T_1 , определяемой уравнением $\bar{F}_0(T) = \frac{1}{n+1}$, лежит при $t \geq 0$ ниже

графика $\bar{F}_0(t)$, то справедливо неравенство $T \geq T_1$. Таким образом, при выполнении обоих условий среднее время T удовлетворяет двойному неравенству

$$T_1 \leq T \leq T_2. \quad (13)$$

Кроме параллельного соединения одинаковых элементов, в теории надежности часто используется более общее соединение, которое называют « m из n ». Это означает, что система из n одинаковых элементов исправна, если в ней исправно не менее m элементов. Для этого случая

$$\bar{F}(t) = \sum_{k=m}^n C_n^k \bar{F}_0^k(t) F_0^{n-k}(t). \quad (14)$$

Здесь можно тоже получить оценку типа (13).

2. МЕТОД ПУТЕЙ И СЕЧЕНИЙ. РЕКУРРЕНТНЫЙ МЕТОД

Рассмотрим два общих метода, которые позволяют находить или оценивать надежность системы с независимо отказывающимися элементами при произвольном соединении этих элементов.

Метод путей и сечений. Пусть, как и выше, система состоит из n элементов, которые занумерованы в определенном порядке. Множество элементов с номерами i_1, i_2, \dots, i_k назовем путем A , если все состояния вида

$e = (\dots 0.0..0\dots 0)$, где на местах с номерами i_1, \dots, i_k стоят нули, а на остальных местах — произвольные символы 0 и 1, являются исправными состояниями системы. Другими словами, система исправна, если исправны элементы с номерами i_1, i_2, \dots, i_k независимо от состояния других элементов.

Более того, из всех путей мы будем рассматривать только минимальные пути, которые удовлетворяют еще одному условию — никакое подмножество элементов пути не является путем.

Пусть A_1, A_2, \dots, A_m — все минимальные пути. Символом A_s мы будем обозначать не только s -й путь, но и событие, состоящее в том, что все элементы этого пути исправны. Тогда надежность системы

$$\begin{aligned} \bar{F}(t) = P \left\{ \bigcup_{s=1}^m A_s \right\} &= \sum_{i=1}^m P \{A_i\} - \\ &- \sum_{i < j} P \{A_i A_j\} + \\ &+ \sum_{i < j > k} P \{A_i A_j A_k\} - \end{aligned} \quad (15)$$

Если число минимальных путей невелико, то формула позволяет эффективно рассчитать надежность. В противном случае, поскольку число слагаемых в правой части равно $2^m - 1$, она практически бесполезна. Правда, существует один прием, позволяющий уменьшить объем вычислений. Два пути назовем пересекающимися, если они содержат общий элемент. Два пути назовем связанными, если их можно соединить цепочкой пересекающихся путей. Отношение связности является отношением эквивалентности, и поэтому все множество путей разбивается на классы связанных путей, скажем,

$$\begin{aligned} K_1 &= \{A_1, A_2, \dots, A_{m_1}\}, \\ K_2 &= \{A_{m_1+1}, \dots, A_{m_2}\}, \dots \end{aligned}$$

Тогда из формулы (15)

$$\bar{F}(t) = P \left\{ \bigcup_{s=1}^{m_1} A_s \right\} P \left\{ \bigcup_{s=m_1+1}^{m_2} A_s \right\} \dots$$

Каждая из вероятностей справа вычисляется по формуле (15), и суммарное число операций уменьшается до $(2^{m_1} + 2^{m_2 - m_1} + \dots)$.

Представление (15) полезно еще и тем, что позволяет оценивать снизу вероятность безотказной работы $\bar{F}(t)$. Ведь очень часто не нужно знать точное значение $\bar{F}(t)$, а требуется лишь убедиться, что $\bar{F}(t) \geq p$, где p — некоторый заранее заданный уровень надежности. В этом случае мы нумеруем пути следующим образом. Номер 1 приписываем пути, для которого $P\{A_1\} = \max P\{A_i\}$,

номер 2 приписываем пути, для которого

$$P\{A_1 \cup A_2\} = \max_{i \neq 1} P\{A_1 \cup A_i\},$$

номер 3 — пути, для которого

$$P\{A_1 \cup A_2 \cup A_3\} = \max_{i \neq 1, 2} P\{A_1 \cup A_2 \cup A_i\},$$

и т. д. Тогда для оценки надежности последовательно вычисляем вероятности

$$p_1 = P\{A_1\} < p\{A_1 \cup A_2\} = p_2 < P\{A_1 \cup A_2 \cup A_3\} = p_3 <$$

до тех пор, пока очередное значение p_k не превзойдет величины p .

Аналогично вводится понятие сечения.

Множество элементов с номерами j_1, j_2, \dots, j_s образует сечение, если отказ всех этих элементов приводит к отказу системы независимо от состояния других элементов. Пусть B_1, B_2, \dots, B_e — все минимальные сечения. Тогда

$$F(t) = P\left\{\bigcup_{i=1}^e B_i\right\} = \sum_{i=1}^e P\{B_i\} - \sum_{i < j} P\{B_i B_j\} + \dots \quad (16)$$

Если система высоконадежна и вероятности отказа ее элементов за заданное время малы, то формула

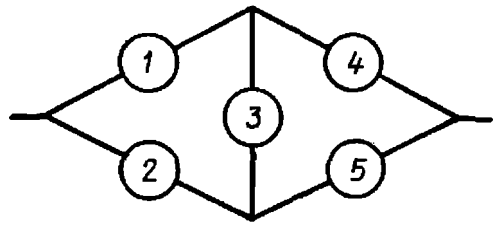


Рис. 1

(16) позволяет эффективно оценивать вероятность отказа системы. В этом случае суммы в формуле (16) очень быстро убывают, а вычисляя последовательно эти суммы, мы получаем попеременно значение $F(t)$ с избытком и недостатком, и, таким образом, первая отброшенная сумма в формуле (16) дает нам оценку абсолютной погрешности вычислений.

Приведем пример. Рассмотрим мостиковую схему (рис. 1), где элементы занумерованы так, как это показано на рисунке, и

$$\bar{F}_1(t) = \bar{F}_2(t) = 0,9; \quad \bar{F}_3(t) = 0,8,$$

$$\bar{F}_4(t) = \bar{F}_5(t) = 0,95.$$

Здесь минимальными сечениями, занумерованными так, как это делалось в методе путей, будут

$$A_1 = (1, 4); \quad A_2 = (2, 3); \\ A_3 = (1, 3, 5); \quad A_4 = (2, 3, 4).$$

Тогда

$$P\{A_1\} = 0,855; \quad P\{A_1 \cup A_2\} = 0,978975;$$

$$P\{A_1 \cup A_2 \cup A_3\} = 0,982395.$$

Если нам нужно проверить, что вероятность безотказной работы не меньше 0,99, то, как видно, на третьем шаге мы в этом убеждаемся.

Посмотрим теперь, что дает для этого примера метод сечений. Здесь минимальными сечениями будут

$$B_1 = \{1, 2\}; \quad B_2 = \{4, 5\}; \\ B_3 = \{1, 3, 5\} \text{ и } B_4 = \{2, 3, 4\}.$$

Вычисляя суммы в формуле (16), получаем

$$F(t) = 0,0127 - 0,00033 + \dots,$$

т. е. вычисление только двух сумм в формуле (16) дает гораздо более точную оценку

$$0,01237 \leq F(t) \leq 0,0127,$$

чем метод путей (и притом двустороннюю оценку).

Рекуррентный метод. Суть этого метода состоит в следующем.

Предположим, что наша система является n -м членом в некоторой последовательности систем $S_1, S_2, \dots, S_n, \dots$, в которой каждая последующая система получается из предыдущей добавлением нескольких новых элементов и связей между элементами. Предположим далее, что каждой системе S_n соответствует некоторое вероятностное пространство $\Omega_n = \{\omega_n\}$, задающее состояние этой системы. Допустим, что мы выбрали эти пространства так, что выполнены два условия:

1) отказ системы S_n является событием, измеримым относительно пространства Ω_n , другими словами, по любому состоянию системы ω_n мы можем однозначно сказать, исправна она или нет;

2) последовательность вероятностных пространств Ω_n образует цепь Маркова, т. е. по состоянию системы S_n однозначно определяются вероятности состояний системы S_{n+1} независимо от состояния предшествующих систем S_k при $k < n$.

Если выполнены эти условия, то мы можем методами марковских цепей рассчитать вероятности $P(\omega_n)$, и тогда вероятность безотказной работы системы S_n

$$\bar{F}(t) = \sum_{\omega_n \in E_+} P(\omega_n).$$

Конечно, этот метод применим только к тем системам, структура которых имеет достаточно правильный, рекуррентный характер. Но даже для таких систем обычно возникают трудности, связанные с выбором пространств состояний Ω_n . Если мы возьмем Ω_n с небольшим числом точек, то не будет выполнено одно из условий (1 или 2). Если взять Ω_n с большим числом точек, то будет чрезмерно велик объем

вычислений. В качестве примера оценим надежность ветвящейся системы.

Предположим, что главный элемент, у которого надежность $\bar{F}_0(t) = p_0$, управляет a_1 одинаковыми элементами первого ранга, для которых $\bar{F}_1(t) = p_1$. Каждый из них в свою очередь управляет a_2 элементами второго ранга и $\bar{F}_2(t) = p_2$ и т. д. Наконец, каждый из элементов $(n-1)$ -го ранга управляет a_n элементами последнего n -го ранга. Эти элементы мы будем называть выходными. Их число, очевидно, равно $N_n = a_1, a_2, \dots, a_n$. Отказ каждого элемента выводит из строя подчиненные ему элементы. Поэтому каждый элемент может отказать по двум причинам — либо он сам отказал, либо отказал элемент, которому он подчинен. Такую систему можно трактовать, например, как систему передачи команд от головного элемента к выходным элементам. Такая система считается исправной, если из N_n выходных элементов работает не менее m элементов. Применим здесь рекуррентный метод. Под системой S_k будем понимать систему, составленную из элементов первых k рангов. В качестве состояний системы S_k возьмем число работающих элементов k -го ранга. Очевидно, $\omega_k = 0, 1, 2, \dots; N_k = a_1, a_2, \dots, a_k$. Нетрудно заметить, что при таком выборе S_k и $\Omega_k = \{\omega_k\}$ выполнены условия 1 и 2. Пусть $P_k(z) = M_z^{\omega_k}$ — производящая функция числа работающих элементов k -го ранга. Тогда справедлива рекуррентная формула

$$P_k(z) = P_{k-1}[(p_k z + q_k)^{a_k}];$$

$$P_0(z) = p_0 z + q_0, \quad q_0 = 1 - p_0,$$

$$q_k = 1 - p_k, \quad n = 1, 2, \dots \quad (17)$$

При небольшом числе выходных элементов эта формула позволяет рассчитать надежность ветвящейся системы. Если же число выходных элементов велико, то эффективно рассчитываются только моменты ω_n . Например,

$$M\omega_n = N_n p_0 p_1 p_2 \dots p_n. \quad (18)$$

3. МОДЕЛИ ЗАВИСИМОСТИ ОТКАЗОВ ЭЛЕМЕНТОВ В СИСТЕМЕ

Рассмотрим некоторые типы зависимых отказов. Другие типы зависимости, описываемые многомерным нормальным и многомерным экспоненциальным распределениями, рассмотрены в монографии [1].

Дискретная зависимость. Предположим, что элементы системы в момент своего отказа мгновенно восстанавливаются. У каждого k -го элемента есть некоторое время безотказной работы, и эти времена независимы. Однако в момент отказа k -го элемента он может с вероятностью $P_k(A)$ вывести из строя некоторое множество A других элементов, которые в этот момент восстанавливаются. Таким образом, с течением времени у каждого элемента происходят либо собственные отказы, либо несобственные отказы, происходящие из-за отказа других элементов. Приведем два характерных примера:

1) при отказе некоторых элементов цветного телевизора в нем может возникнуть пожар, который повредит другие элементы (или даже уничтожит весь телевизор);

2) при замене одного отказавшего элемента мы можем повредить другие элементы. Например, при отвинчивании болта мы можем повредить прокладку, которую также придется заменить.

Ясно, что такого рода зависимость увеличивает интенсивности замен элементов, и ее следует учитывать при расчете ЗИПа. Здесь основной характеристикой является интенсивность отказов каждого элемента с учетом его несобственных отказов. Приведем решение этой задачи для случая, когда собственное время безотказной работы каждого k -го элемента имеет показательное распределение

$$P\{\xi_k > t\} = e^{-\lambda_k t}$$

Если бы несобственных отказов не было, то поток отказов каждого k -го элемента был бы пуассоновским потоком с интенсивностью λ_k . Пусть λ_k — полная интенсивность отказов k -го

элемента с учетом несобственных отказов. Тогда

$$\Lambda_k = \lambda_k + \sum_{A \ni k} \sum_{i \neq k} \lambda_i P_i(A). \quad (19)$$

(Вторая сумма берется по всем множествам элементов, содержащим k -й элемент, а сам поток отказов будет пуассоновским).

Зависимость отказов вследствие зависимости начальных параметров элемента. Случайное время безотказной работы элемента определяется его начальными параметрами и режимом, в котором он работает. Отвлекаясь пока от второго фактора, рассмотрим только первый. В современных системах элементы, из которых она комплектуется, часто изготавливаются на одном заводе и даже на одном станке. Если на завод поступает сырье, которое меняется от партии к партии, то это влияет на изготавливаемые элементы, снижая или повышая их надежность. Если в конце квартала на заводе аврал, то это, как правило, приводит к снижению надежности изготавливаемых в этот период элементов. Но система обычно комплектуется из элементов, изготовленных примерно в одно и то же время, поэтому при таких условиях надежность, по крайней мере группы элементов, будет ниже или выше средней. Если учитывать этот фактор, то можно очень сильно ошибиться при прогнозировании надежности. Приведем один конкретный пример, который носит модельный характер. Предположим, что на станке изготавливаются одинаковые элементы, из которых в конце конвейера комплектуется система, причем каждая система составляется из 100 подряд изготовленных элементов. На станке периодически возникают разладки, и в период от возникновения разладки до ее обнаружения и устранения станок изготавливает неисправные элементы. Пусть на нормальном периоде своей работы станок изготавливает 10 000 элементов, а за период разладки 100 элементов. Когда проводят заводские испытания этих элементов на надежность, то выбирают элементы наугад, и тогда вероятность того, что элемент исправный, по результатам таких ис-

пытаний будет равна $p = 0,99$. Если рассчитывать надежность системы обычным образом, предполагая независимость, то вероятность того, что система исправна, будет равна

$$p_1 = (0,99)^{100} = 0,366.$$

Но ведь при наших условиях 100 дефектных элементов, изготовленных в период разладки станка, попадут в одну или две системы, а за период между разладками будет изготовлено около 100 исправных систем. Поэтому истинная вероятность исправной системы будет находиться в интервале $0,98 \leq p_2 \leq 0,99$.

Этот несколько утрированный пример показывает, что зависимостью начальных параметров (в тех случаях, когда она есть) пренебрегать нельзя.

Зависимость отказов элементов, работающих в общем режиме. Под режимом $\varepsilon(t)$ мы понимаем, вообще говоря, многомерный случайный процесс, описывающий внешние и внутренние условия в системе, от которого зависят времена безотказной работы элементов системы. Предположим, что элементы в системе соединены последовательно и при фиксированном режиме условные времена безотказной работы элементов независимы. Тогда вероятность безотказной работы системы

$$\begin{aligned} \bar{F}(t) &= P \{ \xi > t \} = \\ &= \int_{\Omega} P \{ \xi > t/\varepsilon \} P \{ d\varepsilon \} = \\ &= \int_{\Omega} \prod_{i=1}^n P \{ \xi_i > t/\varepsilon \} P \{ d\varepsilon \} = \\ &= \int_{\Omega} \prod_{i=1}^n \bar{F}_i(t/\varepsilon) P \{ d\varepsilon \}, \end{aligned} \quad (20)$$

где $\bar{F}_i(t/\varepsilon)$ — условная вероятность безотказной работы i -го элемента при условии, что фиксирован режим ε ,

$$\bar{F}_i(t) = \int_{\Omega} \bar{F}_i(t/\varepsilon) P \{ d\varepsilon \},$$

Ω — пространство траекторий режима, $P \{ \cdot \}$ — вероятностная мера на этих траекториях. Формула (20), как пра-

вило, бесполезна, поскольку мы не знаем распределения режима и, главное, не знаем, как зависит от режима надежность элемента. Однако одно полезное следствие из этой формулы вывести можно. Предположим, что для любых i и j и любых режимов ε' и ε'' справедливо неравенство

$$[\bar{F}_i(t/\varepsilon') - \bar{F}_i(t/\varepsilon'')] [\bar{F}_j(t, \varepsilon') - \bar{F}_j(t, \varepsilon'')] \geq 0,$$

т. е. переход от одного режима к другому или повышает надежность всех элементов, или уменьшает ее. Тогда

$$\bar{F}(t) \geq \prod_{i=1}^n \bar{F}_i(t). \quad (21)$$

Другими словами, в этом случае, рассчитывая надежность в предположении независимости отказов, мы можем только занижить истинную надежность. Еще один тип зависимости будет рассмотрен ниже.

4. МАРКОВСКИЕ ОДНОРОДНЫЕ ПРОЦЕССЫ С КОНЕЧНЫМ ИЛИ СЧЕТНЫМ ЧИСЛОМ СОСТОЯНИЙ. ПРОЦЕССЫ ЧИСТОЙ ГИБЕЛИ. ПРОЦЕСС РОЖДЕНИЯ И СМЕРТИ

Определение Случайный процесс κ с состояниями $0, 1, 2, \dots, n$, называют марковским однородным процессом, если для любых состояний i_1, i_2, \dots, i_k, i и $j, i \neq j$ и любых моментов $0 < t_1 < t_2 < \dots < t_k < t < t + h$ условная вероятность

$$\begin{aligned} P \{ \kappa(t+h) = j \mid \kappa(t_1) = i_1, \dots, \kappa(t_k) = i_k, \kappa(t) = i \} &= P \{ \kappa(t+h) = j \mid \kappa(t) = i \} = \lambda_{ij} h + 0(h). \end{aligned} \quad (22)$$

Если моменты t_1, \dots, t_k называть прошлым, момент t — настоящим, а момент $t+h$ — будущим, то приведенное определение означает, что если известно состояние процесса в настоящий момент, то при этом условии вероятности будущего не зависят от прошлого поведения процесса. Обозначим

$$\lambda_{ii} = - \sum_{j \neq i} \lambda_{ij}.$$

Есть еще одно определение такого процесса, которое эквивалентно приведенному выше, — процесс $x(t)$ называют марковским, если:

а) длина каждого интервала ξ_i , на котором процесс находится в состоянии i , не зависит от поведения процесса вне этого интервала и имеет показательное распределение:

$$P \{ \xi_i > t \} = e^{-\lambda_i t}$$

б) последовательность проходимых процессом состояний образует марковскую цепь с вероятностями перехода $p_{ij} = -\frac{\lambda_{ij}}{\lambda_{ii}}$ (если $\lambda_{ii} = 0$, то полагаем $p_{ij} = 0$; в этом случае состояние i будет поглощающим).

При решении некоторых задач, связанных с марковскими процессами, удобно использовать именно это второе определение марковского процесса.

Вероятности состояний процесса. Пусть $p_i(t) = P \{ x(t) = i \}$. Тогда эти вероятности удовлетворяют системе уравнений:

$$\begin{cases} p_j'(t) = \sum_{i=0}^{\infty} p_i(t) \lambda_{ij}, \\ j = 0, 1, 2, \dots; \\ \sum_{j=0}^{\infty} p_j(t) = 1. \end{cases} \quad (23)$$

Чтобы эта система имела единственное решение, необходимо задать начальное распределение процесса

$$P_j(0), \quad j = 0, 1,$$

Систему (23) часто удобно решать, применяя преобразование Лапласа. Обозначим

$$a_i(z) = \int_0^{\infty} e^{-zt} p_i(t) dt.$$

Тогда, умножая обе части уравнений (23) на e^{-zt} , интегрируя и используя известные свойства преобразования Лапласа, находим

$$\begin{cases} -p_j(0) + za_j(z) = \sum_{i=0}^{\infty} a_i(z) \lambda_{ij}; \\ \sum_{j=0}^{\infty} a_j(z) = \frac{1}{z}. \end{cases} \quad (24)$$

В случае, когда число состояний конечно, линейная алгебраическая система решается обычным образом и ее решения $a_j(z)$ будут рациональными дробями, которые легко обращаются и дают представление $p_j(t)$ в виде

$$p_j(t) = \sum_k q_{kj}(t) e^{-\lambda_k t}$$

где $q_{kj}(t)$ — многочлены степени на единицу меньшей кратности корня $-\lambda_k$.

Если число состояний бесконечно, то при решении системы (24) возникают значительные трудности. Иногда эта система решается методом производящих функций, когда мы вводим

$$\text{функцию } A(z, \omega) = \sum_{j=0}^{\infty} a_j(z) \omega^j \text{ и}$$

пытаемся из системы (24) получить уравнение для функции $A(z, \omega)$.

Если марковский процесс эргодический, т. е. существуют стационарные вероятности $p_i = \lim_{t \rightarrow \infty} p_i(t)$, то эти

стационарные вероятности удовлетворяют системе

$$\begin{cases} \sum_{i=0}^{\infty} \lambda_{ij} p_i = p_j; \\ \sum_{j=0}^{\infty} p_j = 1. \end{cases} \quad (25)$$

Вычисление основных характеристик марковского процесса. В теории надежности изменение состояний системы во времени описывается некоторым марковским процессом. Пусть E — множество состояний этого процесса, которое всегда выбирают так, чтобы по состоянию процесса в момент t можно было однозначно сказать, исправна система или нет. Поэтому множество E разбивают на два подмножества $E = E_+ + E_-$, где E_+ — мно-

жество исправных состояний системы; E_- — множество неисправных состояний системы. Переход процесса из E_+ в E_- называют отказом системы, а обратный переход — восстановлением. Обычно предполагают, что в начальный момент $t = 0$ система исправна. Тогда на оси времени будут чередоваться участки исправного состояния системы $\tau'_0, \tau'_1, \dots, \tau'_n$, и участки неисправного состояния системы $\tau''_1, \tau''_2, \dots, \tau''_n$.

Если процесс с течением времени входит в стационарный режим, то распределение этих интервалов в пределе не будет зависеть от номера интервала $\tau'_n \rightarrow \tau', \tau''_n \rightarrow \tau''$ (здесь символом \rightarrow обозначается слабая сходимость, т. е. сходимость распределений). Величину τ'_0 называют временем до первого отказа, а величину τ' — наработкой на один отказ. В теории надежности используется еще одна характеристика τ_t — время от момента t до первого справа отказа; при этом, если момент t попадает на участок неисправного состояния, то $\tau_t = 0$. Соответствующую стационарную характеристику обозначим $\tau, \tau_t \rightarrow \tau$. Основными характеристиками надежности являются следующие величины: $\tau'_0, \tau', \tau'', \tau$, а также их средние $M\tau'_0 = T_0, M\tau' = T_1, M\tau'' = T_2, M\tau = T$ и коэффициент готовности $k_r = \frac{T_1}{T_1 + T_2}$.

Покажем, как для произвольного марковского процесса следует находить эти характеристики. Множество состояний процесса $E = \{0, 1, 2, \dots, n, \dots\}$. Поскольку в задачах надежности множество исправных состояний E_+ всегда конечно, можно, не ограничивая общности, считать, что $E_+ = \{0, 1, 2, \dots, n\}$, $E_- = \{n+1, n+2, \dots\}$.

Обычно предполагается, что в начальный момент все элементы системы исправны, и это состояние обозначается через «0». Таким образом, время до первого отказа τ'_0 есть время перехода процесса из состояния «0» в множество E_- . Введем величины $\tau_i(A)$, $i \notin A$ — время перехода процесса из состояния i в множество A , и их преобразование Лапласа

$$\varphi_i = \varphi_i(z, A) = M e^{-z\tau_i(A)}.$$

Функции $\varphi_i(z, A)$ удовлетворяют алгебраической системе

$$\sum_{j \notin A} (z\delta_{ij} - \lambda_{ij}) \varphi_j(z, A) = 0, \quad (26)$$

$$= \sum_{j \in A} \lambda_{ij} \varphi_j(z, A), \quad i \notin A,$$

где δ_{ij} — символ Кронекера.

Поскольку в нашем случае $A = E_-$, а $\bar{A} = E_+$, эта система будет конечной и функции $\varphi_j(z, A)$ будут некоторыми рациональными дробями, которые легко обращаются и дают явное распределение величин $\tau_i(A)$. Отсюда находят распределение времени до первого отказа $\tau'_0 = \tau_0(E_-)$. Дифференцируя уравнения (26) по z и подставляя $z = 0$, получают систему уравнений для средних $T_i(A) = M\tau_i(A)$

$$\sum_{j \notin A} \lambda_{ij} T_j(A) + 1 = 0, \quad i \notin A, \quad (27)$$

откуда среднее время до первого отказа

$$M\tau_0(E_-) = T_0.$$

Покажем также, как определяется распределение стационарных величин τ', τ'', τ и их средние.

Пусть $q_i(E_+)$ есть вероятность того, что в момент восстановления системы в стационарном режиме процесс попадает в состояние $i \in E_+$, а $q_j(E_-)$ — вероятность того, что в момент восстановления системы в стационарном режиме процесс попадает в состояние $j \in E_-$. Эти вероятности выражаются так:

$$q_i(E_+) = \frac{\sum_{j \in E_-} p_j \lambda_{ji}}{\sum_{i \in E_+} \sum_{j \in E_-} p_j \lambda_{ji}}; \quad (28)$$

$$q_j(E_-) = \frac{\sum_{i \in E_+} p_i \lambda_{ij}}{\sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij}},$$

где p_i — стационарные вероятности процесса.

Тогда искомые распределения находят из выражений

$$\begin{aligned}
 P\{\tau' < t\} &= \sum_{i \in E_+} q_i(E_+) \times \\
 &\times P\{\tau_i(E_-) < t\}; \\
 P\{\tau'' < t\} &= \sum_{j \in E_-} q_j(E_-) \times \\
 &\times P\{\tau_j(E_+) < t\}.
 \end{aligned} \tag{29}$$

Средние $T_1 = M\tau'$ и $T_2 = M\tau''$ находят еще проще:

$$\begin{aligned}
 T_1 &= \frac{\sum_{i \in E_+} p_i}{\sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij}}; \\
 T_2 &= \frac{\sum_{i \in E_-} p_i}{\sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij}}.
 \end{aligned} \tag{30}$$

Распределение остаточного времени жизни τ определяется формулой

$$P\{\tau > t\} = \sum_{i \in E_+} p_i P\{\tau_i(E_-) > t\}. \tag{31}$$

Наконец, коэффициент готовности выражается через стационарные вероятности

$$k_r = \frac{T_1}{T_1 + T_2} \sum_{i \in E_+} p_i. \tag{32}$$

Процесс чистой гибели и его обобщение. Приведенные выше характеристики используются в основном при анализе восстанавливаемых систем. В этой главе, где изучаются лишь невосстанавливаемые системы, мы рассмотрим те марковские процессы, которые возникают при описании невосстанавливаемых систем с зависимыми отказами.

Пусть состояние системы задается, как это уже говорилось выше, двоичным вектором

$$e(t) = [e_1(t), \dots, e_n(t)].$$

Траекторию процесса на интервале $[0, t]$ обозначим e_0^t . Если предположить, что нет зависимости начальных

параметров и нет зависимости отказов из-за общего случайного режима, то единственно возможным типом зависимости может быть зависимость надежности одного элемента в момент t от состояния других элементов до момента t . Эту зависимость в общем случае можно описать так: если известна траектория процесса до момента, то вероятность перейти за время Δt в состояние e равна

$$\lambda(e_0^t, e) \Delta t + o(\Delta t).$$

Эта общая формула учитывает и возможность появления в один момент нескольких отказов. Более частным является тот случай, когда в данный момент произойдет только один отказ. Тогда вероятность появления отказа i элемента на интервале $(t_1 t + \Delta t)$ будет иметь вид

$$\lambda_i(e_0^t) \Delta t + o(\Delta t).$$

Еще более частным является тот случай, когда интенсивность отказа i элемента не зависит от прошлого и определяется только значением процесса в момент t . Тогда

$$\lambda_i(e_0^t) = \lambda_i(e, t).$$

В этом случае процесс $e(t)$ будет марковским неоднородным процессом с конечным числом состояний (такой процесс определяется условием (32), в котором интенсивности $\lambda_{ij} = \lambda_{ij}(t)$ зависят от времени). Наконец, можно рассмотреть случай, когда интенсивность отказа i элемента при заданном состоянии системы e не зависит от времени и равна $\lambda_i(e)$. Такой процесс называют обобщенным процессом чистой гибели, и мы рассмотрим его ниже. Но сначала введем еще более простой процесс, который описывает наведение невосстанавливаемых систем с однотипными элементами и который называют процессом чистой гибели.

Рассмотрим систему, состоящую из n одинаковых элементов, выполняющих одну и ту же функцию. Предположим, что интенсивность отказа каждого элемента зависит только от числа неисправных элементов в системе. Такое предположение естественно, поскольку при отказе части элементов на дру-

гие неотказавшие элементы ложится большая рабочая нагрузка, и эта нагрузка зависит только от числа отказавших элементов. Система отказывает тогда, когда отказывают все ее n элементов. Пусть λ_k — суммарная интенсивность отказа (сумма всех интенсивностей отказа неотказавших элементов) в состоянии, когда в системе неисправно k элементов. В этом случае процесс $x(t)$, равный числу отказавших к моменту t элементов, будет марковским однородным процессом с состояниями $E = \{0, 1, 2, \dots, n\}$, причем множество неисправных состояний состоит из одного состояния $E_- = \{n\}$. Интенсивности перехода λ_{ij} будут здесь такими:

$$\lambda_{i, i+1} = \lambda_i, \quad \lambda_{i, i} = -\lambda_i, \quad \lambda_n = 0,$$

а все другие интенсивности $\lambda_{ij} = 0$. Такой процесс называют процессом чистой гибели. Для него система (23) имеет вид

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t); \\ p'_k(t) = \lambda_{k-1} p_{k-1}(t) - \lambda_k p_k(t); \\ k = 1, 2, \dots, n-1 \\ p'_n(t) = \lambda_{n-1} p_{n-1}(t) - \lambda_n p_n(t), \\ p_0(0) = 1, \quad p_k(0) = 0, \quad k > 0. \end{cases} \quad (33)$$

Эту систему легко решить, например, с помощью преобразования Лапласа:

$$p_k(t) = \lambda_0 \lambda_1 \dots \lambda_{k-1} \times \sum_{i=0}^k \frac{1}{\omega'_k(-\lambda_i)} e^{-\lambda_i t}, \quad (34)$$

$$\text{где } \omega_k(z) = (z + \lambda_0)(z + \lambda_1) \dots (z + \lambda_k).$$

В частности,

$$p_n(t) = P\{\tau'_0 < t\} = 1 + \lambda_0 \lambda_1 \dots \lambda_{n-1} \times \sum_{i=0}^{n-1} \frac{1}{\omega'_n(-\lambda_i)} e^{-\lambda_i t}, \quad (35)$$

есть вероятность отказа системы до момента t .

Среднее время до отказа

$$T_0 = M\tau'_0 = \frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \dots + \frac{1}{\lambda_{n-1}}. \quad (36)$$

Выражение для среднего легче получить не из распределения (35), а из того факта, что процесс чистой гибели последовательно проходит состояния $0, 1, 2, \dots, n-1$, проводя в каждом из них показательно распределенное время. Следовательно, время до отказа есть сумма независимых экспоненциально распределенных величин

$$\tau_0 = \xi_0 + \xi_1 + \dots + \xi_{n-1},$$

$$\text{где } p\{\xi_k > t\} = e^{-\lambda_k t}$$

Это позволяет сразу же найти не только среднее, но и все моменты τ'_0 . Часто рассматривают такую систему на начальном участке времени, где вероятность отказа системы $p_n(t)$ мала. Для этого случая вместо громоздкого выражения (35) можно использовать очень простую оценку:

$$\frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{n!} t^n \left(1 - \frac{1}{n+1} \sum_{k=0}^{n-1} \lambda_k t \right) \leq \leq p_n(t) \leq \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{n!} t^n. \quad (37)$$

Рассмотрим теперь более общий процесс, который описывает поведение произвольной невосстанавливаемой системы с различными элементами. Предположим, что система состоит из n элементов и состояние всех элементов в каждый момент t задается двоичным вектором

$$e(t) = [e_1(t) \dots e_n(t)].$$

Пусть в начальный момент $t = 0$ все элементы исправны и интенсивность каждого неотказавшего элемента зависит только от состояния других элементов $\lambda_i(e) = \lambda_i(e_1, e_2, \dots, e_n)$. Множество E , как всегда, разбито на множество исправных и неисправных состояний $E = E_+ + E_-$, удовлетворяющих условию монотонности.

Процесс $e(t)$, который будет, очевидно, марковским однородным процессом с конечным числом состояний, называют часто обобщенным процессом чистой гибели. Поскольку число состояний такого процесса равно 2^n , система уравнений для вероятностей состояний $p_e(t)$ весьма громоздка и неудобна для эффективного решения. Поэтому предложим метод, позволяющий достаточно эффективно находить надежность такой системы.

Назовем путем π последовательность состояний, которые процесс $e(t)$ проходит от начального нулевого состояния до отказового

$$\pi = (0, e^{(1)}, e^{(2)}, \dots, e^{(n)}),$$

$$e^{(k)} \in E_+ \left| \begin{array}{l} k < m, \\ e^{(m)} \in E_- \end{array} \right.$$

При этом при переходе из состояния $e^{(k-1)}$ в состояние $e^{(k)}$ отказывает элемент с номером i_k . Считая $\lambda_i(e) = 0$, если в состоянии e i -й элемент уже отказал, обозначим

$$\lambda(e) = \sum_{i=1}^n \lambda_i(e).$$

Тогда, как следует из второго определения марковского процесса, $p_i(e) = \frac{\lambda_i(e)}{\lambda(e)}$ есть вероятность того, что при отказе элемента в состоянии e откажет элемент с номером i . Отсюда следует, что вероятность того, что процесс $e(t)$ пройдет путь π ,

$$p(\pi) = p_{i_1}(0) p_{i_2}(e^{(1)}) \dots p_{i_m}(e^{(m-1)}).$$

Вероятность отказа системы

$$F(t) = P\{\tau'_0 < t\} = \sum_{\pi} F(t/\pi) p(\pi), \tag{38}$$

где сумма берется по всем возможным путям π , а $F(t/\pi)$ есть условная вероятность отказа при условии, что процесс прошел путь π . Так как, проходя

путь π , процесс $e(t)$ в каждом состоянии $e^{(k)}$ проводит экспоненциально распределенное время с параметром $\lambda(e^{(k)})$, то условный процесс $e(t)$ при условии, что пройден путь π , будет обычным процессом чистой гибели с интенсивностями перехода $\lambda(e^{(k)})$. Поэтому вероятность $F(t/\pi)$ можно вычислять по формуле (35). Если вероятность отказа системы $F(t)$, мала, можно оценить ее используя неравенство (37):

$$F(t) \leq \sum_{\pi} \frac{\lambda_{i_1}(0) \lambda_{i_2}(e^{(1)}) \dots \lambda_{i_m}(e^{(m-1)})}{m!} t^m, \tag{39}$$

относительная погрешность которой не превосходит

$$\max_{e \in E_+} \lambda(e) t.$$

Рассмотрим в качестве иллюстрации последнего метода один пример. Пусть система состоит из четырех элементов, соединенных, как показано на рис. 2. Предположим, что закодированная информация поступает на элементы 1 и 2, которые параллельно декодируют ее. По этой информации принимается решение; оно поступает на элементы 3 и 4, которые кодируют его и передают по назначению. Пусть интенсивность отказа каждого из элементов 1 и 2, если они оба работают, равна $\lambda_1 = 0,0001 \text{ ч}^{-1}$. Если же один из них отказал, то интенсивность отказа оставшегося элемента будет $\lambda'_1 = 0,00015 \text{ ч}^{-1}$. То же самое имеет место и для второй пары элементов 3 и 4. Для них соответствующие интенсивности равны: $\lambda_2 = 0,0005 \text{ ч}^{-1}$, $\lambda'_2 = 0,0001 \text{ ч}^{-1}$. Система отказывает,

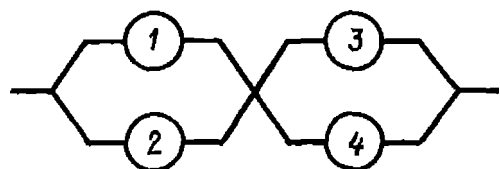


Рис. 2.

если отказывают оба элемента в одной из пар. Требуется найти вероятность отказа такой системы за время $t = 1000$ ч. Воспользуемся неравенством (39). Задавая путь номерами последовательно отказывающихся элементов, видим, что имеется четыре пути длины два: (1,2); (2,1); (3,4); (4,3), из которых несимметричны только два пути, и 16 путей длины три, из которых, опять-таки, согласно симметрии, достаточно взять только два пути — (1, 3, 2) и (1, 3, 4); остальные пути дают те же слагаемые (39), что первый или второй путь. Отсюда

$$F(t) \leq 2 \frac{\lambda_1 \lambda'_1}{2} t^2 + 2 \frac{\lambda_2 \lambda'_2}{2} t^2 + \\ + 8 \frac{\lambda_1 \lambda_2 \lambda'_1}{6} t^3 + 8 \frac{\lambda_1 \lambda_2 \lambda'_2}{6} t^3 = 0,02167,$$

причем относительная погрешность не превосходит 0,225.

В сумме (39) можно брать только минимальные по длине пути; вклад в сумму более длинных путей обычно мал.

СПИСОК ЛИТЕРАТУРЫ

1. Барлоу Р., Прошан Ф. Статистическая теория надежности. М.: Мир, 1978. 488 с.
2. Вопросы математической теории надежности/Под ред. Б. В. Гнеденко. М.: Радио и связь, 1983. 386 с.

1. РЕГЕНЕРИРУЮЩИЕ ПРОЦЕССЫ. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Определение регенерирующего процесса. Предположим, что на некотором вероятностном пространстве определена пара — случайный процесс $\mu(t)$ и момент его остановки ξ , $t < \xi$. Пусть

$$\mu_1(t), \xi_1; \mu_2(t), \xi_2; \mu_3(t), \xi_3,$$

$$\mu_n(t), \xi_n;$$

— независимые одинаково распределенные пары. Тогда процесс

$$x(t) = \mu_n(t - t_{n-1}), \text{ если}$$

$$t_{n-1} \leq t < t_n, \quad t_0 = 0,$$

где $t_n = \xi_1 + \xi_2 + \dots + \xi_n$ называют регенерирующим.

Моменты t_n называют моментами регенерации, а интервалы $[t_{n-1}, t_n)$ — периодами регенерации. Другими словами, в моменты регенерации процесс «забывает» прошлое и начинается заново, причем отсчитываемый от момента t_n процесс имеет то же распределение, что и с момента $t_0 = 0$. Предположим, что на каждом периоде регенерации $[t_{n-1}, t_n)$ в некоторый момент $t_{n-1} + \eta_n$, $0 \leq \eta_n < \xi_n$, может произойти событие A_n , причем пара (A_n, η_n) не зависит от номера n и от поведения процесса $x(t)$ вне данного периода регенерации. В задачах теории надежности процесс $x(t)$ описывает эволюцию системы. Множество состояний процесса E , как и выше, состоит из множества исправных состояний E_+ и множества неисправных состояний E_- . В этом случае событие A_n интерпретируется как отказ системы (т. е. попадание в множество E_-) на n -м периоде регенера-

ции. В этом случае условия на пару (A_n, η_n) выполнены.

Пусть τ'_0 — момент первого наступления события A . Если ν — номер того периода регенерации, на котором первый раз наступило событие A , то $\tau'_0 = t_{\nu-1} + \eta_\nu$, причем ν имеет геометрическое распределение

$$P\{\nu = n\} = p^{n-1}q, \quad p = 1 - q, \text{ а}$$

$$q = P\{A_n\}$$

— вероятность наступления события на данном периоде регенерации.

Нетрудно найти точное распределение величины τ'_0 :

$$a(z) = Me^{-z\tau'_0} = \frac{\varphi_-(z)}{1 - \varphi_+(z)}, \quad (1)$$

где

$$\varphi_-(z) = Me^{-z\eta_1}\chi_1;$$

$$\varphi_+(z) = Me^{-z\xi_1}\bar{\chi}_1;$$

χ_1 — индикатор события A_1 , $\bar{\chi}_1 = 1 - \chi_1$.

Из формулы (1), в частности, получается выражение для среднего времени

$$M\tau'_0 = \frac{T}{q}, \quad T = M(\eta_1\chi_1 + \xi_1\bar{\chi}_1).$$

Представление (1) можно эффективно использовать для нахождения распределения τ'_0 только в тех случаях, когда процесс $x(t)$ на периоде регенерации имеет простую структуру, да и то в этом случае возникает проблема обращения преобразования Лапласа. Поэтому формула (1) в большинстве случаев мало пригодна для эффективной оценки надежности тех восстанавливаемых систем, которые описываются регенерирующим процессом. Однако в

задачах надежности вероятность наступления события A на одном периоде регенерации (т. е. вероятность отказа системы) q мала. Поэтому первый раз это событие наступает в среднем через большое число периодов регенерации. В этих условиях естественно исследовать асимптотическое поведение величины τ'_0 при $q \rightarrow 0$, доказать соответствующие предельные теоремы и, как следствие, получить простые приближенные формулы для характеристик надежности системы.

Предельные теоремы. Пусть $\zeta_1 = \eta_1 \chi_1 + \xi_1 \tilde{\chi}_1$ — укороченный период регенерации и $\alpha = \frac{M\zeta_1^2}{(M\zeta_1)^2} q$. Тогда справедлива теорема 1:

$$\lim_{\alpha \rightarrow 0} P \left\{ \frac{q\tau'_0}{T} > x \right\} = e^{-x}. \quad (2)$$

Из этой теоремы следует, что когда величина α мала, можно использовать приближенную формулу

$$P \{ \tau'_0 > t \} \approx e^{-\frac{qt}{T}} \quad (3)$$

Неудобство теоремы 1 состоит в том, что в ее формулировку входит укороченный период регенерации ζ_1 и его среднее $T = M\zeta_1$, которые трудно вычислить или оценить. Следующая теорема является простым следствием первой и дает более удобную форму для предельного распределения.

Пусть $T_0 = M\xi_1$ и $\alpha_0 = \frac{M\xi_1^2}{(M\xi_1)^2} q$. Тогда справедлива теорема 2:

$$\lim_{\alpha_0 \rightarrow 0} P \left\{ \frac{q\tau'_0}{T_0} > x \right\} = e^{-x}, \quad (4)$$

из которой при малом α_0 следует приближенная формула

$$P \{ \tau'_0 > t \} \approx e^{-\frac{qt}{T_0}} \quad (5)$$

При анализе надежности восстанавливаемых систем часто необходимо знать не только время до первого отказа, но послеотказовые характеристики — величины τ'_k , τ''_k и их средние, вводимые на с. 155. Теперь мы сузим понятие события A , введенного при определении регенерирующего про-

цесса, и под событием A будем понимать переход процесса $\kappa(t)$ из множества E_+ в множество E_- , т. е. на языке надежности отказ системы. Обратный переход из E_- в E_+ , как и выше, будем называть восстановлением. Тогда на одном периоде регенерации событие A может наступить несколько раз. Обозначим через ν_n число наступлений события A на n -м периоде регенерации и введем условное математическое ожидание

$$\beta = M[(\nu_n - 1) | \nu_n \geq 1].$$

Тогда в обозначениях теоремы 2 справедлива теорема 3:

$$\lim_{\substack{\alpha_0 \rightarrow 0 \\ \beta \rightarrow 0}} P \left\{ \frac{q\tau'_k}{T_0} > x \right\} = e^{-x}, \quad (6)$$

причем этот предел равномерен по номеру k .

Из этой теоремы следует, что при малых α_0 и β приближенная формула (5) справедлива для всех периодов безотказной работы системы. Что касается интервалов неисправного состояния системы τ''_k , то для оценки их распределения требуется гораздо больше информации о процессе $\kappa(t)$. Поэтому распределения τ''_k будут оценены ниже для конкретных восстанавливаемых систем.

Примеры. Рассмотрим теперь некоторые простые модели надежности, в которых можно непосредственно использовать приведенные выше предельные теоремы.

A. Ненагруженное дублирование с восстановлением.

Рассмотрим систему, состоящую из двух равнонадежных элементов, из которых один работает, а другой находится в выключенном состоянии. В момент отказа рабочего элемента на его место мгновенно становится и включается в работу резервный элемент, а рабочий элемент поступает на ремонт, который полностью восстанавливает его надежность, и после окончания ремонта становится в резерв. В свою очередь второй элемент, проработав некоторое время, отказывает, на его место становится восстановленный первый элемент, а второй идет в ремонт и после ремонта — в резерв

и т. д. Таким образом, каждый элемент многократно проходит цикл работа—ремонт—резерв. Отказ системы наступает тогда, когда оказываются неисправными оба элемента.

Пусть время безотказной работы каждого элемента не зависит от числа предшествующих ремонтов и имеет распределение

$$P\{\xi < t\} = F(t),$$

а время ремонта каждого элемента имеет распределение

$$P\{\eta < t\} = G(t).$$

Пусть $\xi_0, \xi_1, \xi_2, \dots$ — последовательные интервалы безотказной работы элементов, η_1, η_2, \dots — последовательные времена ремонтов этих элементов. Отказ системы наступает тогда, когда время работы очередного элемента ξ_n оказывается меньше времени другого элемента η_n . Рассмотрим процесс $\kappa(t)$, равный числу неисправных элементов в момент t ,

$$E = \{0, 1, 2\}, \quad E_- = \{2\}.$$

Если отбросить начальный период ξ_0 , на котором не происходит ремонта, то с момента ξ_0 процесс будет регенерирующим, причем моментами регенерации будут те моменты, перед которыми оба элемента исправны и отказывает рабочий элемент.

Используя формулу (1), нетрудно получить в терминах преобразования Лапласа распределение времени до первого отказа:

$$M e^{-z\tau'_0} = \varphi(z) \frac{\varphi_-(z)}{1 - \varphi_+(z)}, \quad (7)$$

где

$$\varphi(z) = \int_0^{\infty} e^{-zx} dF(x);$$

$$\varphi_-(z) = \int_0^{\infty} e^{-zx} \bar{G}(x) dF(x),$$

$$\bar{G}(x) = 1 - G(x);$$

$$\varphi_+(z) = \varphi(z) - \varphi_-(z).$$

Однако представление (7) неудобно для использования, так как надо еще обращать преобразование Лапласа.

Однако, если вероятность отказа на одном периоде регенерации $q =$

$$= \int_0^{\infty} \bar{G}(x) dF(x) \text{ мала, то, используя}$$

теорему 2 и производя некоторые дополнительные оценки, можно получить простую приближенную формулу

$$P\{\tau'_0 > t\} \approx e^{-\frac{f(0) T_1}{T} t} \quad (8)$$

где $T = \int_0^{\infty} \bar{F}(t) dt$ — среднее время

безотказной работы элемента, $T_1 =$

$$= \int_0^{\infty} \bar{G}(t) dt \text{ — среднее время ремонта,}$$

и дополнительно к условиям теоремы 2 предполагается, что плотность $f(x) = F'(x)$ непрерывна в нуле и $f(0) > 0$.

Б. Облегченное дублирование с профилактиками.

Предположим, что система состоит из двух элементов — рабочего с интенсивностью отказа λ и резервного с интенсивностью отказа $\lambda_1 < \lambda$. В момент отказа рабочего элемента на его место автоматически подключается резервный элемент и его интенсивность отказа возрастает до величины λ . Предположим, что через время h , начиная с очередной профилактики, назначается следующая профилактика, во время которой мгновенно все отказавшие элементы заменяются новыми.

Если в течение периода h отказы в системе не обнаруживаются, то очередная профилактика проводится, как и назначено, через время h . Если же отказ обнаруживается, то в этот момент проводится профилактика, а следующая опять назначается через время h . Предположим далее, что отказ рабочего элемента обнаруживается сразу, а отказ резервного элемента мы не наблюдаем, он восстанавливается либо в момент запланированной профилактики, либо тогда, когда отказал рабочий элемент, и мы, проводя досрочную профилактику, обнаруживаем, что отказал и резервный элемент. Отказ системы происходит в тот момент, когда до начала запланирован-

ной профилактики сначала отказывает резервный элемент (его отказ мы не обнаруживаем), а затем отказывает рабочий элемент. Процесс $x(t)$, как и в первом примере, есть число неисправных в момент t элементов. Для него моменты проведения профилактик будут моментами регенерации. Используя теорему 1, легко вывести следующую приближенную формулу: если величина $(\lambda + \lambda_1)h$ мала (а это означает, что мала вероятность отказа на периоде между соседними профилактиками хотя бы одного элемента), то

$$P\{\tau'_0 > t\} \approx e^{-\frac{\lambda\lambda_1 h}{2} t} \quad (9)$$

2. РЕГЕНЕРИРУЮЩИЙ ПРОЦЕСС СПЕЦИАЛЬНОГО ТИПА. МАРКОВСКИЕ ВВЕРХ ПРОЦЕССЫ

Регенерирующий процесс специального типа. Рассмотрим один класс регенерирующих процессов, который очень часто возникает в моделях восстанавливаемых систем и в теории массового обслуживания. Предположим, что каждый n -й период регенерации ξ_n состоит из двух частей, длины которых и траектории процесса $x(t)$ на них независимы

$$\xi_n = \xi'_n + \xi''_n.$$

Первая часть ξ'_n имеет показательное распределение с параметром λ , и в ней событие A_n произойти не может, а вторая часть имеет произвольное распределение со средним $T = M\xi''_n$. В теории массового обслуживания такой процесс возникает в тех случаях, когда в систему обслуживания поступает пуассоновский поток. Если под процессом $x(t)$ понимать число требований, находящихся в системе в момент t , то моментами регенерации будет момент освобождения системы от требований, а период регенерации будет состоять из двух частей. Первая часть — это интервал, на котором $x(t) = 0$, его называют свободным периодом. Вторая часть, которую называют периодом занятости, — это интервал, на котором в системе есть хотя бы одно

требование. Ниже используется эта терминология. В моделях восстанавливаемых систем такой процесс возникает в том случае, когда интенсивность отказа каждого i -го элемента системы зависит только от состояния других элементов и равна $\lambda_i(e)$. В этом случае первая часть периода регенерации — это интервал, где все элементы системы исправны, а вторая часть — это интервал, на котором происходят отказы и восстановление элементов. Для регенерирующего процесса специального типа удастся не только доказать предельные теоремы, но и найти весьма точные двусторонние неравенства для некоторых характеристик надежности.

Теорема 4: справедливо неравенство

$$e^{-\lambda qt} \leq P\{\tau'_0 > t\} \leq e^{-\lambda qt} + \lambda T. \quad (10)$$

Из него, в частности, следует, что при $\lambda T \rightarrow 0$ справедлива приближенная формула

$$P\{\tau'_0 > t\} \approx e^{-\lambda qt}. \quad (11)$$

Здесь q , как и выше, — вероятность наступления события на одном периоде занятости. Приближенная формула (11) одновременно является и пессимистической оценкой надежности, и поэтому ее можно использовать без всякого риска.

Теорема 4 следует из одного более общего утверждения. Пусть ξ и η — неотрицательные и, вообще говоря, зависимые случайные величины

$$P\{\xi > t\} = e^{-\lambda t} \quad \text{Обозначим}$$

$$P\{\eta > x \mid \xi = t\} = F(x, t).$$

Теорема 5: если $F(x, t)$ при любом x монотонно возрастает по t , то

$$e^{-\lambda t} \leq P\{\xi + \eta > t\} \leq e^{-\lambda t} + M(1 - e^{-\lambda \eta}). \quad (12)$$

Если в качестве ξ взять сумму длин всех свободных периодов до наступления события A и применить неравенство (12), то получим неравенство (11).

Марковский вверх процесс. Практическим недостатком теорем 1—4 является то, что в их формулировке входит вероятность q , которая в большинстве случаев не вычисляется в явном

виде. Поэтому сейчас мы введем один содержательный подкласс регенерирующих процессов специального типа, для которых эта вероятность может быть оценена через более простые характеристики процесса.

Рассмотрим процесс с конечным или счетным числом состояний $E = \{0, 1, 2, 3, \dots\}$ и предположим, что это множество состояний частично упорядочено. Это отношение упорядоченности будем обозначать $i < j$, и его не следует смешивать с упорядоченностью вещественных чисел. Предположим, что состояние «0» является единственным минимальным состоянием.

Предположим далее, что множество E разбито на две части, $E = E_+ + E_-$, и выполнены условия: 1) $0 \in E_+$; 2) $|E_+| < \infty$ (множество E_+ — конечно); 3) разбиение E монотонно, т. е. из того, что $i < j$ и $i \in E_-$, следует, что $j \in E_-$.

Рассмотрим процесс $x(t)$ с множеством состояний E . Интервалы, где $x(t) = 0$, будем называть свободными периодами, а интервалы, где $x(t) > 0$ — периодами занятости. Предположим, что процесс $x(t)$ удовлетворяет следующим условиям:

1) мгновенные скачки из состояния i могут происходить только в состоянии j , сравнимые с i , когда либо $j < i$, либо $j < i$;

2) для любых i и j , $i < j$ и любых $t_1 < t_2 < \dots < t_n < t$ условная вероятность

$$P\{x(t+h) = j \mid x(t) = i, x(t_1) = i_1, \dots, x(t_n) = i_n\} = P\{x(t+h) = j \mid x(t) = i\} = \lambda_{ij}h + o(h),$$

т. е. скачки вверх носят марковский характер;

3) если $i > j$, то вероятность скачка за время h из состояния i в состояние j зависит только от траектории процесса $x(t)$ на том периоде занятости, на котором происходит этот скачок;

4) $x(0) = 0$, и для любого $i \in E_+$

$$0 < \sum_{j>i} \lambda_{ij} = \lambda_i < \infty.$$

Такой процесс назовем марковским вверх процессом. Он является регене-

рирующим процессом специального типа, и для него справедливы теоремы 1—4. Поясним, какое отношение к моделям восстанавливаемых систем имеет такой процесс.

Если состоянием системы $x(t)$ считать вектор, задающий состояние ее элементов, и если интенсивность отказа каждого элемента зависит только от состояния других элементов, то процесс $x(t)$ будет марковским вверх процессом. При состоянии «0» это состояние системы, в которой все элементы исправны, а упорядоченность $i < j$ означает, что из состояния i можно перейти в состояние j при отказе одного или нескольких элементов (при неизменном состоянии других элементов).

Предположим, что процесс $x(t)$ эргодический (а для этого достаточно, чтобы период занятости имел конечное среднее), т. е. существуют стационарные вероятности

$$p_i = \lim_{t \rightarrow \infty} P\{x(t) = i\}.$$

Пусть событие A есть попадание процесса в множество E_- . Назовем путем π последовательность проходящих процессом состояний от начала периода занятости и до момента наступления события на этом периоде

$$\pi = \{0, i_1, i_2, \dots, i_m\}$$

$$i_k \in E_+ \text{ при } k < m, \quad i_m \in E_-.$$

Пусть $\Pi = \{\pi\}$ — множество всех возможных путей. Тогда $q = \sum_{\pi \in \Pi} q(\pi)$,

где $q(\pi)$ — вероятность того, что на периоде занятости произойдет событие A и при этом процесс пройдет путь π .

Назовем путь π монотонным, если

$$0 < i_1 < i_2 < \dots < i_m,$$

а множество всех монотонных путей обозначим Π_0 . Тогда $q_0 = \sum_{\pi \in \Pi_0} q(\pi)$

есть вероятность наступления события по монотонному пути. Эта вероятность в конкретных задачах легко вычисляется, поскольку число монотонных путей конечно.

Справедлива теорема 6: для марковского вверх процесса вероятность q удовлетворяет неравенствам

$$q_0 \leq q \leq \frac{1}{\lambda_0 \rho_0} \sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij} = \bar{q}. \quad (13)$$

В конкретных моделях восстанавливаемых систем при некоторых естественных условиях левая и правая части этого неравенства эквивалентны, что дает асимптотически точную оценку вероятности q . В тех случаях, когда стационарные вероятности p_i не вычисляются в явном виде, их, в свою очередь, можно оценить сверху и получить эффективную оценку q сверху.

Приведем еще одно утверждение, которое позволяет для марковского вверх процесса находить введенные выше средние характеристики $T_1 = M\tau'$ и $T_2 = M\tau''$ — среднюю наработку на отказ и среднее время требования системы в неисправном состоянии в стационарном режиме.

Теорема 7: величины T_1 и T_2 выражаются через стационарные вероятности

$$T_1 = \frac{\sum_{i \in E_+} p_i}{\sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij}},$$

$$T_2 = \frac{\sum_{i \in E_-} p_i}{\sum_{j \in E_-} \sum_{i \in E_+} p_i \lambda_{ij}}. \quad (14)$$

3. СХОДИМОСТЬ СЛУЧАЙНЫХ ВЕЛИЧИН С РАЦИОНАЛЬНЫМ ПРЕОБРАЗОВАНИЕМ ЛАПЛАСА К ПОКАЗАТЕЛЬНО РАСПРЕДЕЛЕННОЙ ВЕЛИЧИНЕ

1. Выше было показано, как вычислять основные характеристики однородного марковского процесса с конечным числом состояний. Из приведенных там формул следует, что такие важные характеристики, как времена безотказной работы системы и времена

пребывания в неисправном состоянии, имеют рациональное преобразование Лапласа. Если число состояний марковского процесса велико, обращение рациональной дроби практически труднореализуемо. Однако, если марковский процесс попадает в отказовое множество E_- редко, то согласно теории вероятности время безотказной работы системы приближенно должно иметь показательное распределение. Поэтому естественно найти по возможности простые и эффективные условия, при которых распределение с рациональным преобразованием Лапласа будет близко к показательному распределению. Приведенные ниже утверждения дают такие условия.

Теорема 8. Если ξ — неотрицательная случайная величина со средним $M\xi = 1$ и ее преобразование Лапласа имеет вид

$$M e^{-z\xi} = \frac{1}{(1 + \alpha_1 z)(1 + \alpha_2 z) \dots (1 + \alpha_n z)} = \frac{1}{1 + z + a_2 z^2 + \dots + a_n z^n},$$

$\alpha_k > 0,$

то при любом $x \geq 0$ справедливо неравенство

$$|P\{\xi > x\} - e^{-x}| \leq \frac{1 - \sqrt{1 - 4a_2}}{1 + \sqrt{1 - 4a_2}}. \quad (15)$$

Замечание. Если $a_2 \rightarrow 0$, то можно показать, что

$$\sup_{x \geq 0} |P\{\xi > x\} - e^{-x}| \sim a_2.$$

и, таким образом, неравенство (15) асимптотически неутончено. Величина $a_2 = 1 - \frac{1}{2} M\xi^2$, и поэтому приведенное неравенство позволяет очень точно оценивать скорость сходимости к показательному распределению через первые два момента величины,

Теорема 9. Если ξ — неотрицательная случайная величина со средним

$M\xi = 1$ и ее преобразование Лапласа имеет вид

$$Me^{-z\xi} = \frac{1}{1 + z + a_2 z^2 + \dots + a_n z^n},$$

то условие $a_2 \rightarrow 0$ необходимо и достаточно для сходимости

$$P\{\xi > x\} \rightarrow e^{-x}. \tag{16}$$

Замечание. Используя метод Эссена, можно показать, что

$$\sup_{x>0} |P\{\xi > x\} - e^{-x}| \leq C\sqrt{a_2}$$

и $C < 10$. (17)

Теорема 10. Если ξ — неотрицательная случайная величина со средним $M\xi = 1$ и ее преобразование Лапласа рационально

$$Me^{-\xi z} = \frac{1 + b_1 z + b_2 z^2 + \dots + b_m z^m}{1 + a_1 z + a_2 z^2 + \dots + a_n z^n},$$

$$a_1 - b_1 = 1$$

и степень знаменателя n ограничена, то из условий $b_1 \rightarrow 0$ и $a_2 \rightarrow 0$ следует сходимость

$$P\{\xi > x\} \rightarrow e^{-x}. \tag{18}$$

Процесс рождения и смерти. Теперь на примере марковского процесса рождения и смерти, который имеет широкие приложения в теории массового обслуживания и теории надежности, покажем, как применяются сформулированные выше теоремы. Но сначала определим этот процесс и найдем точные выражения для его характеристик.

Процессом рождения и смерти называют марковский однородный процесс с конечным или счетным числом состояний $0, 1, 2, \dots, k, \dots$, в котором мгновенные переходы возможны только в соседние состояния, т. е.

$$\lambda_{i, i+1} = \lambda_i > 0, \quad \lambda_{i, i-1} = \mu_i > 0$$

и $\lambda_{ij} = 0$, если $|j - i| > 1$.

Кроме того, $\mu_0 = 0$, и если число состояний конечно и равно $N + 1$, то $\lambda_N = 0$. Для такого процесса урав-

нения для вероятностей состояний $p_k(t)$ имеют вид

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t); \\ p'_k(t) = \lambda_{k-1} p_{k-1}(t) - \\ - (\lambda_k + \mu_k) p_k(t); \\ k = 1, 2, \dots, N-1 \\ p'_N(t) = \lambda_{N-1} p_{N-1}(t) - \mu_N p_N(t). \end{cases} \tag{19}$$

Если число состояний бесконечно, то последнего уравнения не будет.

Чтобы процесс рождения и смерти имел стационарное распределение, необходимо и достаточно, чтобы выполнялись условия

$$\sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{k-1}}{\mu_1 \mu_2 \dots \mu_k} < \infty;$$

$$\sum_{k=1}^{\infty} \frac{\mu_1 \mu_2 \dots \mu_k}{\lambda_1 \lambda_2 \dots \lambda_k} = \infty. \tag{20}$$

В этом случае стационарные вероятности имеют вид

$$p_k = \frac{\theta_k}{\sum_{i=0}^{\infty} \theta_i}, \quad \theta_0 = 1,$$

$$\theta_k = \frac{\lambda_0 \lambda_1 \dots \lambda_{k-1}}{\mu_1 \mu_2 \dots \mu_k} \tag{21}$$

Введем случайные величины τ_{ij} — время перехода процесса из состояния i в состояние j .

Пусть, как всегда, множество состояний процесса разбито на два подмножества:

$E_+ = \{0, 1, 2, \dots, n\}$ — исправные состояния;

$E_- = \{n + 1, n + 2, \dots, N\}$ — неисправные состояния.

(Ниже рассмотрен случай конечного числа состояний, что всегда имеет место в теории надежности.) Тогда основные характеристики надежности выражаются так:

$$\begin{aligned} \tau'_0 &= \tau_{0, n+1}, & \tau'_k &= \tau_{n, n+1}, \\ \tau''_k &= \tau_{n+1, n}. \end{aligned} \tag{22}$$

Поскольку распределения τ'_k и τ''_k не зависят от номера K , те же представления имеют место и для стационарных характеристик:

$$\tau' = \tau_{n, n+1}; \quad \tau'' = \tau_{n+1, n}.$$

Таким образом, нахождение основных характеристик надежности сводится к нахождению распределений величин τ_{ij} . Распределение величины $\tau_{0, n+1}$ имеет вид

$$Me^{-z\tau_{0, n+1}} = Me^{-z\tau'_0} = \frac{1}{\Delta_{n+1}(z)}, \quad (23)$$

где многочлены $\Delta_k(z)$ определяются рекуррентным уравнением

$$\Delta_{k+1}(z) = \left(1 + \frac{z + \mu_k}{\lambda_k}\right) \Delta_k(z) - \frac{\mu_k}{\lambda_k} \Delta_{k-1}(z), \quad (24)$$

$$\Delta_0(z) = 1, \quad \Delta_1(z) = 1 + \frac{z}{\lambda_0}.$$

Эти многочлены удовлетворяют следующим свойствам:

1. Все корни $\Delta_k(z)$ различны и отрицательны.

2. Корни соседних многочленов чередуются, т. е. между двумя корнями $\Delta_{k+1}(z)$ лежит ровно один корень $\Delta_k(z)$.

3. Модуль максимального по абсолютной величине корня многочлена $\Delta_k(z)$ не превосходит $A_k = \sum_{i=0}^k (\lambda_i + \mu_i)$.

Эти свойства позволяют быстро и просто вычислять корни многочленов $\Delta_k(z)$.

Интервалы безотказной работы системы $\tau'_k = \tau_{n, n+1}$ имеют распределение

$$Me^{-z, \tau_{n, n+1}} = \frac{\Delta_n(z)}{\Delta_{n+1}(z)}. \quad (25)$$

Без труда находятся и средние характеристики:

$$T_0 = M\tau'_0 = \sum_{k=0}^n \frac{1}{\lambda_k \rho_k} \sum_{s=0}^k \rho_s;$$

$$M\tau'_k = \frac{1}{\lambda_n \rho_n} \sum_{s=0}^n \rho_s = T_1; \quad (26)$$

$$M\tau''_k = \frac{1}{\lambda_n \rho_n} \sum_{s=n+1}^N \rho_s = T_2.$$

Используя приведенные выше теоремы 8—10, найдем условия, при которых величины τ'_0 и τ'_k имеют асимптотически показательное распределение. Пусть

$$\Delta_k(z) = 1 + \Delta_{k1}z + \Delta_{k2}z^2 + \dots + \Delta_{kk}z^k,$$

где

$$\Delta_{k1} = \sum_{i=0}^{k-1} \frac{1}{\lambda_i \rho_i} \sum_{s=0}^i \rho_s; \quad (27)$$

$$\Delta_{k2} = \sum_{i=1}^{k-1} \frac{1}{\lambda_i \rho_i} \sum_{s=1}^i \rho_s \Delta_{s1}.$$

Тогда, как следует из теоремы 8, распределение τ'_0 удовлетворяет неравенству

$$\sup_{t>0} \left| P\{\tau'_0 > t\} - e^{-\frac{t}{T_0}} \right| \leq \leq \frac{1 - \sqrt{1 - 4a_2}}{1 + \sqrt{1 - 4a_2}}, \quad (28)$$

где

$$a_2 = \frac{\Delta_{n+1, 2}}{\Delta_{n+1, 1}^2}.$$

Для наработки на отказ τ'_k справедливо утверждение: если $\frac{\Delta_{n+1, 2}}{\Delta_{n+1, 1}^2} \rightarrow 0$

и $\frac{\Delta_{n1}}{\Delta_{n+1, 1}} \rightarrow 0$, то, как следует из теоремы 10.

$$\lim P\left\{\frac{\tau'_k}{T_1} > x\right\} = e^{-x}. \quad (29)$$

Таким образом, теоремы 8—10 позволяют эффективно доказывать сходимость (и даже оценивать скорость

сходимости) распределения времени безотказной работы системы к показательному распределению.

Марковские модели резервирования с восстановлением. Рассмотрим теперь некоторые модели теории надежности, которые описываются процессом рождения и смерти. Пусть система состоит из $(n + 1)$ элемента, из которых один — рабочий элемент, а другие — резервные. В момент отказа рабочего или резервного элемента он поступает в ремонтное устройство, состоящее из r ремонтных единиц; при этом при отказе рабочего элемента на его место мгновенно становится резервный элемент. Предположим, что интенсивность отказа рабочего элемента не зависит от прошлого и равна λ , а интенсивность отказа резервного элемента также не зависит от прошлого и равна λ' . При поступлении резервного элемента на рабочее место его интенсивность отказа становится равной λ . Предположим, что время ремонта имеет показательное распределение с параметром μ . Если оказываются одновременно неисправными более r элементов, то r из них ремонтируются, а остальные ждут своей очереди. После восстановления элементы поступают в резерв. Отказ системы наступает тогда, когда отказывают все $(n + 1)$ элементов. Процесс $x(t)$, равный числу неисправных элементов в момент t , будет процессом рождения и смерти с интенсивностями

$$\lambda_k = \lambda + (n - k)\lambda' \quad \text{при } k \leq n,$$

$$\lambda_{n+1} = 0;$$

$$\mu_k = k\mu, \quad \text{если } k \leq r, \quad \text{и } \mu_k = r\mu, \\ \text{если } k > r.$$

Такую модель называют марковской моделью облегченного резервирования с восстановлением. При $\lambda' = 0$ получаем модель ненагруженного резервирования с восстановлением; при $\lambda' = \lambda$ — модель нагруженного резервирования с восстановлением. Можно рассматривать и другие модели резервирования с восстановлением, например, модель смешанного резервирования, где есть резервные элементы, находящиеся в ненагруженном, в облегченном и в нагруженном резерве.

Есть модель скользящего резервирования, в которой резервные элементы могут заменять любой отказавший элемент из некоторой группы последовательно соединенных элементов. Наконец, интенсивность ремонта для каждой ремонтной единицы может зависеть от числа неисправных в этот момент элементов. Все эти модели описываются процессом рождения и смерти, и для их анализа можно использовать все точные формулы и оценки, которые получены выше.

4. ОБЩАЯ МОДЕЛЬ РЕЗЕРВИРОВАНИЯ С ВОССТАНОВЛЕНИЕМ

Определение. Выше были рассмотрены марковские модели резервирования с восстановлением, когда поведение системы описывается процессом рождения и смерти. При этом были сделаны два предположения: суммарная интенсивность отказа элементов системы зависит только от числа неисправных элементов, а время ремонта каждого отказавшего элемента имеет показательное распределение. Первое предположение физически оправдано, поскольку каждый элемент сам является устройством, состоящим из нескольких, а часто и большого числа элементов. Отказ такого устройства есть отказ одного из его элементов, который при восстановлении заменяется новым. Поэтому, если рассматривать чистое время работы устройства, его поток отказов будет суммой большого числа потоков восстановления, и согласно известной предельной теореме Ханчина он асимптотически будет пуассоновским потоком. Отсюда интервал между соседними отказами будет иметь показательное распределение. Если дополнительно отказы других устройств меняют нагрузку (режим), приходящуюся на наше устройство, то это равносильно сжатию или растяжению соответствующего интервала времени. И поэтому при неизменном состоянии других устройств интенсивность отказа нашего устройства будет постоянна и соответственно будет постоянной суммарная интенсивность отказа всей системы.

Что касается второго предположения о том, что время ремонта элементов имеет показательное распределение, но оно, как правило, не выполняется со сколько-нибудь хорошим приближением. Действительно, время ремонта складывается из времени, в течение которого проводятся вспомогательные операции, из времени, когда производится поиск неисправности, и из времени восстановления. Поэтому ремонт не может закончиться за время меньше некоторой величины. Отсюда следует, что на начальном участке плотность распределения случайного времени ремонта равна нулю, в то время как для показательного распределения плотность в окрестности нуля максимальна. По этой причине для того, чтобы наша математическая модель была более или менее адекватна реальной восстанавливаемой системе, необходимо предполагать, что время ремонта имеет произвольное распределение. В этом случае получаем следующую модель.

Пусть система состоит из $(n + 1)$ элементов. Если в данный момент k элементов неисправно, то суммарная интенсивность отказа элементов не зависит от прошлого поведения процесса и равна λ_k . Каждый элемент при отказе мгновенно поступает в ремонтное устройство, состоящее из r ремонтных единиц. Каждая единица может одновременно восстанавливать один элемент. Приходящие в ремонтное устройство элементы восстанавливаются или ожидают очереди на восстановление, а после окончания ремонта мгновенно возвращаются в систему. Время восстановления каждого элемента есть случайная величина η с функцией распределения $G(x)$, и разные времена ремонта независимы. Отказ системы наступает тогда, когда оказываются неисправными все $(n + 1)$ элементов. Такая модель включает в себя все известные типы резерва, в том числе и скользящий резерв, для которого надо последовательно соединенную группу работающих элементов считать одним элементом (при этом надо, конечно, предположить, что при отказе системы выключаются до восстановления системы рабочие элементы).

Описанную модель обозначим (λ_k, G, r, n) , где первый символ показывает, какие суммарные интенсивности отказа, второй символ — как распределено время восстановления, третий символ есть число ремонтных, а четвертый — число резервных элементов. Процесс $x(t)$, который описывает такую систему, есть число неисправных в момент t элементов. Легко видеть, что процесс $x(t)$ является регенерирующим процессом, в котором моменты регенерации суть моменты попадания процесса в состояние «0». Более того, этот процесс является регенерирующим процессом специального типа, у которого свободный период, где $x(t) \equiv 0$, имеет показательное распределение с параметром λ_0 , а период занятости, на котором $x(t) > 0$, имеет сложное распределение, которое, вообще говоря, не находится в явном виде. Далее, поскольку интенсивность скачка процесса $x(t)$ из состояния k в $k + 1$ равна λ_k и не зависит от предшествующей траектории процесса, наш процесс является марковским вверх процессом. Поэтому для него справедливы все результаты теорем 1—7.

Сформулируем ряд результатов для общей модели резервирования с восстановлением. Обозначим: $\bar{\lambda} = \max \lambda_k$; $T = m_1 = M\eta$ — среднее время восстановления; $m_k = \int_0^{\infty} t^k dG(t)$ — k -й момент времени

восстановления; $\alpha_n = \frac{m_{n+1}}{m_1^n}$;

$$J_{nr} = \int_0^{\infty} \frac{x^{n-r}}{(n-r)!} \left[\int_x^{\infty} \bar{G}(t) dt \right]^{r-1} \times \frac{\bar{G}(x)}{(r-1)!} dx;$$

$\tau'_0, \tau'_k, \tau''_k$, как и выше, обозначают соответствующие периоды исправного и неисправного состояния системы; q — вероятность отказа системы на одном периоде занятости.

Теорема 11. Если в системе (λ_k, G, r, n) $\bar{\lambda}T \rightarrow 0$, то для всех $k = 0, 1, 2$,

$$\lim P \{ \lambda_0 q \tau'_0 > x \} = e^{-x}. \quad (30)$$

Теорема 12. Если в системе (λ_k, G, r, n) параметры λ_k, r, n фиксированы, а распределение $G(t)$ меняется так, что $\alpha_n \rightarrow 0$, то при любом $k = 0, 1, 2$,

$$\lim P \{ \Lambda \tau'_k > x \} = e^{-x}, \quad (31)$$

причем предел равномерен по номеру k . Здесь $\Lambda = \lambda_0 \lambda_1 \dots \lambda_n J_{nr}$.

Поскольку эти пределы равномерны по номеру k , обе теоремы справедливы и для стационарной наработки на отказ τ'

Из теоремы 12 следует, что при малом α_n справедливо приближенное равенство

$$P \{ \tau'_k > t \} \approx e^{-\Lambda t}. \quad (32)$$

Отметим частные случаи.

1. $r = n$, т. е. в исправной системе все отказавшие элементы восстанавливаются без ожидания. В этом случае

$$\Lambda = \frac{\lambda_0 \lambda_1 \dots \lambda_n}{n!} T^n. \quad (33)$$

2. $r = 1$ — имеется одна ремонтная единица.

Тогда

$$\Lambda = \frac{\lambda_0 \lambda_1 \dots \lambda_n}{n!} m_n. \quad (34)$$

Из последнего равенства видно, какую ошибку дает предположение о показательном распределении времени ремонта. Для показательного распределения момент $m_n = n! T^n$. Если же на самом деле время ремонта постоянно, то $m_n \approx T^n$, и, используя показательное распределение, мы увеличиваем интенсивность отказа системы в $n!$ раз.

Для периодов неисправного состояния τ''_k при более сильных условиях справедлива теорема 13: пусть для системы (λ_k, G, r, n) распределение $G(t)$ фиксировано, существует момент $m_n < \infty$, а интенсивности λ_k имеют вид $\lambda_k = \lambda_k^{(0)} \epsilon$, где $\lambda_k^{(0)}$ фиксировано, а $\epsilon \rightarrow 0$. Тогда

$$\lim_{\epsilon \rightarrow 0} P \{ \tau''_k > x \} = \frac{J_{nr}(x)}{J_{nr}}, \quad (35)$$

где

$$J_{nr}(x) = \int_0^x \frac{(t-x)^{n-r}}{(n-r)!} \times \\ \times \left[\int_t^\infty \bar{G}(u) du \right]^{r-1} \frac{\bar{G}(t)}{(r-1)!} dt.$$

Двусторонние неравенства для характеристик надежности. Предельные теоремы, приведенные выше, имеют тот недостаток, что из них не следует, когда можно применять соответствующее приближенное равенство с заданной точностью. Поэтому ниже приведем асимптотически точные двусторонние неравенства для характеристик надежности, из которых, в частности, следуют соответствующие предельные теоремы.

Теорема 14. Для системы (λ_k, G, r, n) справедливо неравенство

$$- \lambda_0 q t \leq P \{ \tau'_0 > t \} \leq e^{\lambda_0 q t} + \\ + \frac{\bar{\lambda} T}{1 - \bar{\lambda} T}. \quad (36)$$

Из него, в частности, следует утверждение теоремы 11.

Для того чтобы эффективно использовать это неравенство, необходимо найти двусторонние оценки для вероятности q . Здесь асимптотически точные оценки удается найти только для крайних случаев $r = 1$ и $r = n$.

Теорема 15. Для системы $(\lambda_k, G, 1, n)$ вероятность q удовлетворяет неравенству

$$b_{1n} \leq q \leq b_{1n} A_{n-1}(\gamma), \quad (37)$$

где $b_{1n} = \int_0^\infty p_{1n}(t) dG(t)$, $p_{1n}(t)$ есть вероятность перехода процесса чистой гибели с параметрами λ_k за время t

из состояния 1 в состояние $(n + 1)$;

$$A_{n-1}(\gamma) = \sum_{k=0}^{\infty} (k + 1)^{n-1} \gamma^k,$$

$$\gamma = \bar{\lambda} \int_0^{\infty} e^{-\lambda t} \bar{G}(t) dt$$

$$\bar{\lambda} = \max_{k \leq n} \lambda_k, \quad \lambda = \min_{k \leq n} \lambda_k.$$

Величину b_{1n} , которая представляет собой вероятность отказа системы за время ремонта первого элемента, поступившего в начале периода занятости, в свою очередь можно оценить.

Если существует момент m_{n+1} , то

$$\lambda_1 \lambda_n \left(\frac{m_n}{n!} - \bar{\lambda} \frac{m_{n+1}}{n!} \right) \leq b_{1n} \leq \leq \lambda_1 \lambda_n \frac{m_n}{n!}. \quad (38)$$

Величина γ также оценивается сверху $\gamma \leq \bar{\lambda} T$

Если $\frac{m_{n+1}}{m_n} \rightarrow 0$, то неравенства (37) и (38) дают асимптотически точные оценки для q , выраженные через моменты распределения $G(t)$.

Для системы $(\lambda, G, 1, n)$, у которой все $\lambda_k = \lambda$ (случай ненагруженного резерва), неравенство (37) имеет вид

$$b_{n-1} \leq q \leq b_{n-1} A_{n-1}(\gamma), \quad (39)$$

где

$$b_{n-1} = \int_0^{\infty} \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} \bar{G}(x) dx.$$

В свою очередь b_{n-1} оценивается неравенством (38).

Теорема 16. Для системы (λ_k, G, n, n) вероятность q удовлетворяет неравенству

$$\lambda_1 \lambda_2 \dots \lambda_n \bar{\lambda} \int_0^{\infty} e^{-\lambda t} \frac{\psi^n(t)}{n!} dt \leq \leq q \leq \frac{\lambda_1 \lambda_2 \dots \lambda_n}{n!} T^n, \quad (40)$$

где $\psi(t) = \int_0^t \bar{G}(x) dx.$

Если существует момент $m_2 < \infty$, то справедливо более простое неравенство

$$\frac{\lambda_1 \lambda_2 \dots \lambda_n}{n!} T^n \left(1 - \frac{n \bar{\lambda} m_2}{2T} \right) \leq q \leq \leq \frac{\lambda_1 \lambda_2 \dots \lambda_n}{n!} T^n. \quad (41)$$

5. МОДЕЛЬ СЛОЖНОЙ ВОССТАНАВЛИВАЕМОЙ СИСТЕМЫ

Описание системы. Рассмотрим систему, состоящую из n элементов. Поведение этой системы во времени будем задавать двоичным вектором $e(t) = [e_1(t), \dots, e_n(t)]$, где $e_k(t) = 0$, если в момент t k -й элемент исправен, и $e_k(t) = 1$, если в этот момент он неисправен. Предположим, что состояние элементов системы однозначно определяет состояние самой системы. Тогда все множество состояний $E = \{e\}$ разбивается на два подмножества $E = E_+ + E_-$; E_+ — исправные состояния системы, E_- — неисправные состояния системы. Как и выше, предполагается, что это разбиение монотонно. Предположим далее, что интенсивность отказа каждого k -го элемента зависит только от состояния других элементов в данный момент; обозначим ее $\lambda_k(e)$. Если k -й элемент в состоянии $e = (e_1, \dots, e_n)$ неисправен, т. е. $e_k = 1$, то полагаем $\lambda_k(e) =$

$$= 0. \text{ Обозначим } \lambda(e) = \sum_{k=1}^n \lambda_k(e). \text{ Каж-}$$

дый отказавший элемент мгновенно поступает в ремонтное устройство, состоящее из ремонтных единиц. Каждая ремонтная единица может одновременно ремонтировать только один элемент. Каждому элементу соответствует некоторое непустое множество ремонтных единиц, которые могут его восстанавливать. Если хотя бы одна ремонтная единица, доступная для элемента, свободна, элемент сразу же начинает ремонтироваться. Предположим, что начавшийся ремонт не прерывается. После окончания ремонта элемент сразу же возвращается на свое место в системе. Если все ремонтные единицы, доступные для

элемента, заняты, он ожидает в очереди. При этом дисциплина выбора элемента из очереди ремонтной единицы, которая закончила ремонт, может быть произвольной.

Построенный таким образом случайный процесс $e(t)$ описывает поведение сложной восстанавливаемой системы. Этот процесс является, очевидно, регенерирующим процессом специального типа, у которого свободный период, где $e(t) \equiv 0 = (0, 0, \dots, 0)$, имеет показательное распределение с параметром $\lambda(0)$. Более того, этот процесс является марковским вверх процессом, и поэтому к нему применимы все предельные теоремы и неравенства, выведенные для этого случая.

Предельные теоремы. Пусть $\lambda = \max_{e \in E_+} \lambda(e)$, $G_{ij}(t)$ — функция рас-

пределения времени ремонта i -го элемента, поступившего на j -ю единицу (учитывается, что элементы — разные и ремонтные единицы (рабочие) могут иметь разную квалификацию).

Обозначим

$$G(t) = \min G_{ij}(t); \quad T = \int_0^\infty t dG(t);$$

q — вероятность отказа системы на одном периоде занятости.

Теорема 17.

$$\lim_{\lambda T \rightarrow 0} P \{ \lambda(0) q \tau'_0 > x \} = e^{-x}. \quad (42)$$

Для того чтобы с помощью этой теоремы эффективно оценивать надежность, нужно оценить еще вероятность q . Для этого введем некоторые понятия и обозначения.

Назовем путем π последовательность проходимых процессом $e(t)$ состояний от начала периода занятости и до момента отказа на этом периоде:

$$\pi = \{ 0, e^{(1)}, e^{(2)}, \dots, e^{(m)} \},$$

$$e^{(k)} \in E_+ \text{ при } k < m, \quad e^{(m)} \in E_-.$$

Множество всех путей обозначим через Π , а число $m = m(\pi)$ назовем длиной пути. Путь π назовем монотонным, если $0 < e^{(1)} < e^{(2)} < \dots < e^{(m)}$, причем под упорядоченностью векторов понимается покомпонентная

упорядоченность. Каждому монотонному пути соответствует последовательность номеров i_1, i_2, \dots, i_m , где i_k — номер элемента, отказывающего при переходе из состояния $e^{(k-1)}$ в $e^{(k)}$, $k = 1, \dots, m$. Монотонный путь назовем допустимым, если $\lambda_{i_1}(0) \lambda_{i_2}(e^{(1)}) \dots \lambda_{i_m}(e^{(m-1)}) > 0$.

Пусть, наконец, $m_0 = \min_{\pi} m(\pi)$, где минимум берется по всем монотонным допустимым путям. Монотонный допустимый путь назовем минимальным, если его длина $m(\pi) = m_0$. Множество монотонных минимальных допустимых путей обозначим Π_0 . Предположим, наконец, что все распределения $G_{ij}(t)$ фиксированы, а интенсивности $\lambda_i(e)$ имеют вид $\lambda_i(e) = \lambda_i^{(0)}(e) \varepsilon$, где $\lambda_i^{(0)}(e)$ фиксированы, а $\varepsilon \rightarrow 0$.

Теорема 18. Если существует момент

$$\int_0^\infty x^{m_0} dG(x) < \infty, \text{ то}$$

$$\lim_{\varepsilon \rightarrow 0} P \{ \Lambda \tau'_k > x \} = e^{-x}, \quad (43)$$

причем сходимость равномерна по k . Здесь

$$\Lambda = \sum_{\pi \in \Pi_0} \lambda_{i_1}(0) \lambda_{i_2}(e^{(1)}) \dots \lambda_{i_{m_0}}(e^{(m_0-1)}) J(\pi);$$

$$J(\pi) = \int_{t_1 > t_2 > \dots > t_{m_0} > 0} \prod_{e=1}^s \bar{G}_{i_k e}(t_{ke}) \times dt_1 dt_2 \dots dt_{m_0};$$

$i_{k1} = i_1, i_{k2}, \dots, i_{ks}$ — номера элементов, которые поступают на ремонтные единицы: j_1, j_2, \dots, j_s — соответствующие номера ремонтных единиц, на которые они поступают.

Из этой теоремы следует, что в случае, когда величины $\lambda_i(e) T$ малы и притом одного порядка малости, справедлива приближенная формула

$$P \{ \tau'_k > t \} \approx e^{-\Lambda t}. \quad (44)$$

Формулировка теоремы в силу общности предположений весьма гро-

моздка. Однако для ряда частных случаев выражение для Λ сильно упрощается.

Рассмотрим некоторые частные случаи общей модели.

1. Пусть $G_{ij}(x) = G(x)$ и каждая ремонтная единица доступна для любого элемента. Тогда

$$\Lambda = \sum_{\pi \in \Pi_0} \lambda_{i_1}(0) \lambda_{i_{m_0}}(e^{(m_0-1)}) \times J_{m_0-1, r}, \quad (45)$$

где $J_{n, r}$ — определена в теореме 12.

2. Пусть дополнительно к условиям случая 1 $\lambda_i(e) = \lambda_i$, если $e_i = 0$. Тогда

$$\Lambda = \sum_{e^{(m_0)}} J_{m_0-1, r} m_0! \lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_{m_0}}, \quad (46)$$

где k_1, k_2, \dots, k_{m_0} — номера элементов, отказавших в состоянии $e^{(m_0)}$, сумма берется по всем состояниям $e^{(m_0)}$, которые являются последними состояниями для путей из класса Π_0 .

3. Пусть дополнительно к условиям случая 1 $E_+ = \{e: \|e\| < m\}$, где

$$\|e\| = \sum_{i=1}^n e_i \text{ и } \lambda_i(e) = \lambda(\|e\|). \text{ Это}$$

означает, что элементы входят в систему симметрично. Тогда

$$\Lambda = \lambda(0) \lambda(1) \lambda(m-1) J_{m-1, r}. \quad (47)$$

В этом случае процесс $\|e(t)\|$ совпадает с процессом $x(t)$, описывающим общую модель резервирования с восстановлением.

4. Пусть каждый элемент имеет свою ремонтную единицу

$$G_{ij}(t) = G_i(t), \quad T_i = \int_0^\infty t dG_i(t),$$

$\lambda_i(e) = \lambda_i$, если $e_i = 0$. Тогда

$$\Lambda = \sum_{e^{(m_0)}} \lambda_{k_1} \dots \lambda_{k_{m_0}} \left(\frac{1}{T_{k_1}} + \dots + \frac{1}{T_{k_{m_0}}} \right) T_{k_1} \dots T_{k_{m_0}}, \quad (48)$$

где индексы k_1, k_2, \dots, k_{m_0} и суммирование определяются так, как и в формуле (46).

Двусторонняя оценка надежности одной восстанавливаемой системы. Для характеристик надежности восстанавливаемых систем можно получать неравенства такого же типа, какие выведены выше. Однако эти оценки получаются менее точными и более громоздкими. Приведем один пример, в котором удается найти сравнительно простые двусторонние неравенства. Рассмотрим случай 4, где величина Λ выражена формулой (48). Справедлива теорема 19: вероятность безотказной работы удовлетворяет неравенству

$$e^{-\lambda \bar{q} t} \leq P\{\tau_0 > t\} \leq e^{-\lambda q t} +$$

$$+ \prod_{i=1}^n (\lambda_i T_i + 1) - 1,$$

$$\text{где } \lambda = \sum_{k=1}^n \lambda_k;$$

$$\lambda q = \sum_{e \in \Gamma_+} \lambda_-(e) \left[1 - \sum_{k=1}^n \frac{\lambda m_k}{2T_k} \right] \times$$

$$\times \prod_{k=1}^n (\lambda_i T_i)^{e_i};$$

$$\lambda \bar{q} = \sum_{e \in \Gamma_+} \lambda_-(e) \prod_{k=1}^n (\lambda_i T_i)^{e_i};$$

$$m_k = \int_0^\infty t^2 dG_k(t); \Gamma_+ \text{ — множество гра-}$$

ничных состояний из E_+ , из которых можно перейти в отказовое состояние при отказе одного элемента; $\lambda_-(e)$ — суммарная интенсивность отказа системы в состоянии e . Нетрудно убедиться, что выражение для $\lambda \bar{q}$ совпадают с (48), и поэтому предел в теореме 18 является одновременно пессимистической оценкой.

1. ФИЗИЧЕСКОЕ ОПИСАНИЕ ЗАДАЧИ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ

Проблема технического обслуживания возникает на этапе эксплуатации любой технической системы. Техническое обслуживание есть вмешательство (целенаправленное внешнее воздействие) в работу системы, направленное на получение максимального эффекта от эксплуатации системы, поэтому проблема технического обслуживания заключается в разработке оптимального процесса технического обслуживания эксплуатируемой системы или оптимального процесса управления.

При таком подходе возникают три вопроса: чем управлять, каков объект управления? Как управлять, как воздействовать на объект? Для чего управлять, каковы цели управления?

Объектом управления является техническая система, которая в каждый момент времени t характеризуется некоторым состоянием $x(t)$ из множества возможных состояний X ($x(t) \in X$).

Функция $x(t)$ характеризует эволюцию системы во времени. Если система функционирует под воздействием случайных факторов, то эволюция состояний технической системы во времени описывается некоторым случайным процессом $\xi(t)$.

Вывод: в физической и математической модели объектом управления является процесс $x(t)$ или $\xi(t)$, характеризующий эволюцию состояний системы во времени.

Воздействие на объект. В задачах технического обслуживания (ТО) воздействовать на объект, т. е. управлять объектом, можно выбором сроков проведения ТО и характером ТО.

Характер ТО определяется *глубиной восстановления*. Глубина восстановления и сроки проведения этого восстановления определяют вид восстановительной работы.

Для системы задается множество (конечное или бесконечное) возможных в системе восстановительных работ $D = (d_1, d_2, \dots, d_n, \dots)$.

Классификация восстановительных работ. В реальной ситуации восстановительные работы (воздействия на систему) осуществляются не мгновенно, а занимают некоторое время. Поэтому классификацию целесообразно провести по факторам, влияющим на длительность восстановительной работы: 1) момент начала восстановительной работы (известен ли заранее этот момент или нет); 2) состояние системы в момент начала проведения восстановительной работы; 3) состояние системы к моменту окончания восстановительной системы (глубина восстановления).

Если момент начала восстановительной работы известен, то такую восстановительную работу называют *плановой*. В противном случае работу называют *внеплановой*. Если в начале восстановительной работы система была работоспособной, то такую восстановительную работу называют *предупредительной*, если же система была неработоспособной, то работу называют *аварийной*.

Виды восстановительных работ по предлагаемой классификации приведены в таблице.

Цель управления (цели ТО) — получение от эксплуатации технической системы максимального эффекта (максимального значения показателя эффективности функционирования).

Глубина восстановления системы	Восстановительные работы с системой	
	работоспособной	неработоспособной
Никакого обновления в системе не проводится	Плановый (внеплановый) осмотр или проверка работоспособности	—
Проводится полное обновление	Плановая (внеплановая) предупредительная профилактика	Плановый (внеплановый) аварийно-профилактический ремонт
Проводится обновление части системы	Плановая (внеплановая) предупредительная профилактика части системы	Плановый (внеплановый) аварийно-профилактический ремонт части системы

Величина эффекта зависит от эволюции (траектории) процесса $\xi(t)$, поэтому показатель эффективности функционирования (и ТО) за $[0, T]$ есть функционал, построенный на траекториях процесса $\xi(t)$:

$$J = J(\xi(t), t \in [0, T]).$$

Так как показатель J есть случайная величина, то при постановке задачи оптимизации за критерий оптимизации выбирают

$$MJ = MJ(\xi(t), t \in [0, T]).$$

Задача оптимизации (физическая формулировка). Определить правило выбора управлений в зависимости от эволюции процесса $\xi(t)$, т. е. стратегию управления, обеспечивающую максимальное значение функционала качества функционирования MJ .

Замечание. Техническое обслуживание (проведение восстановительных работ) связано, как правило, с прерыванием процесса функционирования, занимает некоторое время (случайное или детерминированное) и не может осуществляться непрерывно, а реализуется периодически, т. е. периоды функционирования чередуются с периодами технического обслуживания (периодами проведения восстановительных работ). Такое управление называют *дискретным*.

2. МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

2.1. ДИСКРЕТНОЕ УПРАВЛЕНИЕ ПРОЦЕССОМ $\xi(t)$ С ДИСКРЕТНЫМ ВРЕМЕНЕМ

Случайный процесс. Пусть $\xi(t)$ — случайный процесс, принимающий значения из некоторого измеримого пространства (X, \mathfrak{X}) , $\xi(t) \in X$, \mathfrak{X} , есть σ — алгебра подмножеств множества X , $t \in N$, где $N = (0, 1, 2, \dots)$ — множество неотрицательных целых чисел.

Процесс с дискретным временем задается набором функций распределения $n > 1$

$$F_n(A, x_0, x_1, \dots, x_{n-1}) = P(\xi(n) \in A / \xi(0) = x_0, \xi(n-1) = x_{n-1})$$

и распределением начальных состояний $P(\xi(0) \in A)$, $A \in \mathfrak{X}$.

Пространство управлений — измеримое пространство (U, \mathfrak{B}) . Как правило, будем считать $U = R^{(1)}$, или U — счетное множество, или U — конечное множество.

Управляемый объект задается семейством распределений:

$$F_0(A) = P(\xi(0) \in A),$$

$$F_n(A, x_0, \dots, x_{n-1}, u_0, \dots, u_{n-1}) = P(\xi(n) \in A / \xi(0) = x_0, \dots, \xi(n-1) = x_{n-1}, u_0, \dots, u_{n-1}) \quad (1)$$

— условным распределением значений процесса в момент n при условии, что в моменты $k < n$ процесс принимал значения x_k , и управления равны u_k , $u_k \in U$, $A \in \mathfrak{A}$, $k = 0, 1, 2, \dots, n-1$.

Замечания. 1. Если $X = R^{(1)}$, то в качестве множества A можно брать множество $(-\infty, x)$.

2. Семейство условных распределений (1) порождает меру на множествах вида при $\bar{x} = (x_0, \dots, x_n)$

$$A^{(n)} = \{\bar{x} : x_0 \in A_0, \dots, x_n \in A_n\},$$

$$A_i \in \mathfrak{A}.$$

При фиксированном $\bar{u} = (u_0, u_1, \dots, u_n)$ имеем

$$\mu(A^{(n)}/\bar{u}) = P(\bar{\xi}(n) \in A^{(n)}/\bar{u}) =$$

$$= \int_{A_0} dF_0(x_0) \int_{A_1} dF_1(x_1, x_0, u_0)$$

$$\int_{A_n} dF_n(x_n, x_0, \dots, x_{n-1},$$

$$u_0, \dots, u_{n-1}),$$

где $\bar{\xi}(n) = (\xi(0), \xi(1), \dots, \xi(n))$.

Выбор управления. Стратегия управления. Управление $\eta(n)$ в момент $n > 0$ задается семейством распределений $\Phi_n(u, x_0, \dots, x_n, u_0, \dots, u_{n-1})$ — вероятностью того, что управление в момент n выбирается из множества $B = (-\infty, u) \in \mathfrak{B}$ при условии, что в моменты $k (k < n)$ $\xi(k) = x_k$, принимались управления u_k и $\xi(n) = x_n$. При $n = 0$ задается $\Phi_0(u, x)$ — вероятность того, что управление в момент 0 выбирается из множества $B = (-\infty, u)$ при условии $\xi(0) = x$.

Семейство условных распределений Φ_n , $n \geq 0$ или порожденную им меру $\nu(\cdot/\cdot)$ называют стратегией управления.

Замечание. Семейство условных распределений Φ_n , $n \geq 0$, порождает меру на множествах вида

$$B^{(n)} = \{\bar{u} : u_0 \in B_0, \dots, u_n \in B_n\};$$

$$B_i \in \mathfrak{B};$$

$$\bar{u} = (u_0, \dots, u_n) \in U^{(n+1)} =$$

$$= U \times U \times \dots \times U.$$

При любом фиксированном $\bar{x} = (x_0, x_1, \dots, x_n)$ имеем

$$\nu(B^{(n)}/\bar{x}) = P(\bar{\eta}(n) \in B^{(n)}/\bar{x}) =$$

$$= \int_{B_0} d\Phi_0(u_0, x_0) \int_{B_1} d\Phi_1(u_1, x_0, x_1, u_0) \dots$$

$$\int_{B_n} d\Phi_n(u_n, x_0, \dots, x_n, u_0, \dots, u_{n-1}),$$

(2)

где $\bar{\eta}(n) = (\eta(0), \eta(1), \dots, \eta(n))$.

Управляемым случайным процессом с дискретным временем (управляемая последовательность) называют последовательность пар $(\xi(n), \eta(n), n \geq 0)$, принимающих значения из пространства $X \times U$, причем

$$P(\xi(n) \in A/\xi(0), \eta(0), \dots, \xi(n-1), \eta(n-1)) = F_n(A, \xi(0), \xi(1),$$

$$\xi(n-1), \eta(0), \dots, \eta(n-1));$$

$$P(\eta(n) \in B/\xi(0), \eta(0), \dots, \xi(n-1), \eta(n-1), \xi(n)) = \Phi_n(B, \xi(0),$$

$$\xi(n), \eta(0), \dots, \eta(n-1)),$$

$$A \in \mathfrak{A}, B \in \mathfrak{B}.$$

Конечномерные распределения последовательности определяются равенством

$$P(\xi(0) \in A_0, \eta(0) \in B_0, \dots, \xi(n) \in A_n,$$

$$\eta(n) \in B_n) = \int_{A_0} dF_0(x_0) \int_{B_0} d\Phi_0(u_0, x_0) \dots$$

$$\int_{A_n} dF_n(x_n, x_0, \dots, x_{n-1},$$

$$u_0, \dots, u_{n-1}) \int_{B_n} d\Phi_n(u_n, x_0, \dots, x_n,$$

$$u_0, \dots, u_{n-1}).$$

2.2. ДИСКРЕТНОЕ УПРАВЛЕНИЕ СЛУЧАЙНЫМ ПРОЦЕССОМ $\xi(t)$ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Случайный процесс. Пусть $\xi(t)$ — случайный процесс, принимающий значения из некоторого измеримого пространства (X, \mathfrak{A}) и $t \in [0, T]$ ($T < \infty$ либо $T = \infty$), $X^{[0, T]}$ — пространство

функций (траекторий процесса $\xi(t)$), определенных на $[0, T]$ и принимающих значения из X .

Процесс с непрерывным временем задается семейством согласованных распределений

$$F(A_0, \dots, A_n, t_0, \dots, t_n) = P\left(\bigcap_{k=0}^n \{\xi(t_k) \in A_k\}\right) \quad (3)$$

для любого $n \geq 0$, любого набора $0 = t_0 < t_1 < \dots < t_n \leq T$ и любых $A_k \in \mathfrak{A}$, $k \leq n$ или условных распределений

$$\tilde{F}(A_0, \dots, A_n, t_0, \dots, t_n) = P\left(\xi(t_n) \in A_n \middle/ \bigcap_{k=0}^{n-1} \{\xi(t_k) \in A_k\}\right) \quad (4)$$

и начальным распределением

$$F(A_0, t_0) = P(\xi(t_0) \in A_0).$$

Семейство распределений (3) порождает меру

$$\mu(A), \quad A \in \mathcal{X}^{[0, T]} \quad (5)$$

в пространстве траекторий $\mathcal{X}^{[0, T]}$ процесса $\xi(t)$.

Семейство условных распределений (4) порождает семейство мер $\mu_t(A/\xi(\cdot))$, определенное при $t \in [0, T]$, $A \in \mathfrak{A}^{[t, T]}$ при условии фиксированной траектории $\xi(s)$, $s \in [0, t]$, где $\mathfrak{A}^{[t, T]}$ — σ -алгебра из $\mathcal{X}^{[0, T]}$, порожденная цилиндрическими множествами при $s \in [t, T]$.

Пространство управлений. Выбор управления. Пусть задано измеримое пространство управлений (U, \mathfrak{B}) ; $U^{[0, T]}$ — множество ступенчатых функций, определенных на $[0, T]$ со значениями из U .

Управление $\eta(t)$ как случайный процесс задается семейством условных распределений моментов скачков процесса $\eta(t)$ (моментов выбора управления) и значений этого процесса:

$$\begin{aligned} &\lambda_k(t/\xi(\cdot), u_0, \dots, u_k, \tau_0, \dots, \tau_k); \\ &\nu_k(B_k/\xi(\cdot), u_0, \dots, u_{k-1}, \tau_0, \dots, \tau_k); \\ &k \geq 0, \quad B_k \in \mathfrak{B}; \\ &0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_k < t, \quad (6) \end{aligned}$$

где $\lambda_k(\cdot/\cdot)$ есть распределение момента θ_k k -го скачка процесса $\eta(t)$ при условии $\theta_s = \tau_s$, $\eta(\tau_s) = u_s$, $s \leq k$ и фиксирована траектория $\xi(\cdot)$ до момента τ_k ; $\nu_k(\cdot/\cdot)$ есть распределение управлений (значения процесса $\eta(t)$) в момент τ_k при условии $\theta_s = \tau_s$, $s \leq k$, $\eta(\tau_s) = u_s$, $s < k$ и фиксирована траектория $\xi(\cdot)$ до момента τ_k .

Семейство условных распределений (6) или порождаемое ими совместное условное распределение пары $(\eta(\theta_k), \theta_k)$ называют стратегией управления.

Семейство условных распределений λ_k и ν_k определяет конечномерные распределения случайного процесса $\eta(t)$ при условии фиксированной траектории $\xi(t)$. Тем самым порождается мера $\nu/(B/\xi(\cdot))$ на пространстве траекторий $U^{[0, T]}$, $B \in U^{[0, T]}$.

Управляемым объектом называют семейство $\mu(A/\eta(\cdot))$ вероятностных мер, $A \in \mathcal{X}^{[0, T]}$, $\eta(\cdot) \in U^{[0, T]}$, т. е. для управляемого объекта предполагается, что порожденная мера $\mu(A)$ [см. (5)] зависит от ступенчатой функции $\eta(\cdot)$.

Управляемый случайный процесс с непрерывным временем и дискретным (ступенчатым) управлением. Управляемый случайный процесс есть пара $(\xi(t), \eta(t))$; $\xi(t) \in \mathcal{X}^{[0, T]}$, $\eta(t) \in U^{[0, T]}$, конечномерные распределения которой задаются следующим образом:

1. При $\theta_0 = 0$ начальное распределение $\xi(0)$ определяется из формулы (4)

$$P(\xi(0) \in A) = F(A, 0).$$

2. Управление $\eta(0)$ задается распределением (6)

$$\nu_0(B_0/\xi(0), \theta_0).$$

3. Интервал выбора следующего управления θ_1 задается распределением (6)

$$\lambda_0(t/\xi(0), \eta(0), \theta_0).$$

4. При $0 \leq t \leq \theta_1$ полагаем $\eta(t) = \eta(0)$, а распределение компоненты $\xi(t)$ определяется мерой $\mu(A/\eta(0))$, т. е.

$$P(\xi(t) \in A/\eta(0)) = \int_A \mu(dx/\eta(0)).$$

5. В момент θ_1 процесс построения распределений повторяется, т. е. при известной траектории $\xi(t), 0 \leq t \leq \theta_1$ и управлении $\eta(0)$:

определяется управление $\eta(\theta_1)$ с распределением

$$v_1(B_1/\xi(\cdot), \eta(0), \theta_0, \theta_1);$$

определяется момент выбора следующего управления θ_2 с распределением

$$\lambda_1(t/\xi(\cdot), \eta(0), \eta(\theta_1), \theta_0, \theta_1);$$

при $\theta_1 < t \leq \theta_2$ полагаем $\eta(t) = \eta(\theta_1)$, а распределение компоненты $\xi(t)$ определяется мерой $\mu(A/\eta(\cdot))$.

Таким образом, определяется поведение $(\xi(t), \eta(t))$ на $\theta_1 \leq t \leq \theta_2$ и так далее.

2.3. КЛАССЫ СТРАТЕГИЙ И ЗАДАЧА ОПТИМИЗАЦИИ

Рандомизированная стратегия. Управление процессом с дискретным временем. Стратегию, определяемую семейством невырожденных условных распределений Φ_n (см. п. 2.1), называют **рандомизированной**.

Управление процессом с непрерывным временем. Стратегию называют **рандомизированной**, если условные распределения (6), определяющие момент выбора управления и значение управления, невырожденные.

Нерандомизированная стратегия. Управление процессом с дискретным временем. Стратегию, определяемую семейством вырожденных при любом наборе $(x_0, \dots, x_n, u_0, \dots, u_{n-1})$ (сосредоточенных в одной точке) условных распределений Φ_n , называют **нерандомизированной**.

Замечание. Если для любого $n > 0$ и любого набора $(x_0, \dots, x_n, u_0, \dots, u_{n-1})$ функция $\Phi_n(\cdot)$ вырожденная, т. е.

$$\Phi_n(u, x_0, \dots, x_n, u_0, \dots, u_{n-1}) = \begin{cases} 0 & \text{при } u \leq u^*, \\ 1 & \text{при } u > u^*, \end{cases}$$

то управление u^* , принимаемое в момент n , есть функция набора $(x_0, \dots, x_n, u_0, \dots, u_{n-1})$, т. е. управление

в момент n есть функция поведения управляемого процесса в прошлом $\{\xi(k), \eta(k), k < n\}$ и настоящего значения $\xi(n)$.

Управление процессом с непрерывным временем. Стратегию называют **нерандомизированной**, если распределения (6) вырожденные при любых значениях условий.

Замечание. Аналогично предыдущему замечанию, для нерандомизированных стратегий управление, принимаемое в момент t , есть детерминированная функция траектории процесса $\xi(s), s \leq t$ и управления $\eta(s), s < t$.

Марковская стратегия. Стратегия, определяемая семейством условных распределений $\Phi_n, n \geq 0$ или $(\lambda_k, v_k), k \geq 0$, называется **марковской**, если эти распределения зависят только от настоящего значения процесса $\xi(t)$ и не зависят от прошлого, т. е.

$$\begin{aligned} \Phi_n(u, x_0, \dots, x_n, u_0, \dots, u_{n-1}) &= \\ &= \Phi_n(u, x_n), \\ \lambda_k(t/\xi(\cdot), \dots) &= \lambda_k(t/\xi(\tau_k), \tau_k), \\ v_k(t/\xi(\cdot), \dots) &= v_k(t/\xi(\tau_k), \tau_k). \end{aligned} \quad (7)$$

Однородная стратегия. Управление процессом с дискретным временем. Марковскую стратегию, определяемую семейством $\Phi_n(u, x)$, называют **однородной**, если $\Phi_n(u, x_1) \equiv \Phi_m(u, x_2)$ при $x_1 = x_2$ при любых m и n .

Замечание. Однородность для марковской стратегии означает независимость от времени принимаемых управлений. Если в разные моменты времени $m \neq n$ процесс принял одно и то же значение $\xi(n) = x_1 = x_2 = \xi(m)$, то управление определяется одним и тем же законом

$$\Phi_n(u, x_1) \equiv \Phi_m(u, x_2).$$

Управление процессом с непрерывным временем. Марковскую стратегию, определяемую семейством условных распределений (7), называют **однородной**, если функции v_k и λ_k не зависят от τ_k и k .

$$\begin{aligned} \lambda_k(t/\xi(\tau_k), \tau_k) &= \lambda(t/\xi(\tau_k)), \\ v_k(t/\xi(\tau_k), \tau_k) &= v(t/\xi(\tau_k)). \end{aligned}$$

Пусть задан управляемый случайный процесс $\{\xi(t), \eta(t)\}$ (или последовательность $\{\xi(k), \eta(k), k \geq 0\}$), пусть $\{\bar{\xi}(s), \bar{\eta}(s), s \leq t\}$ — фиксированная траектория процесса $\xi(t)$ и принятые управления $\eta(t)$ на интервале времени $[0, t]$, пусть $J(\bar{\xi}, \bar{\eta})$ есть функционал, зависящий от траектории и принятых управлений.

Качество принятой стратегии управления, определяемой мерой $v(\cdot/\cdot)$, характеризуется математическим ожиданием функционала $J(\bar{\xi}, \bar{\eta})$ по мере $v(\cdot/\cdot)$

$$J(v) = M_v J(\bar{\xi}, \bar{\eta}).$$

Задача оптимизации состоит в определении такой стратегии (меры v^*) из некоторого класса допустимых стратегий \mathfrak{N} , для которой

$$J(v^*) = \max_{v \in \mathfrak{N}} J(v) = \max_{v \in \mathfrak{N}} M_v J(\bar{\xi}, \bar{\eta})$$

или

$$J(v^*) = \min_{v \in \mathfrak{N}} J(v) = \min_{v \in \mathfrak{N}} M_v J(\bar{\xi}, \bar{\eta}),$$

если экстремумы (max и min) достигаются, т. е. оптимальная стратегия в классе допустимых стратегий \mathfrak{N} существует.

2.4. ОБЩИЕ СВОЙСТВА ОПТИМАЛЬНЫХ СТРАТЕГИЙ

Дискретное время. Если функции $\Phi_n(u, x_0, \dots, x_n, u_0, \dots, u_{n-1})$ измеримы относительно $\mathfrak{A}_n \times \mathfrak{B}^{(n-1)}$, где \mathfrak{A}_n есть σ -алгебра подмножеств $A^{(n)}$, $\mathfrak{B}^{(n)}$ — σ -алгебра подмножеств $B^{(n)}$ и для всех допустимых стратегий управления \mathfrak{N} , $v(\cdot/\cdot) \in \mathfrak{N}$ для всех n $P(\bar{\eta}(n) \in \Gamma_n) = 1$, где $\Gamma_n \in U^{(n+1)}$, то класс управлений \mathfrak{N} называют классом управлений с ограничениями.

Теорема. Если множество допустимых стратегий \mathfrak{N} является классом управлений с ограничениями, то каково бы ни было $v \in \mathfrak{N}$, найдется нерандомизированная стратегия $\bar{v} \in \mathfrak{N}$, такая, что

$$J(\bar{v}) \geq J(v) \quad (8)$$

(в задаче определения максимума $J(v)$).

Замечание. Неравенство (8) означает, что для каждой стратегии найдется нехудшая нерандомизированная стратегия \bar{v} , следовательно,

$$\max_{v \in \mathfrak{N}} J(v) = \max_{v \in \bar{\mathfrak{N}}} J(v) = J(\bar{v}^*),$$

где $\bar{\mathfrak{N}}$ — класс нерандомизированных стратегий из \mathfrak{N} ($\bar{\mathfrak{N}} \subset \mathfrak{N}$), т. е. оптимальную стратегию можно определять из множества нерандомизированных стратегий.

Теорема существования. Если $U \subseteq R^{(1)}$ и $X \subseteq R^{(1)}$ для управляемого объекта $\mu(\cdot/\bar{u})$ при любой функции $g(x) \in \bar{C}$, где \bar{C} — множество непрерывных ограниченных на $R^{(1)}$ функций,

$$\int_U g(u) d\Phi_n(u, x_0, \dots, x_n, u_0, \dots, u_{n-1})$$

непрерывен по совокупности переменных и функционал $J(\bar{\xi}, \bar{\eta})$ ограничен и непрерывен, то существует оптимальная стратегия в классе всех стратегий.

Теорема (управление марковской цепью). Если управляемый объект $\xi(n)$ есть марковская цепь с конечным множеством состояний X и пространство управлений U конечно, то при управлении за конечное время ($T < \infty$) оптимальная стратегия существует и может быть выбрана из класса марковских нерандомизированных стратегий.

3. УПРАВЛЯЕМЫЕ ПОЛУМАРКОВСКИЕ ПРОЦЕССЫ

Определение управляемого объекта. Пусть $\xi(t)$ есть полумарковский процесс, принимающий значения из конечного пространства $E = \{e_1, \dots, e_N\}$, $N < \infty$, определяемый семейством полумарковских матриц $\{Q_{ij}(t, \alpha)\}$,

$$0 \leq Q_{ij}(t, \alpha) \leq 1, \quad Q_{ij}(+\infty, \alpha) = p_{ij}(\alpha), \quad \sum_{j=1}^N p_{ij}(\alpha) = 1, \quad (9)$$

монотонных по $t \in [0, +\infty)$ при любых $i, j \in [1, 2, \dots, N]$ и $\alpha \in [0, +\infty)$ (α — параметр). Матрица $(p_{ij}(\alpha))$ определяет эволюцию состояний полумарковского процесса, которая описы-

вается цепью Маркова с вероятностями перехода $p_{ij}(\alpha)$. Время $\tau(i, j)$ — время пребывания процесса $\xi(t)$ в состоянии e_i при переходе в состояние e_j . Оно имеет распределение

$$F_{ij}(t) = P(\tau(i, j) < t) = \frac{Q_{ij}(t, \alpha)}{Q_{ij}(+\infty, \alpha)}. \quad (10)$$

Если $\tau(i)$ — время пребывания процесса $\xi(t)$ в состоянии e_i до первого выхода из него, то

$$F_i(t, \alpha) = P(\tau(i) < t) = \sum_{j=1}^N Q_{ij}(t, \alpha). \quad (11)$$

Пространство управлений и стратегий. Пусть пространство управлений $U = R^{(N)}$, а множество допустимых стратегий \mathfrak{M} совпадает с множеством марковских однородных рандомизированных стратегий, определяемых конечным набором распределений

$$\Phi(\alpha, e_i) = \Phi^{(i)}(\alpha),$$

$$0 \leq \Phi^{(i)}(\alpha) \leq 1,$$

монотонно возрастающих по $\alpha \in [0, +\infty)$, $\Phi^{(i)}(0) = 0$, $\Phi^{(i)}(+\infty) = 1$ при любом $i \in [1, 2, \dots, N]$, причем в силу марковости и однородности стратегии

$$\lambda_k(t, \xi(\tau_k), \tau_k) = \lambda_k(t - \tau_k / \xi(\tau_k), 0) = F_{\xi(\tau_k)}(t - \tau_k, \alpha) =$$

$$= \sum_{j=1}^N Q_{\xi(\tau_k)j}(t - \tau_k, \alpha);$$

$$v_k(B_k / \xi(\tau_k), \tau_k) = v_k(B_k / \xi(\tau_k)) = \int_{B_k} d\Phi^{(\xi(\tau_k))}(\alpha); \quad B_k \in [0, +\infty).$$

Замечание. Указанное построение меры, определяющей управление $\alpha \in B_k$ в момент τ_k и момент t принятия последующего управления, означает, что управления изменяются в моменты изменения состояний полумарковского процесса $\xi(t)$, т. е. для управляемого полумарковского процесса $(\xi(t), \eta(t))$ со ступенчатым (дискретным)

управлением моменты скачков обеих компонент совпадают.

Управляемый однородный полумарковский процесс $(\xi(t), \eta(t))$, $\xi(t) \in E$, $\eta(t) \in [0, +\infty)$ задается семейством полумарковских матриц $\{Q_{ij}(t, \alpha)\}$

$$Q_{ij}(t, \alpha) = P(\xi(\tau) = e_j, \tau = \tau(i, j) < t / \xi(0) = e_i, \eta(0) = \alpha), \quad (12)$$

$$\lambda_k(t / \xi(0) = e_i, \tau_k = 0) = \sum_{j=1}^N Q_{ij}(t, \alpha), \quad (13)$$

распределением

$$v_k(B / \xi(0) = e_i, \tau_k = 0) = P(\eta(0) \in B / \xi(0) = e_i) = \int_B d\Phi^{(i)}(\alpha) \quad (14)$$

и начальным распределением

$$P(\xi(0) \in A). \quad (15)$$

Описание эволюции управляемого полумарковского процесса $(\xi(t), \eta(t))$:

1. Определяется начальное состояние $\xi(0) = e_i$ в соответствии с распределением (15).

2. Определяется управление $\eta(0) = \alpha$ в соответствии с распределением (14).

3. Определяется момент $\tau(i)$, момент выхода из состояния e_i и состояние e_j , куда перейдет процесс $\tau(i)$ в соответствии с распределением (12) и (13).

4. Далее процесс повторяется для состояния e_j (в этом случае распределения (12) и (13) определяют интервал следующего перехода, отсчитываемого от момента $\tau(i)$; это справедливо вследствие однородности управляемого процесса).

Функционал дохода. Пусть $\tau = \tau(i, j, \alpha)$ есть время перехода процесса $\xi(t)$ из состояния e_i в состояние e_j , если на этом периоде принять управление α . Доход от пребывания процесса в состоянии e_i за время t ($0 \leq t \leq \tau$) определяется функцией $R_{ij}(t, \tau, \alpha) > 0$. Полный доход за все время пребывания в состоянии

$$R_{ij}(\tau, \alpha) = R_{ij}(\tau, \tau, \alpha).$$

Если $\xi(t) = e_{i_k}$ при $t \in [t_k, t_{k+1})$, $k = 0, n+1$, $0 = t_0 < t_1 < t_2 < \dots < t_n < T < t_{n+1}$ и в моменты

перехода $\{t_k\}$ принимались управления $\{\alpha_k\}$, то суммарный доход за период $[0, T]$ определяется равенством

$$J(\bar{\xi}(t), \bar{\eta}(t), t \in [0, T]) = \\ = J(\{t_k\}, \{e_i\}, \{\alpha_k\}, k = 0, \overline{n+1}) = \\ = \sum_{k=0}^{n-1} R_{i_k i_{k+1}}(t_{k+1} - t_k, \alpha_k) + \\ + R_{i_n i_{n+1}}(T - t_n, t_{n+1} - t_n, \alpha_n).$$

Качество принятой стратегии, определяемой распределением (14), характеризуется в соответствии с п. 2.4 функционалом

$$J(\Phi_1, \Phi_2, \dots, \Phi_N, T) = \\ = MJ(\bar{\xi}(t), \bar{\eta}(t), t \in [0, T]). \quad (16)$$

Если $T \rightarrow \infty$, то функционал (16) неограниченно возрастает. В этом случае качество стратегии характеризуют пределом

$$J(\Phi_1, \dots, \Phi_N) = \\ = \lim_{T \rightarrow \infty} \frac{J(\Phi_1, \Phi_2, \dots, \Phi_N, T)}{T}. \quad (17)$$

Постановка и решение задачи оптимизации. Определить функции распределения $\Phi_1^*, \Phi_2^*, \dots, \Phi_N^*$, для которых

$$J(\Phi_1^*, \dots, \Phi_N^*) = \\ = \max_{\Phi_i \in \mathfrak{N}, i=1, N} J(\Phi_1, \dots, \Phi_N),$$

где \mathfrak{N} — класс допустимых функций распределения, для которых функционал J существует.

Пусть

$$Q_{ij}(t) = \int_0^\infty Q_{ij}(t, \alpha) d\Phi_i(\alpha); \\ \rho_{ij}(\alpha) = Q_{ij}(+\infty, \alpha); \\ \rho_{ij} = \int_0^\infty \rho_{ij}(\alpha) d\Phi_i(\alpha)$$

и существуют стационарные вероятности распределения состояний цепи Мар-

кова с матрицей перехода $P = (p_{ij})$ для любых $\{\Phi_i\}$. Тогда

$$J(\Phi_1, \dots, \Phi_N) = \frac{\sum_{j=1}^N \rho_j \pi_j}{\sum_{j=1}^N T_j \pi_j}, \quad (18)$$

где π_j — стационарные вероятности распределения состояний цепи Маркова с вероятностями перехода p_{ij} , величины π_i определяются равенствами

$$\pi_k = \sum_{i=1}^N p_{ik} \pi_i, \quad \sum_{k=1}^N \pi_k = 1;$$

ρ_k — средний доход за период пребывания процесса в состоянии e_k ,

$$\rho_k = \int_0^\infty \left[\int_0^\infty \sum_{j=1}^N R_{kj}(\tau, \alpha) \times \right. \\ \left. \times d_\tau Q_{kj}(\tau, \alpha) \right] d\Phi_k(\alpha);$$

T_k — среднее время пребывания процесса в состоянии e_k до первого выхода из него

$$T_k = \int_0^\infty \int_0^\infty \sum_{j=1}^N \tau d_\tau Q_{kj}(\tau, \alpha) d\Phi_k(\alpha).$$

Числитель выражения (18) есть математическое ожидание дохода за один период между соседними моментами изменения состояния; знаменатель есть средняя длительность этого периода.

Структура функционала. **Теорема.** Функционал $J(\Phi_1, \dots, \Phi_N)$, определяемый равенством (18), есть дробно-линейный функционал относительно функций распределения $\Phi_1, \Phi_2, \dots, \Phi_N$ и имеет вид

$$J(\Phi_1, \Phi_2, \dots, \Phi_N) = \\ = \frac{\int \int A(\alpha_1, \alpha_2, \dots, \alpha_N) \times \\ \times d\Phi_1(\alpha_1) \dots d\Phi_N(\alpha_N)}{\int \int B(\alpha_1, \dots, \alpha_N) \times \\ \times d\Phi_1(\alpha_1) \dots d\Phi_N(\alpha_N)}. \quad (19)$$

Замечания: 1. Функции $A(\cdot)$ и $B(\cdot)$ зависят только от функций $R_{ij}(t, \alpha)$ и $Q_{ij}(t, \alpha)$, $i, j \in [1, \dots, N]$ и не зависят от распределений $\{\Phi_k(\alpha)\}$.

2. Функциям $A(\cdot)$ и $B(\cdot)$ можно дать следующую вероятностную интерпретацию:

функция $A(\alpha_1, \dots, \alpha_N)$ есть условное математическое ожидание дохода на периоде между соседними моментами изменения состояний управляемого полумарковского процесса при условии, что в состоянии e_i ($i = 1, 2, \dots, N$) принимается управление α_i ;

функция $B(\alpha_1, \dots, \alpha_N)$ есть условное математическое ожидание длительности периода изменения состояния этого процесса при аналогичных условиях.

Структура экстремальных функций. *Теорема.* Пусть функции $A(\cdot)$ и $B(\cdot)$ ограничены на любом ограниченном множестве и пусть $B(\cdot) > 0$, тогда, если существует максимум функционала (19) по множеству функций распределения \mathfrak{M} , для которого этот функционал существует, то он достигается на некоторых вырожденных распределениях $\Phi_1^*, \Phi_2^*, \dots, \Phi_N^*$

$$\Phi_k^*(\alpha) = \begin{cases} 0 & \text{при } \alpha \leq \alpha_k^*; \\ 1 & \text{при } \alpha > \alpha_k^*, k=1, 2, \dots, N, \end{cases}$$

и имеет место равенство

$$\begin{aligned} \max_{\substack{\Phi_k \in \mathfrak{M} \\ k=1, N}} J(\Phi_1, \Phi_2, \dots, \Phi_N) &= \\ &= \max_{\substack{x_k \in R(+), \\ k=1, N}} \frac{A(x_1, x_2, \dots, x_N)}{B(x_1, x_2, \dots, x_N)} = \\ &= \frac{A(\alpha_1^*, \alpha_2^*, \dots, \alpha_N^*)}{B(\alpha_1^*, \alpha_2^*, \dots, \alpha_N^*)}. \end{aligned} \quad (20)$$

Замечания: 1. Теорема позволяет утверждать, что введение случайного эксперимента (рандомизации) в процесс принятия решений не улучшает стратегию.

2. Можно показать, что существование максимума функции $A(\cdot)/B(\cdot)$ есть необходимое и достаточное условие существования оптимальной стра-

тегии в классе нерандомизированных при несчетном пространстве управлений.

3. Определение оптимальной стратегии (определение величин α_k^*) сводится к исследованию на экстремум функции $A(\cdot)/B(\cdot)$.

Алгоритм решения задачи

1. Исходные данные: полумарковская матрица $\{Q_{ij}(t, \alpha)\}$; матрица функций дохода $\{R_{ij}(t, \alpha)\}$.

2. Последовательность решения задачи:

определяем матрицу $\{p_{ij}(\alpha)\}$

$$p_{ij}(\alpha) = \lim_{t \rightarrow \infty} Q_{ij}(t, \alpha);$$

решаем систему алгебраических уравнений относительно

$$\pi_k = \pi_k(\alpha_1, \dots, \alpha_{k-1},$$

$$\alpha_{k+1}, \dots, \alpha_N)$$

$$\pi_k = \sum_{i=1}^N p_{ik}(\alpha_i) \pi_i, \sum_{k=1}^N \pi_k = 1;$$

определяем средний доход за период

$$\rho_k = \rho_k(\alpha_k) \int_0^{\infty} \sum_{i=1}^N R_{ki}(t, \alpha_k) \times \\ \times d_t Q_{ki}(t, \alpha_k);$$

определяем среднюю длительность периода

$$T_k = T_k(\alpha_k) = \int_0^{\infty} \sum_{i=1}^N \tau d_{\tau} Q_{ki}(\tau, \alpha_k);$$

определяем функцию

$$\begin{aligned} A(\alpha_1, \alpha_2, \dots, \alpha_N) &= \\ &= \sum_{k=1}^N \rho_k(\alpha_k) \pi_k; \end{aligned}$$

определяем функцию

$$\begin{aligned} B(\alpha_1, \alpha_2, \dots, \alpha_N) &= \\ &= \sum_{k=1}^N T_k(\alpha_k) \pi_k; \end{aligned}$$

определяем максимум отношения

$$\begin{aligned} \max_{\substack{\alpha_k \in R(+), \\ k=1, N}} \frac{A(\alpha_1, \dots, \alpha_N)}{B(\alpha_1, \dots, \alpha_N)} = \\ = \frac{A(\alpha_1^*, \dots, \alpha_N^*)}{B(\alpha_1^*, \dots, \alpha_N^*)} \end{aligned}$$

и точку, в которой этот максимум достигается, $\alpha_1^*, \dots, \alpha_N^*$.

Вывод: Отношение

$$\frac{A(\alpha_1^*, \dots, \alpha_N^*)}{B(\alpha_1^*, \dots, \alpha_N^*)}$$

есть величина оптимального дохода, а $\alpha_1^*, \dots, \alpha_N^*$ — оптимальные управления, которые нужно принимать в состояниях соответственно e_1, \dots, e_N .

4. ЧАСТНЫЕ СЛУЧАИ

Управление однородным марковским процессом с конечным числом состояний. Однородный марковский процесс с конечным числом состояний и непрерывным временем определяется полумарковской матрицей $\{Q_{ij}(t, \alpha)\}$, для которой справедливо представление

$$\begin{aligned} Q_{ij}(t, \alpha) = p_{ij}(\alpha) [1 - e^{-\lambda_i(\alpha)t}], \\ \lambda_i(\alpha) > 0, \quad i, j \in [1, 2, \dots, N], \\ N < \infty. \end{aligned}$$

Замечание. Как следует из определения, время пребывания марковского процесса в состоянии e_i распределено по экспоненциальному закону с параметром $\lambda_i(\alpha)$, зависящим от номера состояния i и параметра α .

$$\begin{aligned} F_i(t, \alpha) = P(\tau(i) < t) = \\ = \sum_{j=1}^N Q_{ij}(t, \alpha) = 1 - e^{-\lambda_i(\alpha)t}, \end{aligned} \quad t > 0,$$

$$\text{так как } \sum_{j=1}^N p_{ij}(\alpha) = 1.$$

Алгоритм решения задачи

1. Исходные данные: матрица $(p_{ij}(\alpha))$; вектор $(\lambda_i(\alpha))$; матрица функций дохода $R_{ij}(t, \alpha)$.

2. Последовательность решения задачи:

решаем систему алгебраических уравнений относительно

$$\begin{aligned} \pi_k = \pi_k(\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \\ \alpha_{k+1}, \dots, \alpha_N) \end{aligned}$$

$$\pi_k = \sum_{i=1}^N p_{ik}(\alpha_k) \pi_i, \quad k = 1, 2, \dots, N,$$

$$\sum_{k=1}^N \pi_k = 1;$$

определяем средний удельный доход за период

$$\begin{aligned} \rho_k = \rho_k(\alpha_k) = \int_0^{\infty} \left[\sum_{j=1}^N R_{kj}(t, \alpha_k) \times \right. \\ \left. \times p_{kj}(\alpha_k) \right] \lambda_k(\alpha_k) e^{-\lambda_k(\alpha_k)t} dt; \end{aligned}$$

определяем среднюю длительность периода

$$T_k = T_k(\alpha_k) = \frac{1}{\lambda_k(\alpha_k)};$$

определяем функцию

$$A(\alpha_1, \dots, \alpha_N) = \sum_{k=1}^N \rho_k(\alpha_k) \pi_k;$$

определяем функцию

$$B(\alpha_1, \dots, \alpha_N) = \sum_{k=1}^N \frac{\pi_k}{\lambda_k(\alpha_k)};$$

определяем максимум отношения A и B

$$\begin{aligned} \max_{(\alpha_1, \dots, \alpha_N)} \frac{A(\alpha_1, \dots, \alpha_N)}{B(\alpha_1, \dots, \alpha_N)} = \\ = \frac{A(\alpha_1^*, \dots, \alpha_N^*)}{B(\alpha_1^*, \dots, \alpha_N^*)} \end{aligned}$$

и точку $(\alpha_1^*, \dots, \alpha_N^*)$, в которой он достигается.

Вывод: максимальный доход

$$\frac{A(\alpha_1^*, \dots, \alpha_N^*)}{B(\alpha_1^*, \dots, \alpha_N^*)}$$

достигается при управлении α_k^* в состоянии e_k .

Управление однородной цепью Маркова с конечным множеством состояний. Однородная цепь Маркова с конечным множеством состояний определяется полумарковской матрицей $\{Q_{ij}(t, \alpha)\}$, для которой справедливо представление

$$Q_{ij}(t, \alpha) = p_{ij}(\alpha) I(t),$$

$$i, j \in [1, 2, \dots, N], N < \infty,$$

$$I(t) = \begin{cases} 0 & t \leq 1; \\ 1 & t > 1. \end{cases}$$

Замечание. Для цепи Маркова длительность пребывания в фиксированном состоянии e_i есть неслучайная величина, равная единице (единице изменения времени для данного процесса).

Алгоритм решения задачи

1. Исходные данные: матрица вероятностей перехода $\{p_{ij}(\alpha)\}$; матрица функций дохода $\{R_{ij}(t, \alpha)\}$.

2. Последовательность решения задачи:

определяем стационарное распределение $\pi_k = \pi_k(\alpha_1, \dots, \alpha_{k-1}, \alpha_{k+1}, \dots, \alpha_N)$ как решение системы алгебраических уравнений

$$\pi_k = \sum_{i=1}^N p_{ik} \pi_i, \quad k = \overline{1, N}, \quad \sum_{k=1}^N \pi_k = 1;$$

определяем средний удельный доход за период (за один шаг)

$$\rho_k = \rho_k(\alpha_k) = \sum_{i=1}^N R_{ki}(1, \alpha_k) p_{ki}(\alpha_i);$$

определяем функцию

$$A(\alpha_1, \dots, \alpha_N) = \sum_{k=1}^N \rho_k(\alpha_k) \pi_k;$$

определяем максимум функции

$$\max_{(\alpha_1, \dots, \alpha_N)} A(\alpha_1, \dots, \alpha_N) = A(\alpha_1^*, \dots, \alpha_N^*)$$

и точку $(\alpha_1^*, \dots, \alpha_N^*)$, в которой он достигается.

Замечание. Для рассматриваемого случая $T_k = 1$ и $B(\alpha_1, \dots, \alpha_N) = 1$ при любых $\{\alpha_k\}$.

Вывод: максимальный доход $A(\alpha_1^*, \dots, \alpha_N^*)$ достигается при управлении α_k^* в состоянии e_k .

5. УПРАВЛЕНИЕ РЕГЕНЕРИРУЮЩИМ СЛУЧАЙНЫМ ПРОЦЕССОМ

Пусть заданы вероятностное пространство (X, \mathfrak{A}, P) и последовательность $\{\xi_k(t), \xi_k\}$ независимых при различных k ($k = 1, 2, \dots$) одинаково распределенных пар. Каждая пара (при любом k) характеризуется начальным распределением $P(\xi_k(0) \in A), A \in \mathfrak{A}$ и конечномерными распределениями, т. е. вероятностями вида

$$P(\xi_k(\tau_1) \in A_1, \xi_k(\tau_s) \in A_s, \tau_s < \xi_s < t),$$

определенными для любых $s > 0, A_s \in \mathfrak{A}, \tau_1 < \tau_2 < \dots < \tau_s$.

Регенерирующим процессом назовем процесс $\xi(t) = \xi_k(t - t_{k-1})$ при $t \in [t_{k-1}, t_k], t_k = \sum_{s=1}^k \xi_s$, ξ_s — периоды регенерации.

Если $X = \{e_0, e_1, \dots, e_N\}$ — конечное множество и $P(\xi_k(0) = e_0) = 1$, то для такого регенерирующего процесса $\xi(t)$ траектории есть ступенчатые функции и на каждом периоде регенерации процесс $\xi(t)$ стартует из состояния e_0 .

Такой регенерирующий процесс задается семейством распределений

$$F(\tau_1, \tau_s, e_{i_1}, \dots, e_{i_s}, \tau) = P(\xi(t_k + \tau_1) = e_{i_1}, \xi(t_k + \tau_s) = e_{i_s}, \xi_{k+1} > \tau) \quad (21)$$

для любого $s > 0$, любого набора e_{i_1}, \dots, e_{i_s} и набора $\{\tau_i\}$, для которого $\tau_1 < \tau_2 < \dots < \tau_s < \tau$.

Замечание. Так как последовательность $\{\xi_k(t), \xi_k\}$ имеет одинаково распределенные пары, то распределение (21) не зависит от номера k и t_k .

Управляемый регенерирующий процесс. Считаем, что вероятностные ха-

ра характеристики регенерирующего процесса $\xi(t)$ зависят от некоторого параметра (управления) $\alpha \in U$ с функцией распределения $F(\tau_1, \dots, \tau_s, e_{i_1}, e_{i_s}, \tau, \alpha)$, где U — пространство управлений (например, считаем $U = [0, +\infty)$), множество допустимых стратегий совпадает с множеством марковских однородных рандомизированных стратегий, определяемых распределением $\Phi(\alpha)$ с носителем $U = [0, \infty)$.

Для определения дискретного (ступенчатого) управления (см. общие определения) в силу марковости и однородности стратегии положим

$$\lambda_k(t/\xi(\tau_k), \tau_k) = \lambda_k(t - \tau_k/\xi(\tau_k), 0) = P(\xi_{k+1} < t - \tau_k), \quad (22)$$

$$v_k(B/\xi(\tau_k), \tau_k) = v_k(B/\xi(\tau_k) = e_0) = \int_B d\Phi(\alpha),$$

$$B \subset [0, +\infty).$$

Замечание. Указанное построение распределений λ_k и v_k , определяющих управление α в момент τ_k и момент t принятия последующего управления, означает, что управление изменяется в моменты регенерации случайного процесса $\xi(t)$.

Управляемый регенерирующий процесс $(\xi(t), \eta(t))$ определяется семейством распределений (21), зависящим от параметра $\alpha: F = F(\tau_1, \dots, \tau_s, \dots, \tau, \alpha)$, распределениями $\lambda_k(t/\xi(0) = e_0, \tau_k = 0) = P(\xi_1 < t)$ (распределение случайной величины ξ_1 зависит от α и определяется через функцию F), $v_k(B/\xi(0) = e_0, \tau_k = 0) = \int_B d\Phi(\alpha)$

и начальным распределением $P(\xi(0) = e_0) = 1$.

Описание эволюции управляемого регенерирующего процесса $(\xi(t), \eta(t))$:

1. При известном начальном значении $\xi(0) = e_0$ определяется управление $\eta(0) = \alpha$ в соответствии с распределением $\Phi(\alpha)$.

2. При фиксированном α определяется длительность $\xi_1 = t_1$ периода регенерации в соответствии с распределением $P(\xi_1 < t)$.

3. При заданном управлении α и длительности периода регенерации t_1 определяются любые конечномерные распределения процесса $\xi(t)$ на $[0, t_1]$.

4. Далее процесс принятия решений повторяется для момента t_1 , так как $\xi(t_1) = e_0$.

Функционал дохода. Обозначим через $X_i(\tau)$ суммарное время, которое процесс $\xi(t)$ провел в состоянии e_i за период регенерации, который имел длительность τ , т. е.

$$\sum_{j=1}^N X_j(\tau) = \tau.$$

Суммарный доход за период определим как функцию величин $X_i = X_i(\tau)$

$$R = R(X_1, X_2, \dots, X_N, \tau, \alpha).$$

Суммарный доход за несколько периодов регенерации определим как сумму доходов за каждый период регенерации (аддитивность дохода), т. е. доход $R(T)$ за период $[0, T]$ определяется равенством

$$R(T) = R(T, v(T), \alpha_1, \alpha_2,$$

$$\alpha_{v(T)+1}) = \sum_{i=1}^{v(T)} R(X_i^{(i)},$$

$$X_N^{(i)}, \tau_i, \alpha_i) + \tilde{R}(\tilde{X}_1, \dots, \tilde{X}_N, T - \sum_{i=1}^{v(T)} \tau_i, \alpha_{v(T)+1}),$$

где $v(T)$ — число периодов регенерации, закончившихся до момента T ; τ_i — длительность периодов регенерации; \tilde{X}_i — время пребывания регенерирующего процесса в состоянии e_i на

интервале $\left[\sum_{i=1}^{v(T)} \tau_i, T \right]$; \tilde{R} — доход за

$$\text{период } \left[\sum_{i=1}^{v(T)} \tau_i, T \right].$$

Качество стратегии управления, определяемой распределениями (22), характеризуется функционалом

$$J(\Phi, T) = MR(T) = MR(T, v(T), \alpha_1, \dots, \alpha_{v(T)+1}).$$

При бесконечной длительности управления $T \rightarrow \infty$ качество стратегии характеризуем пределом

$$J(\Phi) = \lim_{T \rightarrow \infty} \frac{J(\Phi, T)}{T}. \quad (23)$$

Структура функционала качества. Теорема. Функционал $J(\Phi)$, определяемый равенством (23), есть отношение математического ожидания дохода за период регенерации к средней длительности этого периода и имеет вид дробно-линейного функционала относительно распределения $\Phi(\alpha)$

$$J(\Phi) = \frac{MR(X_1, \dots, X_N)}{M\tilde{X}} = \frac{\int A(\alpha) d\Phi(\alpha)}{\int B(\alpha) d\Phi(\alpha)}, \quad (24)$$

где \tilde{X} — длительность периода регенерации; X_i — суммарное время пребывания процесса в состоянии e_i за период регенерации; $R(X_1, \dots, X_N)$ — доход за период регенерации, на котором времена пребывания в состояниях e_1, \dots, e_N равны соответственно X_1, \dots, X_N .

Замечания: 1. Функции $A(\alpha)$ и $B(\alpha)$ не зависят от распределения $\Phi(\alpha)$, а зависят от функции $R(\cdot)$.

2. Функция $A(\alpha)$ есть математическое ожидание дохода на периоде регенерации при условии, что на этом периоде принято решение α ; функция $B(\alpha)$ есть математическое ожидание длительности периода регенерации при том же условии.

Структура экстремальной функции. Экстремум функционала (24) достигается (если он существует) на вырожденной функции распределения

$$\Phi(\alpha) = \begin{cases} 0 & \text{при } \alpha \leq \alpha^*, \\ 1 & \text{при } \alpha > \alpha^* \end{cases}$$

[частный случай теоремы, см. (20)].

Вывод: оптимальную стратегию можно определить в классе нерандомизированных, т. е. в момент регенерации принимается одно и то же управле-

ние α^* , которое определяется как точка максимума функции

$$\frac{A(\alpha)}{B(\alpha)} : \frac{A(\alpha^*)}{B(\alpha^*)} = \max_{\alpha} \frac{A(\alpha)}{B(\alpha)},$$

величина максимума $\frac{A(\alpha^*)}{B(\alpha^*)}$ равна максимальному доходу.

Алгоритм решения задачи оптимизации сводится к определению функции $A(\alpha)$ — условного математического ожидания дохода за период регенерации при условии, что принято управление α ;

функции $B(\alpha)$ — условного математического ожидания длительности периода регенерации при условии, что принято управление α ;

точки α^* максимума функции $\frac{A(\alpha)}{B(\alpha)}$ и величины этого максимума $\frac{A(\alpha^*)}{B(\alpha^*)}$

6. ПРИМЕРЫ МОДЕЛЕЙ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ

6.1. МОДЕЛЬ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ СИСТЕМЫ БЕЗ УЧЕТА СТРУКТУРЫ

Описание модели функционирования. Пусть в начальный момент $t = 0$ начинается эксплуатация новой системы, у которой время безотказной работы ξ распределено по некоторому закону $F(x) = P(\xi < x)$, а время самостоятельной индикации отказа ζ по закону $Q(x) = P(\zeta < x)$. В момент $t = 0$ планируется проведение плановой предупредительной профилактики через случайное время η , распределенное по некоторому закону $\Phi(x) = P(\eta < x)$. Если к назначенному моменту система не отказала $\xi > \eta$, то проводится плановая предупредительная профилактика, которая начинается в момент η и длится в среднем время $T_{\text{пп}} = M\xi_{\text{пп}}$. По определению плановая предупредительная профилактика полностью обновляет системы. Если к назначенному моменту система отказала, но отказ не проявился $\xi \leq \eta < \xi + \zeta$, то факт отказа обнаруживается в назначенный момент η , и

начинается плановый аварийно-профилактический ремонт, средняя длительность которого равна $T_{ап} = M\bar{\xi}_{ап}$ и который также полностью обновляет систему. Наконец, если система отказала и отказ проявился до назначенного момента $\xi + \zeta \leq \eta$, то в момент обнаружения отказа начинается внеплановый аварийно-профилактический ремонт, средняя длительность которого равна $\bar{T}_{ап} = M\bar{\xi}_{ап}$ и после окончания которого система полностью обновляется. После окончания любой восстановительной работы система полностью обновлена, и весь процесс обслуживания повторяется независимо от прошлого.

Описание управляемого случайного процесса. Определим случайный процесс $\xi(t)$, характеризующий состояние системы в момент t , положив

$$\xi(t) = \begin{cases} E_1, & \text{если в момент } t \text{ система работоспособна;} \\ E_2, & \text{если в момент } t \text{ система простаивает в неработоспособном состоянии (скрытый отказ);} \\ E_3, & \text{если в момент } t \text{ в системе проводится плановая предупредительная профилактика;} \\ E_4, & \text{если в момент } t \text{ в системе проводится плановый аварийно-профилактический ремонт;} \\ E_5, & \text{если в момент } t \text{ в системе проводится внеплановый аварийно-профилактический ремонт;} \end{cases}$$

Определенный выше случайный процесс $\xi(t)$ является регенерирующим (моментами регенерации будут моменты попадания в состояние E_1), так как при окончании ремонта весь процесс обслуживания повторяется независимо от прошлого.

Из описания модели функционирования следует, что вероятностные распределения, определяющие управления имеют вид

$$\begin{aligned} \lambda_k(t/\xi(0) = E_1, \tau_k = 0) &= \\ &= P(X_1 + X_2 + \dots + X_5 < t); \\ \nu_k(B/\xi(0) = E_1, \tau_k = 0) &= \int_B d\Phi(\alpha), \end{aligned}$$

где X_i — времена пребывания в состоянии E_i на периоде регенерации определяются равенствами:

$$\begin{aligned} X_1 &= \min(\xi, y), \\ X_2 &= \max(0, \min(\zeta, y - \xi)), \\ X_3 &= \begin{cases} 0, & \text{если } \xi \leq y, \\ \xi_{шт}, & \text{если } \xi > y, \end{cases} \\ X_4 &= \begin{cases} 0, & \text{если } \xi \geq y \text{ или } \xi + \zeta < y, \\ \xi_{ап}, & \text{если } \xi < y \leq \xi + \zeta, \end{cases} \\ X_5 &= \begin{cases} 0, & \text{если } \xi + \zeta \geq y, \\ \bar{\xi}_{ап}, & \text{если } \xi + \zeta < y, \end{cases} \end{aligned}$$

если на этом периоде регенерации принято управление y .

Структура дохода. Считаем

$$\begin{aligned} R(X_1, X_2, \dots, X_5) &= \\ &= \sum_{k=1}^5 c_k(X_k), \end{aligned}$$

где $c_k(t)$ — некоторые функции.

Вычисление функционала:

$$J(\Phi) = \frac{\int_0^{\infty} A(y) d\Phi(y)}{\int_0^{\infty} B(y) d\Phi(y)},$$

где

$$\begin{aligned} A(y) &= c_1(y) \bar{F}(y) + \\ &+ \int_0^y c_1(t) dF(t) + \int_0^y c_2(t) \times \\ &\times \left[\int_t^{\infty} f(t-y) dQ(x) + \right. \\ &\left. + \int_0^{y-t} f(x) dx q(t) \right] dt + c_3 \bar{F}(y) + \\ &+ c_4 \int_0^y \int_{y-x}^{\infty} dF(x) dQ(z) + \\ &+ c_5 \int_0^y F(y-x) dQ(x), \end{aligned}$$

$$B(y) = \int_0^y \left[1 - \int_0^x F(y-z) dQ(z) \right] \times \\ \times dx + M\xi_{\text{пп}} + (M\xi_{\text{ап}} - M\xi_{\text{пп}}) \times \\ \times F(y) + (M\bar{\xi}_{\text{ап}} - M\xi_{\text{ап}}) \times \\ \times \int_0^y F(y-x) dQ(x),$$

$$\bar{F}(x) = 1 - F(x), \quad c_k = \int_0^{\infty} c_k(t) dF_k(t),$$

$$F_3(t) = P(\xi_{\text{пп}} < t), \quad F_4(t) = P(\xi_{\text{ап}} < t),$$

$$F_5(t) = P(\bar{\xi}_{\text{ап}} < t), \quad q(t) = Q'(t).$$

Решение задачи оптимизации. Определяется $\max_{\alpha} \frac{A(\alpha)}{B(\alpha)}$ и точка α^* , в которой он достигается.

6.2. МОДЕЛЬ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ ДУБЛИРОВАННОЙ СИСТЕМЫ

Описание модели функционирования. Пусть основной элемент в дублированной системе с холодным резервом имеет длительность безотказной работы ξ с функцией распределения $F(x) = P(\xi < x)$. Появившийся отказ основного элемента самостоятельно проявляется через случайное время ζ , распределенное по закону $\Phi(x)$ (случайные величины ξ и ζ считаем независимыми). В системе возможно проведение трех видов восстановительных работ: плановой предупредительной профилактики, длительность которой ξ_0 распределена по закону $F_0(x)$, внепланового или планового аварийно-профилактического ремонта, длительность которого ξ_1 имеет распределение $F_1(x)$, и внеплановой предупредительной профилактики, длительность которой ξ_2 имеет распределение $F_2(x)$. Для проведения восстановительных работ имеется одна бригада, т. е. в произвольный момент времени может восстанавливаться не более одного элемента.

Кроме того, если в какой-то момент нужно начать новую восстановительную работу (плановую или аварийную), но нет резервного элемента (он еще ремонтируется), то начало новой вос-

становительной работы задерживается.

Обозначим через t_k ($k = 0, 1, \dots$) моменты начала восстановительных работ в исследуемой системе (один элемент начинает восстанавливаться, второй начинает работать), $z_k = t_{k+1} - t_k$ — интервалы между этими моментами.

Пусть в некоторый момент t_k начинается восстановление i -го вида для одного элемента, а второй начинает работать, тогда $\min(\eta_i, \xi + \zeta)$ есть интервал времени, через которое нужно начать следующее восстановление (в момент $\xi + \zeta$ будет известно наличие отказа, в момент η_i наступает срок проведения плановой работы), а ξ_i — интервал, через который можно начать новую восстановительную работу (будет новый работоспособный резервный элемент). Тогда $Z_k^{(i)} = \max(\xi_i, \min(\eta_i, \xi + \zeta))$, где $Z_k^{(i)}$ — длительность интервала между соседними моментами начала восстановительных работ при условии, что в начале периода проводилась восстановительная работа i -го вида.

В зависимости от соотношений величин $\xi, \xi_i, \eta_i, \zeta$ в конце периода $Z_k^{(i)}$ начинаются различные восстановительные работы и, следовательно, начинаются различные периоды новых плановых работ:

1. Если $\xi > \eta_i > \xi_i$, то в момент η_i начинается плановая предупредительная профилактика работающего элемента, так как он к этому моменту не отказал и есть работоспособный резервный элемент. Этот резервный элемент начинает выполнять функции основного, а ранее функционировавший элемент идет на плановую предупредительную профилактику.

2. Если $\xi_i > \xi > \eta_i$, то в момент η_i нет работоспособного резервного элемента, поэтому запланированная восстановительная работа переносится, но происходит отказ второго (основного) элемента (отказ всей системы), система простаивает в неработоспособном состоянии время $\xi_i - \xi$ до появления работоспособного резервного элемента, в момент ξ_i заканчивается ремонт резервного элемента и он начинает выполнять функции основного, а не-

работоспособный элемент идет во внеплановый аварийно-профилактический ремонт.

3. Если $\xi > \xi_i > \eta_i$, то в η_i нет работоспособного резервного элемента, поэтому запланированная восстановительная работа переносится, в момент ξ_i заканчивается восстановление резервного элемента, работоспособный элемент идет на внеплановую предупредительную профилактику, а резервный начинает выполнять функции основного.

4. Если $\eta_i > \xi + \zeta > \xi_i$, то в момент $\xi + \zeta$ мы узнаем об отказе основного элемента, и так как к этому моменту закончился ремонт резервного элемента, начинается внеплановый аварийно-профилактический ремонт отказавшего элемента, а другой начинает выполнять функции основного.

5. Если $\min(\eta_i, \xi_i) > \xi + \zeta$, то в момент $\xi + \zeta$ мы узнаем об отказе основного элемента, но начать ремонт не можем, так как еще продолжается ремонт другого элемента. Внеплановый аварийно-профилактический ремонт начинается в момент ξ_i , когда закончится ремонт резервного элемента.

6. Если $\xi < \eta_i < \min(\xi_i, \xi + \zeta)$, то в момент η_i мы узнаем об отказе основного элемента, но ремонт начать не можем до момента ξ_i , когда появится работоспособный резервный элемент, в момент ξ_i начинается внеплановый аварийно-профилактический ремонт.

7. Если $\max(\xi, \xi_i) < \eta_i < \xi + \zeta$, то в момент η_i узнаем об отказе основного элемента и начинаем плановый аварийно-профилактический ремонт, для простоты отождествляя его с внеплановым аварийно-профилактическим ремонтом.

Во всех описанных случаях в момент ухода одного элемента в ремонтную бригаду для другого нового элемента, начинающего выполнять функции основного, назначается проведение плановой предупредительной профилактики через случайное время η_i с распределением $G_i(x)$, $i = 0, 1, 2$, причем $i = 0$, если начинается плановая предупредительная профилактика; $i = 1$, если начинается внеплановая предупредительная профилактика; $i = 2$,

если начинается внеплановый аварийно-профилактический ремонт.

Введем случайный процесс $\xi(t)$, принимающий три значения,

$$\xi(t) = \begin{cases} E_0 & \text{при } t_i \leq t < t_{i+1}, \text{ если} \\ & \text{устройство ушло на плановую} \\ & \text{предупредительную профилактику в} \\ & \text{момент } t_i; \\ E_2 & \text{при } t_i \leq t < t_{i+1}, \text{ если} \\ & \text{устройство ушло во внеплановый} \\ & \text{аварийный ремонт в момент } t_i; \\ E_1 & \text{при } t_i \leq t < t_{i+1}, \text{ если} \\ & \text{устройство ушло во внеплановую} \\ & \text{предупредительную профилактику} \\ & \text{в момент } t_i. \end{cases}$$

Значения процесса в моменты t_1, t_2, \dots образуют цепь Маркова, распределения длительности периодов $Z_k = t_{k+1} - t_k$ зависят от значений процесса в моменты t_k . Таким образом, процесс $\xi(t)$ есть полумарковский, причем интервалы $Z_k^{(i)}$ зависят от управления η_i , т. е. $\xi(t)$ — управляемый полумарковский процесс.

Структура дохода. Между двумя соседними моментами t_k — начала восстановительных работ могут происходить отказы элемента, могут отказы проявляться, т. е. элементы могут находиться в различных состояниях. Доход на периоде $Z_k = t_{k+1} - t_k$ свяжем с состоянием элементов (устройств).

Рассмотрим далее процесс $\xi(t)$, характеризующий состояния элементов исследуемой системы, в произвольный момент t . Состояние это зависит от состояния ремонтируемого устройства (либо оно ремонтируется, либо находится в резерве) и от состояния основного устройства (либо оно работает исправно, либо скрытый отказ, либо явный — обнаруженный отказ). В таком случае возможны следующие состояния элементов дублированной системы:

- e_1 — основное устройство работает исправно, другое устройство ремонтируется;
- e_2 — основное устройство работает исправно, другое находится в резерве;

- e_3 — у основного устройства скрытый отказ, другое устройство ремонтируется;
- e_4 — у основного устройства скрытый отказ, другое устройство находится в резерве;
- e_5 — у основного устройства обнаружен отказ, другое устройство ремонтируется.

Через τ_i обозначим случайное время, проведенное процессом $\xi(t)$ в состоянии e_i на периоде между двумя соседними моментами начала ремонтов.

Далее предположим, что прибыль и потери пропорциональны времени, проведенному элементами в различных состояниях.

Пусть $c^{(0)}$ — прибыль, получаемая за единицу времени исправного функционирования;

c_i — потери за единицу времени, проведенного устройством во внеплановом аварийном, или внеплановом, или плановом профилактическом ремонте, $i = 0, 1, 2$;

\bar{c}_i — потери за единицу времени, проведенного устройством в состоянии скрытого или явного отказа, $i = 0, 1$.

Из описания задачи следует

$$\begin{aligned} \tau_1 &= \min(\xi, \xi_i), \\ \tau_2 &= \max(0, \min(\xi, \eta_i) - \xi_i), \\ \tau_3 &= \max(0, \min(\xi_i, \xi + \zeta) - \xi), \\ \tau_4 &= \max(0, \min(\eta_i, \xi + \zeta) - \max(\xi_i, \xi)), \\ \tau_5 &= \max(0, \xi_i - \xi - \zeta). \end{aligned}$$

Тогда прибыль на периоде $[t_k, t_{k+1}]$ определим как сумму

$$\begin{aligned} R_{ij}(Z_k, \eta_i) &= R_{ij}(\tau_1, \tau_2, \dots, \tau_5, \eta_i) = \\ &= c^{(0)}\tau_1 - c_i\tau_1 + c^{(0)}\tau_2 - \\ &- (c_i + \bar{c}_0)\tau_3 - c^{(0)}\tau_4 - (c_i + \bar{c}_1)\tau_5 = \\ &= c^{(0)}(\tau_1 + \tau_2) - c_i(\tau_1 + \tau_3 + \tau_5) - \\ &- \bar{c}_0(\tau_3 + \tau_4) - \bar{c}_1\tau_5. \end{aligned}$$

Решение задачи оптимизации. В соответствии с алгоритмом, изложенным в параграфе 3, для нерандомизированной стратегии ($\eta_i = y_i$, т. е. $G_i(x) = \begin{cases} 0, & x \leq y_i \\ 1, & x > y_i \end{cases}$). Определяем:

1. Матрицу переходных вероятностей

$$\begin{aligned} p_{i0}(y_i) &= P(\xi > y_i > \xi_i) = \\ &= \bar{F}(y_i)F_i(y_i), \end{aligned}$$

$$\begin{aligned} p_{i2}(y_i) &= P(\xi > \xi_i > y_i) = \\ &= \int_{y_i}^{\infty} \bar{F}(x) dF_i(x), \end{aligned}$$

$$\begin{aligned} p_{i1}(y_i) &= 1 - p_{i0}(y_i) - p_{i2}(y_i) = \\ &= 1 - \bar{F}(y_i)F_i(y_i) - \int_{y_i}^{\infty} \bar{F}(x) dF_i(x). \end{aligned}$$

2. Решаем систему алгебраических уравнений

$$\begin{aligned} \pi_k &= \sum_{i=0}^2 p_{ik}(y_i) \pi_i, \quad k = 1, 2 \\ \sum_{i=0}^2 \pi_k &= 1. \end{aligned}$$

Если обозначить через $|P|$ определитель матрицы

$$P = \begin{pmatrix} p_{00} - 1 & p_{10} & p_{20} \\ p_{01} & p_{11} - 1 & p_{21} \\ 1 & 1 & 1 \end{pmatrix},$$

$$\text{то } \pi_0 = \frac{1}{|P|} (p_{10}p_{21} + (1 - p_{11})p_{20}),$$

$$\pi_1 = \frac{1}{|P|} ((1 - p_{00})p_{21} + p_{01}p_{20}),$$

$$\pi_2 = \frac{1}{|P|} ((1 - p_{00})(1 - p_{11}) - p_{01}p_{10}).$$

3. Определяем средний доход за период

$$\begin{aligned} \rho_i &= MR_{ij}(\tau_1, \tau_5, y_i) = \\ &= c^{(0)}(M\tau_1 + M\tau_2) - c_i(M\tau_1 + \\ &+ M\tau_3 + M\tau_5) + \bar{c}_0(M\tau_3 + M\tau_4) - \\ &- \bar{c}_1M\tau_5, \end{aligned}$$

где

$$\begin{aligned} M\tau_1 + M\tau_2 &= \int_0^{\infty} \bar{F}(x) dx - \int_{y_i}^{\infty} \bar{F}(x) \times \\ &\times F_i(x) dx, \end{aligned}$$

$$M\tau_1 + M\tau_3 + M\tau_5 = \int_0^{\infty} \bar{F}_i(x) dx,$$

$$M\tau_6 + M\tau_8 = \int_0^{\infty} \int_0^{\infty} \Phi(z) \bar{F}_i(x+z) \times$$

$$\times dF(x) dz + \int_0^{y_i} \left[F_i(x) \times \right. \\ \left. \times \int_0^x \Phi(x-z) dF(z) \right] dx,$$

$$M\tau_5 = \int_0^{\infty} \int_0^{\infty} \Phi(z) \bar{F}_i(x+z) dF(x) dz.$$

4. Определяем среднюю длительность периода

$$T_i = MZ_k^{(i)} = \\ = \int_0^{\infty} \bar{F}_i(x) dx + \int_0^{y_i} F_i(x) \left[1 - \right. \\ \left. - \int_0^x F(x-u) d\Phi(u) \right] dx.$$

5. Определяем функцию

$$A(y_0, y_1, y_2) = \sum_{i=0}^2 \rho_i \pi_i.$$

6. Определяем функцию

$$B(y_0, y_1, y_2) = \sum_{i=0}^2 T_i \pi_i.$$

7. Определяем максимум отношения $\frac{A(y_0, y_1, y_2)}{B(y_0, y_1, y_2)}$ и точку, в которой максимум достигается, (y_0^*, y_1^*, y_2^*) .

Отличительной особенностью данных о надежности оборудования в эксплуатации является то, что они отражают поведение системы в целом, а не только отдельных ее компонент. Предпосылкой успешного статистического анализа таких данных являются:

учет и объединение информации об испытании на надежность отдельных компонент системы с данными об ее эксплуатации в целом;

подбор простой математической модели, описывающей функционирование системы и ее отдельных компонент.

1. АНАЛИЗ НАДЕЖНОСТИ ПОКАЗАТЕЛЬНЫХ ПОСЛЕДОВАТЕЛЬНЫХ СИСТЕМ С НЕЗАВИСИМЫМИ КОМПОНЕНТАМИ (СЛУЧАЙ ДВУХ КОМПОНЕНТ)

Пусть время жизни (безотказной работы) τ_i элемента l_i случайно и имеет показательное распределение с параметром λ_i , $i = 1, 2$. Предположим, что из элементов l_1 и l_2 составлена система последовательного типа, т. е. время жизни системы $\tau_0 = \min(\tau_1, \tau_2)$ и надежность системы

$$R(t, \bar{\lambda}) = \exp\{-\lambda_0 t\}, \quad \lambda_0 = \lambda_1 + \lambda_2, \quad \bar{\lambda} = (\lambda_1, \lambda_2).$$

Предположим, что элементы l_i испытываются по плану $[n_i, B, n_i]$, $i = 1, 2$, а система в целом испытывается по плану $[n_0, B, n_0]$. При этом неизвестно, какой из элементов l_i отказал. Суммарные наработки S_0 , S_1 и S_2 системы и элементов образуют достаточную статистику. Совместная плотность вероятностей $(S_0, S_1$ и $S_2)$ имеет вид

$$\begin{aligned} p(S_0, S_1, S_2; \bar{\lambda}) &= \\ &= \left(\prod_{i=0}^2 \frac{\lambda_i^{n_i} S_i^{n_i-1}}{\Gamma(n_i)} \right) \times \\ &\times \exp\{-\lambda_0 S_0 - \lambda_1 S_1 - \lambda_2 S_2\}, \end{aligned} \quad (1)$$

где $\lambda_0 = \lambda_1 + \lambda_2$.

Если необходимо проверить гипотезу $H_0: R(T) > R_0$ против альтернативы $H_1: R(T) < R_1$, то можно перейти к проверке эквивалентных гипотез $H'_0: \lambda_0 < \lambda_{00}$ против альтернативы $H'_1: \lambda_0 > \lambda_{01}$.

В дальнейшем удобно от параметров (λ_1, λ_2) перейти к параметрам (λ_0, λ_2) и рассматривать задачу проверки гипотезы H'_0 против H'_1 в присутствии мешающего параметра λ_2 .

Плотность вероятностей (1) переписывается следующим образом:

$$\begin{aligned} p(S_0, S_1, S_2; \bar{\lambda}) &= \\ &= \frac{\lambda_0^{n_0} (\lambda_0 - \lambda_2)^{n_1} \lambda_2^{n_2} S_0^{n_0-1} S_1^{n_1-1} S_2^{n_2-1}}{\Gamma(n_0) \Gamma(n_1) \Gamma(n_2)} \times \\ &\times \exp\{-\lambda_0 (S_0 + S_1) - \lambda_2 (S_2 - S_1)\}. \end{aligned} \quad (2)$$

Из (2) получаем, что минимальной достаточной статистикой будет вектор-строка $(S_0 + S_1, S_2 - S_1)$. Условная плотность распределения суммы $(S_0 + S_1) | (S_2 - S_1 = v)$ не будет зависеть от λ_2 , а потому статистику $S_2 - S_1$ можно рассматривать как подчиненную. Таким образом, для проверки H'_0 против H'_1 можно будет предложить условный критерий при условии $S_2 - S_1 = v$, $-\infty < v < +\infty$.

Обозначим $p(u | v; \lambda_0)$ условную плотность распределения случайной величины $S_0 + S_1$ при условии $S_2 -$

— $S_1 = v$. После стандартных вычислений получим

$$p(u/v; \lambda_0) = \sum_{j=0}^{n_2-1} \alpha_j \gamma(u, n_0 + n_1 + j, \lambda_0), \quad (3)$$

где

$$\alpha_j = a_j \left/ \sum_{r=0}^{n_2-1} a_r, j = 0, \dots, n_2 - 1; \right. \quad (4)$$

$$a_r = C_{n_2-1}^r (\lambda_0 v)^{-r} \Gamma(n_1 + r); \quad (5)$$

$$\gamma(u, r, \lambda) = \frac{\lambda^r u^{r-1} \exp\{-\lambda u\}}{\Gamma(r)}, \quad (6)$$

$\Gamma(r)$ — гамма-функция, $r = 0, 1, \dots, n_2 - 1$.

Плотность распределения $p(u|v; \lambda_0)$ принадлежит однопараметрическому семейству распределений экспоненциального типа. Отсюда вытекает, что при проверке гипотезы H'_0 против H'_1 существует несмещенный равномерно наиболее мощный критерий. Гипотеза H'_0 $\lambda_0 < \lambda_0^b$ будет отвергаться в пользу H'_1 на уровне α , когда $\{S_0 + S_1 < f_\alpha(v, \lambda_0, n_0, n_1, n_2)\}$,

где $f_\alpha(v, \lambda_0, n_0, n_1, n_2)$ — α -квантиль условного распределения статистики $(S_0 + S_1 | S_2 - S_1 = v)$. Пусть

$$F(u, v, n_2, n_1, n_0, \lambda_0) = \int_0^u p(z|v, \lambda_0) \times$$

$\times dz$ — условная функция распределения статистики $(S_0 + S_1 | S_2 - S_1 = v)$. Из (3)–(6) получаем, что

$$F(u, v, n_2, n_1, n_0, \lambda_0) = \sum_{j=0}^{n_2-1} \alpha_j \Gamma(u, n_0 + n_1 + j, \lambda_0), \quad (7)$$

где $\Gamma(u, r, \lambda) = \int_0^u \gamma(z, r, \lambda) dz$ — функция гамма-распределения.

Представление $F(u, v, n_2, n_1, n_0, \lambda_0)$ в виде взвеси от гамма-распределений предоставляет широкие возможности для аппроксимации этой функции и

связанных с ней характеристик посредством соответствующим образом подобранного нормального распределения, рядов Эджворта или хи-квадрат распределениями.

Как правило, эти аппроксимации получают на основе приравнивания моментов исходного и аппроксимирующего распределения. Приведем два первых момента статистики $T_1 = S_1 + S_0$ при условии, что $T_2 = S_2 - S_1 = v, v > 0$:

$$E(T_1 | T_2 = v) = \sum_{j=0}^{n_2-1} \alpha_j \frac{(n_0 + n_1 + j)}{\lambda_0};$$

$$E(T_1^2 | T_2 = v) = \sum_{j=0}^{n_2-1} \alpha_j \times \frac{(n_0 + n_1 + j)(n_0 + n_1 + j + 1)}{\lambda_0^2}, \quad (8)$$

величины α_j определены в (4).

Из представления (7) легко получить неравенство

$$\begin{aligned} \Gamma(u, n_0 + n_1, \lambda_0) &\geq \\ &\geq F(u, v, n_2, n_1, n_0, \lambda_0) \geq \\ &\geq \Gamma(u, n_0 + n_2 + n_1 - 1, \lambda_0). \end{aligned} \quad (9)$$

Отсюда же получаем неравенство для квантилей

$$\begin{aligned} \gamma^{-1}(\beta, n_0 + n_1, \lambda_0) &\leq \\ &\leq f_\beta(v, \lambda_0, n_0, n_1, n_2) \leq \\ &\leq \gamma^{-1}(\beta, n_0 + n_1 + n_2 - 1, \lambda_0). \end{aligned} \quad (10)$$

В (10) $\gamma^{-1}(\beta, r, \lambda)$ — β -квантиль гамма-распределения $\Gamma(u, r, \lambda)$. Для расчета квантилей $\gamma^{-1}(\beta, r, \lambda)$ можно воспользоваться таблицами, программами или аппроксимационными формулами для квантилей гамма-распределения.

Функция распределения $F(u, n_2, n_1, n_0, \lambda_0)$ является возрастающей функцией λ_0 . Отсюда верхняя β -доверительная граница λ_0^b для λ_0 получится как решение уравнения

$$F(T_1, n_2, n_1, n_0, \lambda_0^b) = \beta.$$

На основе неравенств (9) и (10) легко получить

$$\lambda_0^{1B} \leq \lambda_0^B \leq \lambda_0^{2B},$$

где

$$\Gamma(T_1, n_0 + n_1, \lambda_0^{1B}) = \beta;$$

$$\Gamma(T_1, n_0 + n_1 + n_2 - 1, \lambda_0^{2B}) = \beta.$$

Таким образом, задача построения оптимальных критериев и доверительных границ для λ_0 решена.

Иногда на основе наблюдений (S_0, S_1, S_2) необходимо построить доверительные границы для $\lambda = c_1\lambda_1 + c_2\lambda_2$, $c_i > 0$. При c_i — целых величина λ — интенсивность отказа последовательной системы, состоящей из c_1 элемента l_1 и c_2 элементов l_2 . Для вывода соответствующих этому случаю формул необходимо от параметров (λ_1, λ_2) перейти в (1) к параметрам $(\bar{\lambda}, \lambda_2)$.

Для случая, когда последовательная система состоит из k элементов, обобщение делается стандартным образом.

Если распределение времен жизни τ_i отличается от показательного или планы испытаний таковы, что плотность распределения вероятностей данных об испытании системы и ее компонент не будет принадлежать экспоненциальному семейству с естественной параметризацией (последний случай имеет место, когда при испытании последовательной системы можно установить, какой именно элемент отказал), то изложенный метод не проходит. В качестве общего подхода к построению статистических выводов рекомендуется использовать метод вклада выборки. Кратко его можно описать следующим образом.

Пусть $p(x, \bar{\theta})$ — плотность распределения вероятностей данных x об испытании системы на надежность, $\bar{\theta} = (\theta_1, \dots, \theta_m)$ — параметр, $\bar{\theta} \in \Theta \subset R^m$. Всегда можно предполагать, что θ_1 — интересующий нас параметр, а все остальные — мешающие параметры. Статистика $\lambda(x, \bar{\theta}) = \text{grad}_{\bar{\theta}} \ln p(x, \bar{\theta})$ носит название вклада выборки. Ее компонента $\lambda_1(x, \theta_1, \theta_2, \dots, \theta_m) = \frac{\partial}{\partial \theta_1} \ln p(x, \bar{\theta})$ — минимально достаточная статистика. При фиксиро-

ванном $\bar{\theta}$ распределение $\lambda_1(x, \theta_1, \theta_2, \dots, \theta_m)$ часто удается приблизить нормальным распределением. Пусть $\hat{\theta}_j$ — состоятельные оценки для θ_j , $j = 2, \dots, m$. При фиксированном θ_1 иногда также удается найти достаточно точное приближенное распределение для величины $\lambda_1(x, \theta_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$. В этом случае $\lambda_1(x, \theta_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$ можно использовать для построения доверительных интервалов для θ_1 и проверки гипотез о значении параметра θ_1 .

2. АНАЛИЗ НАДЕЖНОСТИ ПОКАЗАТЕЛЬНЫХ ПОСЛЕДОВАТЕЛЬНЫХ СИСТЕМ С ЗАВИСИМЫМИ КОМПОНЕНТАМИ

Пусть время жизни τ_i компоненты l_i случайно и имеет показательное распределение с параметром λ_i , $i = 1, \dots, m$. Функцией надежности (ф. н.), отвечающей совместному распределению величин (τ_1, \dots, τ_m) , называют

$$R(t_1, \dots, t_m) = P \left\{ \prod_{j=1}^m (\tau_j > t_j) \right\}. \tag{11}$$

Положим $\bar{t} = (t_1, \dots, t_m)^T$, $t_i \geq 0$; Λ — упорядоченный набор параметров λ_j ; G — совокупность всех подмножеств множества $\{1, 2, \dots, m\}$, $J \in G$. Если

$$R(t_1, \dots, t_m) = \bar{E}_m(\bar{t}, \Lambda) = \exp \left\{ - \sum_{J \in G} \lambda_J \max_{i \in J} t_i \right\}, \tag{12}$$

то соответствующее распределение \bar{E}_m назовем m -мерным экспоненциальным. Его функцию распределения обозначим $E_m(\bar{t}, \Lambda)$. Это единственное m -мерное распределение, у которого все одномерные маргинальные распределения показательны и которое вместе со всеми маргинальными распределениями удовлетворяет следу-

ющему многомерному «свойству отсутствия памяти»:

$$P \left\{ \prod_{i=1}^m (\tau_i > t_i + \Delta) \mid \prod_{i=1}^m (\tau_i > t_i) \right\} = P \left\{ \prod_{i=1}^m (\tau_i > \Delta) \right\}, \quad (13)$$

где $\Delta > 0$.

Зависимость между τ_i для распределения E_m весьма специфична. Рассмотрим следующую модель с множественными отказами. Пусть на систему из элементов $\{l_1, \dots, l_m\}$ поступают независимые друг от друга пуассоновские потоки $z_j(v, \lambda_j)$ отказов с интенсивностями λ_j , v — параметр времени, $v > 0$. Момент отказа элемента l_j определяется моментом наступления первого события суммарного потока $\sum_{j \in G(i)} z_j(v, \lambda_j)$, где $G(j) =$

$\{ \cup J : J \in G, j \in J \}$. В этом случае $\bar{E}_m(\bar{t}, \Lambda)$ — совместная ф. н. времен жизни τ_i элементов l_1, \dots, l_m . Существуют и более сложные модели, приводящие к ф. н. $\bar{E}_m(\bar{t}, \Lambda)$.

Перечислим некоторые свойства распределения с ф. н. $\bar{E}_m(\bar{t}, \Lambda)$.

1. Преобразование Лапласа —

Стильтьеса

Пусть $m = 2$. Имеем

$$\int_0^\infty \int_0^\infty \exp \{ -t_1 u_1 - t_2 u_2 \} \times \\ \times dE_m((u_1, u_2)^T, \Lambda) = \\ = \frac{(\lambda + t_1 + t_2)(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12}) + t_1 t_2 \lambda_{12}}{(\lambda + t_1 + t_2)(\lambda_1 + \lambda_{12} + t_1) \times (\lambda_2 + \lambda_{12} + t_2)}, \quad (14)$$

где $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$.

Из (14) получаем формулы моментных характеристик времен жизни τ_i :

$$E\tau_i = \gamma_i^{-1}; \quad D\tau_i = \gamma_i^{-2};$$

$$E\tau_1^i \tau_2^j = j\Gamma(i+1) \times$$

$$\times \sum_{k=0}^{i-1} \frac{\Gamma(j+k)}{\Gamma(k+1) \gamma_1^{i-k} \lambda^{j+k}} + i\Gamma(j+1) \sum_{k=0}^{j-1} \frac{\Gamma(i+k)}{\Gamma(k+1) \gamma_2^{j-k} \lambda^{i+k}}, \quad (15)$$

где $\gamma_i = \lambda_i + \lambda_{12}$, $i = 1, 2$.

Коэффициент корреляции $\rho = \lambda_{12}/\lambda$.

2. Распределение экстремальных значений

Пусть $\tau_{(1)} = \min \tau_i$ имеет ф. р. $F_{(1)}(u)$ и $\tau_{(m)} = \max_{i=1, \dots, m} \tau_i$ имеет ф. р. $F_{(m)}(u)$. Тогда

$$F_{(1)}(u) = \bar{E}_m(u\bar{e}, \Lambda) = \exp \{ -\lambda u \}, \\ \bar{e} = (1, \dots, 1)^T, \quad \lambda = \sum_{j \in G} \lambda_j;$$

$$F_{(m)}(u) = E_m(u\bar{e}, \Lambda).$$

3. Свертка распределений E_m

При $m = 2$ k -кратная свертка

$$E_2^{(k)}(\bar{t}; \Lambda) = k \sum \frac{\lambda_{12}^{k_0} \lambda_1^{k_1} \lambda_2^{k_2}}{k_0! k_1! k_2!} \times \\ \times \sum_{i > k_2} \sum_{j > k_1} \frac{\bar{\lambda}_1^i \bar{\lambda}_2^j}{i! j!} \exp \{ -\bar{\lambda}_1 t_1 - \bar{\lambda}_2 t_2 \} \times \\ \times S_{ij}^k(\bar{t}, \lambda_{12}), \text{ для } t_1 \leq t_2 \\ t = (t_1, t_2)^T, \bar{\lambda}_i = \lambda_i + \lambda_{12}, i = 1, 2, \quad (16)$$

где суммирование во внешней сумме производится для всех натуральных k_0, k_1, k_2 , таких, что $k_0 + k_1 + k_2 = k$,

$$S_{ij}^k(\bar{t}, \alpha) = \sum_{l=0}^j (-1)^{j-l} C_j^l \times \\ \times B(k+j-l, i+1) t_1^{k+i+j-l} \times \\ \times t_2^l \Phi(k+j-l, k+i+j-l+1, \alpha t_1),$$

где $\Phi(a, b, z)$ — вырожденная гипергеометрическая функция.

4. Сингулярность многомерного показательного распределения E_m

Если в (12) $\lambda_j > 0, J \in G$, то с положительной вероятностью все τ_i , для которых $i \in J$, примут одинаковые значения. Это означает, что у соответствующего распределения E_m имеется сингулярная часть. В частности, при $m = 2$ имеет место следующее разложение:

$$\bar{E}_2(\bar{t}, \Lambda) = \frac{\lambda_1 + \lambda_2}{\lambda} \bar{F}_a(\bar{t}) + \frac{\lambda_{12}}{\lambda} \bar{F}_c(\bar{t}), \quad (17)$$

где $\bar{F}_c(\bar{t}) = \exp\{-\lambda \max(t_1, t_2)\}$ — сингулярная часть и

$$\bar{F}_a(\bar{t}) = \frac{\lambda}{\lambda_1 + \lambda_2} \exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)\} - \frac{\lambda_{12}}{\lambda_1 + \lambda_2} \exp\{-\lambda \max(t_1, t_2)\} \quad (18)$$

— абсолютно непрерывная часть распределения E_2 .

5. Условные распределения

Пусть $m = 2$, тогда

$$P\{\tau_2 > t_2 | \tau_1 = t_1\} = \begin{cases} \exp\{-\lambda_2 t_2\} & \text{при } t_1 > t_2; \\ \frac{\lambda_1}{\lambda} \exp\{-\lambda_{12}(t_2 - t_1) - \lambda_2 t_2\} & \\ \text{при } t_1 \leq t_2; \end{cases}$$

$$E\{\tau_2 | \tau_1 = t_1\} = \lambda_2^{-1} - \lambda \lambda_{12} \times \{\lambda_2(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12})\}^{-2} \cdot \exp\{-\lambda_2 t_1\}.$$

Сингулярностью распределения E_m во многом предопределяется сложность построения статистических выводов относительно его параметров. Общий подход к точечному оцениванию функционалов от E_m связан с получением оценок максимума правдоподобия для Λ с последующей подстановкой их в соответствующие функционалы.

Рассмотрим многомерное показательное распределение E_m , у которого $\lambda_j = 0$ при $J \neq i, i = 1, \dots, m$ и $J \neq (1, 2, \dots, m)$.

Пусть $(\bar{T}_1, \dots, \bar{T}_n)$ — случайная выборка объема n из распределения E_m , $\bar{T}_j = (T_{j1}, \dots, T_{jm}), T_{jr}$ — время жизни элемента l_r в j -й реализации. Обозначим: n_i — число реализаций, в которых элемент l_i отказал не последним; n_0 — число реализаций, когда отказали по крайней мере два элемента; $n_0(i)$ — число реализаций, когда l_i отказал в группе; $n_i(c)$ — число реализаций, когда l_i отказал последним, но не в группе с другими элементами.

Пусть

$$\theta_i = \int_0^\infty \gamma_i \exp\{-\gamma_i - u\} \times \prod_{j \neq 0, i} F_j(u) du, \quad i = 1, \dots, m,$$

где $F_j(u) = 1 - \exp\{-\lambda_j u\}$, $E\tau_i = \gamma_i^{-1}$, $E\tau_{(1)} = \lambda^{-1}$; $\lambda_0 E\tau_{(m)} = 1 - \sum_{i=1}^m (\lambda_i/\gamma_i) \theta_i$; $T_{j(m)}$ — значение $\tau_{(m)}$ в j -й реализации.

Если $n_i > 0, i = 0, 1, \dots, m$, то $\hat{\Lambda}$ — оценки максимума правдоподобия для Λ существуют и определяются как единственный корень системы

$$(n_i/\lambda_i) + (n_i(c)/\gamma_i) = \sum_j T_{ji}, \quad i = 1, \dots, m; (n_0/\lambda_0) + \sum_{i=1}^m (n_i(c)/\gamma_i) = \sum_j T_{j(m)}. \quad (19)$$

Если $n_r = 0$ для некоторого $r = 0, 1, \dots, m$, то оценка максимума правдоподобия выпишется в явном виде:

$$\hat{\lambda}_i = n [n_i / (n - n_i(c))] / \sum_j T_{ji}, \quad (20)$$

если $n_i(c) < n; \hat{\lambda}_i = n / \sum_j T_{ji}$,

если $n_i(c) = n, i = 1, \dots, k;$

$$\hat{\lambda}_0 = \left\{ n - \sum_{i=1}^m n_i(c) [n_i / (n - n_i(c))] \right\} \times \left\{ \sum_j T_{(j)m} \right\}^{-1},$$

если $n_i(c) < n$ для всех $i = 1, \dots, m$; $\hat{\lambda}_0 = 0$, если $n_i(c) = n$ для некоторого $i = 1, \dots, m$.

Оценки (20) единственны, за исключением случая, когда $n_r(c) = n$. Однако это может осуществиться лишь для одного r . Поэтому $\hat{\lambda}_i$ единственны, за исключением $\hat{\lambda}_r$ и $\hat{\lambda}_0$.

Для решения системы (19) можно рекомендовать итеративную процедуру вида

$$\lambda_i^{(r+1)} = (n_i + \xi_i^{(r)} n_i(c)) / \sum_j T_{ji},$$

$$i = 1, \dots, m;$$

$$\lambda_0^{(r+1)} = \left(n - \sum_{i=1}^m \xi_i^{(r)} n_i(c) \right) / \sum T_{j(m)},$$

где

$$\xi_i^{(0)} = n_i / (n_i + n_0(i)) =$$

$$= n_i / (n - n_i(c)), \quad i = 1, \dots, m;$$

$$\xi_i^{(r)} = \lambda_i^{(r)} / (\lambda_i^{(r)} + \lambda_0^{(r)}),$$

$$i = 1, \dots, m, \quad r = 1, 2,$$

Пусть $x = (\bar{T}_1, \dots, \bar{T}_n)$ — случайная выборка объема n из распределения E_m , ε_j — число различных значений среди координат $\bar{T}_j = (T_{j1}, \dots, T_{jm})$. Обозначим:

$$A_1 = \{1, 2, \dots, m\}; \quad \psi_{j1} = T_{j(1)} =$$

$$= \min_{1 \leq r \leq m} T_{jr},$$

$$A_{j2} = \{r : T_{jr} > \psi_{j1}\};$$

$$\psi_{j2} = \min_{r \in A_{j2}} T_{jr} - \psi_{j1},$$

$$A_{j3} = \{r : T_{jr} > \psi_{j1} + \psi_{j2}\};$$

$$\psi_{j3} = \min_{r \in A_{j3}} T_{jr} - \psi_{j1} - \psi_{j2},$$

$$A_{j\varepsilon_j} = \{r : T_{jr} = T_{j(m)} = \max_{1 \leq r \leq m} T_{jr}\};$$

$$\psi_{j\varepsilon_j} = T_{j(m)} - \psi_{j1} - \dots - \psi_{j(\varepsilon_j-1)}.$$

Обозначим $\{M_1, \dots, M_R\}$, перенумерованные в некотором порядке без повторения множества A_{ji} , $i = 1, \dots, \varepsilon_j$; $j = 1, \dots, n$, положив $A_1 = M_1$.

Здесь $R \leq N, N = \sum_{j=1}^n \varepsilon_j$. Определим

$$\text{статистики } z_r = \sum_{A_{j\omega} \in M_r} \psi_{j\omega}, \quad \kappa_r =$$

$$= \{\text{числу множеств } A_{j\omega} \in M_r, \omega =$$

$$= 1, 2, \dots, \varepsilon_j; j = 1, \dots, n\}. \text{ Условное рас-}$$

$$\text{пределение величин } \left\{ z_r \sum_{J \in Q} \lambda_J, Q = (J$$

$J \cap M_r \neq \emptyset), r = 1, \dots, R \}$ при фиксированных значениях $(\kappa_r, r = 1, \dots, R)$ совпадает с распределением R независимых гамма-случайных величин с функциями распределения $\Gamma(u, \kappa_r, 1)$. Этот факт дает возможность построить доверительные границы для функций вида $W(\Lambda) = \sum \lambda_j b_j$, где b_j — известные постоянные, а также и для монотонных функций от $W(\Lambda)$, к классу которых принадлежит и (12).

Пусть $\Gamma(n, 1)$ — случайная величина с функцией распределения $\Gamma(u; n, 1)$, $\xi(\gamma; n)$ и $\bar{\xi}(\gamma; n)$ — величины, для которых

$$P \{ \xi(\gamma; n) < \Gamma(n, 1) < \bar{\xi}(\gamma; n) \} = \gamma.$$

Определим множества

$$H_\gamma^1(x) = \left\{ \Lambda : \xi(\gamma^{\eta_1}; n) <$$

$$< \left(\sum_{J \in G} \lambda_J \right) \sum_{j=1}^n T_{j(1)} < \bar{\xi}(\gamma^{\eta_1}; n);$$

$$\bigcap_{q=2}^R (\xi(\gamma^{\eta_q}; \kappa_q) < z_r \times$$

$$\times \sum_{J: J \cap M_r \neq \emptyset} \lambda_J < \bar{\xi}(\gamma^{\eta_q}; \kappa_q) \right\};$$

$$H_\gamma^2(x) = \left\{ \Lambda : \xi(\gamma; N) <$$

$$< \sum_{J \in G} \lambda_J \sum_{i=1}^n \max_{t \in J} T_{jt} < \bar{\xi}(\gamma; N) \right\}.$$

Здесь $\eta_r = \kappa_r / N$, $\eta_1 \equiv n$, $\sum_{r=1}^R \eta_r = 1$.

Система множеств $H_\gamma^k(x)$ является γ -доверительной системой для параметра Λ , $k = 1, 2$.

Доверительные границы для $W(\Lambda)$ получим как \min и \max .

$$\Lambda \in H_j^k(x) \quad \Lambda \in H_j^k(x)$$

Аналогично рекомендуется получать и доверительные границы для более сложных функций от Λ . Вопрос о том, какую из двух γ -доверительных систем выбирать, зависит от оцениваемой функции.

3. АНАЛИЗ СИСТЕМ, ОПИСЫВАЕМЫХ МАРКОВСКИМ ПРОЦЕССОМ С ДИСКРЕТНЫМ ВРЕМЕНЕМ

Приведем результаты об асимптотических свойствах оценок максимального правдоподобия параметров марковской цепи при длительном наблюдении эволюции этой цепи.

Пусть $\{z_n, n = 0, 1, 2, \dots\}$ — марковский процесс с дискретным временем, Z — множество его состояний. На пространстве (Z, F_Z) задано семейство однородных по времени переходных функций $\{P_{\bar{\theta}}(\cdot, A), \bar{\theta} \in \Theta, A \subset F_Z\}$, имеющих плотность относительно одной и той же меры $\lambda(\cdot)$. Параметрическое пространство $\Theta \in R^r$ является открытым множеством. Предполагается, что каждому значению векторного параметра $\bar{\theta} = (\theta_1, \dots, \theta_r)^T$ отвечает единственное невырожденное стационарное распределение $P_{\bar{\theta}}(\cdot)$, а разным значениям $\bar{\theta}$ отвечают различные стационарные распределения. Далее пусть $f(y_1, y_2; \bar{\theta}) = d_{y_2} P_{\bar{\theta}}(y_1, y_2) / d\lambda(y_2)$, $f(y_0; \bar{\theta})$ — плотность начального распределения процесса, $y_i \in Z, i = 0, 1, 2$. Обозначим τ момент прекращения наблюдений, а $\varphi(t; \bar{\theta})$ — его плотность вероятности, $t = 0, 1, \dots$. При $\tau = t$ логарифм функции правдоподобия (ЛФП) принимает вид

$$L(\bar{\theta}) = \ln f(z_0; \bar{\theta}) +$$

$$+ \sum_{j=0}^{t-1} \ln f(z_j, z_{j+1}; \bar{\theta}) + \ln \varphi(t; \bar{\theta}).$$

(21)

Сформулируем необходимые условия регулярности, при которых $\hat{\bar{\theta}}$ — оценки максимума правдоподобия для $\bar{\theta}$ — будут асимптотически нормальными.

1. Для любого $y' \in Z$ множество точек y'' , для которых $f(y'; y''; \bar{\theta}) > 0$, не зависит от $\bar{\theta}$.

Пусть $f_u = \frac{\partial f}{\partial \theta_u}$; $f_{uv} = \frac{\partial^2 f}{\partial \theta_u \partial \theta_v}$ и т. д.

$$g(\cdot, \bar{\theta}) = \ln f(\cdot, \bar{\theta}).$$

2. Для любых $y', y'' \in Z$ функции $f_u(y', y''; \bar{\theta})$, $f_{uv}(y', y''; \bar{\theta})$ и $f_{uvw}(y', y''; \bar{\theta})$ существуют и непрерывны всюду на Θ .

Таким образом, для любого y' функции $g(y', y''; \bar{\theta})$, $g_u(y', y''; \bar{\theta})$, $g_{uv}(y', y''; \bar{\theta})$ и $g_{uvw}(y', y''; \bar{\theta})$ существуют и непрерывны на Θ , за исключением множеств $P_{\bar{\theta}}(y' \cdot)$ — меры нуль, $u, v, w = 1, \dots, r$.

3. Для любого $\bar{\theta} \in \Theta$ существует такая окрестность $N_{\bar{\theta}}$, что для любых u, v, w и y' имеем

$$\int_Z \sup_{\bar{\theta}' \in N_{\bar{\theta}}} |f_u(y', y''; \bar{\theta}')| \lambda(dy'') < \infty;$$

$$\int_Z \sup_{\bar{\theta}' \in N_{\bar{\theta}}} |f_{uv}(y', y''; \bar{\theta}')| \lambda(dy'') < \infty;$$

$$E_{\bar{\theta}} \left\{ \sup_{\bar{\theta}' \in N_{\bar{\theta}}} |g_{uvw}(z_1, z_2; \bar{\theta}')| \right\} < \infty.$$

Ожидание $E_{\bar{\theta}} \{ \cdot \}$ вычисляется по вероятностной мере, порождаемой единственным стационарным распределением в качестве начального.

4. Для $u = 1, \dots, r$ и некоторого $\delta > 0$, которое может зависеть от $\bar{\theta}$,

$$E_{\bar{\theta}} \{ |g_u(z_1, z_2; \bar{\theta})|^{2+\delta} \} < \infty.$$

Определим

$$I_{uv}(\bar{\theta}) = E_{\bar{\theta}} \{ g_u(z_1, z_2; \bar{\theta}) g_v(z_1, z_2; \bar{\theta}) \}.$$

5. Матрица

$$I(\bar{\theta}) = (I_{uv}(\bar{\theta})), \quad u, v = 1, \dots, r,$$

не вырождена.

6. Для каждого $\bar{\theta} \in \Theta$ марковская цепь неразложима, а каждое состояние является возвратным и ненулевым.

Когда $t \rightarrow \infty$, вклад $\ln f(z_0; \bar{\theta})$ и $\ln \varphi(t, \bar{\theta})$ в $L(\bar{\theta})$ пренебрежимо мал, поэтому эти слагаемые можно опустить и далее не учитывать.

Пусть $\bar{\theta}^0$ — истинное значение параметра $\bar{\theta}$, $\hat{\bar{\theta}}_t = (\hat{\bar{\theta}}_{t1}, \dots, \hat{\bar{\theta}}_{tr})^T$ — ОМП, являющаяся решением системы

$$\frac{\partial}{\partial \theta_u} L(\bar{\theta}) = \sum_{j=0}^{t-1} g_u(z_j, z_{j+1}; \bar{\theta}) = 0, \\ u = 1, \dots, r. \quad (22)$$

Оценка $\hat{\bar{\theta}}_t$ — состоятельная асимптотическая эффективная оценка $\bar{\theta}^0$.

Если $\tilde{\bar{\theta}}_t$ — какое-либо другое решение системы (22), то $P_{\bar{\theta}^0} \{ \hat{\bar{\theta}}_t = \tilde{\bar{\theta}}_t \} \rightarrow 1$ при $t \rightarrow \infty$.

Пусть $\bar{q}(t) = (q_1(t), \dots, q_r(t))^T$, $q_i(t) = \sqrt{t}(\hat{\bar{\theta}}_{ti} - \theta_i^0)$, $i = 1, \dots, r$. Тогда при $t \rightarrow \infty$

$$\bar{q}(t) = N_r(\bar{0}, I^{-1}(\bar{\theta}^0)) + O_p(\bar{1}). \quad (23)$$

В задаче проверки простой гипотезы $H_0: \bar{\theta} = \bar{\theta}^0$ в качестве статистики критерия можно взять

$$Q_t(\hat{\bar{\theta}}_t, \bar{\theta}^0) = 2 \{ L(\hat{\bar{\theta}}_t) - L(\bar{\theta}^0) \}.$$

Критические уровни рассчитывают при этом с учетом того, что

$$Q_t(\hat{\bar{\theta}}_t, \bar{\theta}^0) \approx \bar{q}^T(t) I(\bar{\theta}^0) \bar{q}(t) = \\ = \chi_r^2 + O_p(1). \quad (24)$$

Здесь и далее $\bar{0}$ и $\bar{1}$ — векторы, элементы которых состоят соответственно из нулей и единиц; $N_r(\bar{\mu}, \Sigma)$ — r -мерная нормальная случайная величина; χ_r^2, δ^2 — хи-квадрат — случайная величина с r степенями свободы и параметром нецентральности δ^2 , $\chi_r^2 = \chi_{r,0}^2$.

Мощности предложенного критерия при близких гипотезах $\theta^{(t)} = \theta^0 + t^{-1/2}h^{(t)}$, где $h^{(t)} \rightarrow h$ при

$t \rightarrow \infty$; h — фиксированный вектор, вычисляется на основе того, что

$$2 \{ L(\hat{\bar{\theta}}_t) - L(\bar{\theta}^{(t)}) \} \approx \bar{q}^T I(\bar{\theta}^0) \bar{q}(t) = \\ = \chi_r^2, \delta^2 + O_p(1). \quad (25)$$

Здесь $\delta^2 = h^T I(\bar{\theta}^0) h$, $\bar{q}(t) = \sqrt{t}(\hat{\bar{\theta}}_t - \bar{\theta}^{(t)})$. Требование достижения заданной мощности определяет необходимую продолжительность испытаний.

Пусть необходимо проверить сложную гипотезу $H_0: \bar{\theta} \in \Theta_0$ против сложной альтернативы $H_1: \bar{\theta} \in \Theta_1$, $\Theta \cup \Theta_1 = \Theta$, $\Theta_0 \cap \Theta_1 = \emptyset$. Если Θ и Θ_0 — подпространства размерностей соответственно r и k , $r > k$, то

$$2 [\max_{\Theta_0} L(\bar{\theta}) - L(\bar{\theta}^0)] = \chi_k^2 + O_p(1); \quad (26)$$

$$2 [L(\hat{\bar{\theta}}_t) - \max_{\Theta_0} L(\bar{\theta})] = \\ = \chi_{r-k}^2 + O_p(1). \quad (27)$$

Таким образом, (23) и (27) можно использовать при построении критериев в задаче проверки сложных гипотез. Статистики $2 [\max_{\Theta_0} L(\bar{\theta}) - L(\bar{\theta}^0)]$ и $2 [L(\hat{\bar{\theta}}_t) - \max_{\Theta_0} L(\bar{\theta})]$

асимптотически независимы.

Замечания. 1. Элементы матрицы $I(\bar{\theta}^0)$ вычислить, как правило, не просто. Во многих задачах вместо $I(\bar{\theta}^0)$ допустимо использование ее состоятельной оценки $\hat{I}(\hat{\bar{\theta}}_t)$, элементами которой являются

$$\frac{1}{t} \sum_{j=0}^t g_u(z_j, z_{j+1}, \hat{\bar{\theta}}_t) g_v(z_j, z_{j+1}, \hat{\bar{\theta}}_t),$$

$u, v = 1, \dots, r$.

2. Уравнение (22) в замкнутом виде решается довольно редко. Если удастся найти другую состоятельную оценку $\tilde{\bar{\theta}}$, то оценка вида

$$\tilde{\bar{\theta}}_t = \tilde{\bar{\theta}} + \hat{I}^{-1}(\hat{\bar{\theta}}_t) \text{grad } L(\tilde{\bar{\theta}})$$

при $t \rightarrow \infty$ является асимптотически эффективной, т. е.

$$\sqrt{t} (\bar{\theta}_t - \bar{\theta}^0) = N_r(\bar{0}, I^{-1}(\bar{\theta}^0)) + o_p(\bar{1}).$$

Величину $I^{-1}(\bar{\theta}_t) \text{grad } L(\bar{\theta}_t)$ часто называют поправками Ле Кама.

4. АНАЛИЗ СИСТЕМ, ОПИСЫВАЕМЫХ МАРКОВСКИМ ПРОЦЕССОМ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Пусть $\{z(u), 0 \leq u < \infty\}$ — скачкообразный марковский сепарабельный процесс с пространством состояний Z ,

$$P_{\bar{\theta}}(u, a, A) = P_{\bar{\theta}}\{z(u+s) \in A \mid z(s) = a\},$$

$$\bar{\theta} \in \Theta \subset R^r;$$

$$q(a, \bar{\theta}) = \lim_{u \rightarrow 0} (1 - P_{\bar{\theta}}(u, a, \{a\}))/u,$$

$$a \in Z.$$

Пусть $0 < q(a, \bar{\theta}) < \infty$ для всех $\bar{\theta} \in \Theta$. Обозначим $v(u)$ — число скачков процессов $z(u)$ на интервале $(0, u)$. Траекторию процесса $z(u)$ на интервале $(0, t)$ можно задать посредством последовательности $\{(z_n, \rho_n), n = 0, \dots, v(t)\}$, где $z_0 = z(0)$; z_i — положение процесса после i -го скачка; ρ_i — время, проведенное процессом в состоянии z_i . Таким образом, $v(u) = \max\{k : \rho_1 + \dots + \rho_k < u\}$, $z(u) = z_{v(u)}$.

Из общей теории марковских процессов известно, что

$$P_{\bar{\theta}}\{\rho_{n+1} > u \mid \rho_1, \dots, \rho_n,$$

$$z_1, \dots, z_{n+1}\} = \exp\{-q(z_{n+1}; \bar{\theta})u\}$$

$$P_{\bar{\theta}}\{z_{n+1} \in A \mid \rho_1, \dots, \rho_n, z_1, \dots, z_n\} = q(z_n, A; \bar{\theta})/q(z_n; \bar{\theta}).$$

В приложениях процесс $z(u)$ задают, как правило, посредством переходных интенсивностей $q(a, A; \bar{\theta})$. Семейство мер $\{P_{\bar{\theta}}(\cdot), \bar{\theta} \in \Theta\}$ доминируется мерой $P_{\bar{\theta}^0}(\cdot)$, если существуют производные Радома —

$$\text{Никодима } \delta(a, \bar{\theta}) = dP_{\bar{\theta}}(z(0) = a)/dP_{\bar{\theta}^0}(z(0) = a) \text{ и}$$

$$\delta(a, y; \bar{\theta}) = dq(a, y; \bar{\theta})/dq(a, y; \bar{\theta}^0).$$

В этом случае ЛФП, построенная для траектории $\{z(u), 0 \leq u \leq t\}$ по мере $P_{\bar{\theta}^0}(\cdot)$, принимает вид

$$\begin{aligned} \dot{L}_t(\bar{\theta}) &= \ln \delta(z_1; \bar{\theta}) + \\ &+ \sum_{k=1}^{v(t)} \rho_k [q(z_k; \bar{\theta}^0) - q(z_k; \bar{\theta})] + \\ &+ \left(t - \sum_{k=1}^{v(t)} \rho_k\right) \times \\ &\times [q(z_{v(t)+1}; \bar{\theta}^0) - q(z_{v(t)+1}; \bar{\theta})] + \\ &+ \sum_{k=1}^{v(t)} \ln \delta(z_k; z_{k+1}; \bar{\theta}). \end{aligned}$$

При $t \rightarrow \infty$ первым слагаемым можно пренебречь.

Заметим, что $\{(z_n, \rho_n), n = 1, 2, \dots\}$ является вложенной марковской цепью на $Z \times R_+$, где $R_+ = \{y : y \geq 0\}$. Если $\pi_{\bar{\theta}}(a, A) = q(a, A; \bar{\theta})/q(a, \bar{\theta})$, то этот вложенный процесс имеет переходные вероятности

$$P_{\bar{\theta}}(z_{n+1} \in A, \rho_{n+1} \in B \mid z_n = a,$$

$$\rho_n = \alpha) = \int_A \pi_{\bar{\theta}}(a, dy) \int_B q(y; \bar{\theta}) \times$$

$$\times \exp\{-q(y; \bar{\theta})\beta\} d\beta,$$

где $A \subset F_Z$, $B \subset J_{R_+}$. Используя результаты теории статистического вывода для марковских процессов с дискретным временем, можно строить статистический анализ для $\{(z_n, \rho_n), n = 1, 2, \dots\}$. Он будет эквивалентен анализу, основанному на $\{z(u), 0 < u < \infty\}$.

Пример. Процессы гибели и рождения. Пусть $\lambda_i(\bar{\theta})$ — интенсивность перехода процесса из состояния E_i в состояние E_{i+1} , $\mu_i(\bar{\theta})$ — интенсивность перехода из E_i в E_{i-1} . Предположим, что выполнены следующие условия:

- 1) $\mu_0(\bar{\theta}) = 0$;
- 2) $\mu_j(\bar{\theta}) > 0$ для $j = 1, 2, \dots$;

3) $\lambda_j(\theta) > 0$ для целых $j < c$, где $M \leq \infty$;

4) $\lambda_j(\theta) = 0$ для $j = c, c + 1$, если c — конечно;

5) для каждого состояния среднее время возвращения конечно.

Процесс гибели и рождения $\{z(u), 0 \leq u \leq t\}$ является марковским процессом с кусочно-постоянными траекториями, а условия 1—5 гарантируют выполнение условий (22)—(26). Однако модель сама по себе интересна и заслуживает того, чтобы результаты были сформулированы специально для нее. Пусть u_i — число переходов $E_i \rightarrow E_{i+1}$, d_i — число переходов $E_i \rightarrow E_{i-1}$, γ_i — суммарное время пребывания $z(u)$ в E_i . Тогда с точностью до $o_p(1)$

$$L_t(\theta) = \sum_{i=0}^{\infty} u_i \ln \lambda_i(\theta) + \\ + \sum_{j=1}^{\infty} d_j \ln \mu_j(\theta) - \\ - \sum_{i=0}^{\infty} \gamma_i (\lambda_i(\theta) + \mu_i(\theta)).$$

Большой интерес представляют следующие две модели.

Модель 1.

$$\lambda_j(\bar{\theta}) = \lambda_j > 0$$

для $j = 1, \dots, m-1, \lambda_m = 0$;

$$\mu_j(\theta) = \mu_j > 0$$

для $j = 1, \dots, m, \mu_0 = 0$.

В этом случае $\bar{\theta}^T = (\lambda_0, \dots, \lambda_{m-1}, \mu_1, \dots, \mu_m)$. ОМП параметров λ_j и μ_j имеют вид

$$\hat{\lambda}_j = u_j / \gamma_j \quad \text{для } j = 1, \dots, m-1;$$

$$\hat{\mu}_j = d_j / \gamma_j \quad \text{для } j = 1, \dots, m.$$

Пусть

$$I(\theta) = \left(\frac{1}{2} R \right) (I_i \delta_{ij});$$

$$R = \sum_{j=0}^{m-1} \lambda_j P_j = \sum_{j=1}^m \mu_j P_j;$$

$$P_j = \lim_{t \rightarrow \infty} P\{z(t) = j | z(0) = i\};$$

вектор

$$\bar{l} = (l_1, \dots, l_{2m})^T = \\ = (P_0/\lambda_0, \dots, P_{m-1}/\lambda_{m-1}, \\ P_1/\mu_1, \dots, P_{m-1}/\mu_{m-1})^T$$

Свойство (23) в применении к рассматриваемой ситуации означает, что

$$\sqrt{v(t)} (\hat{\theta} - \bar{\theta}^0) = \\ = N_{2m}(\bar{0}, I^{-1}(\bar{\theta}^0)) + o_p(\bar{1}), \\ v(t) = \sum (u_j + d_j).$$

Поскольку $I^{-1}(\bar{\theta}^0)$ — диагональная матрица, то $\hat{\lambda}_j$ и $\hat{\mu}_k$ асимптотически независимы.

Асимптотические дисперсии

$$D_{ac} \{ \sqrt{v(t)} (\hat{\lambda}_j - \lambda_j) \} = 2\lambda_j R / P_j,$$

$$j = 0, \dots, m-1;$$

$$D_{ac} \{ \sqrt{v(t)} (\hat{\mu}_j - \mu_j) \} = 2\mu_j R / P_j,$$

$$j = 1, \dots, m.$$

При проверке гипотезы $H: \bar{\theta} = \bar{\theta}^0$ статистика критерия

$$2 [L_t(\hat{\theta}) - L_t(\bar{\theta}^0)] = \\ = \sum_{j=0}^{m-1} u_j \ln (u_j / \lambda_j^0 \gamma_j) + \\ + \sum_{j=1}^m d_j \ln (d_j / \mu_j^0 \gamma_j) + \\ + \sum_{j=0}^m \gamma_j (\lambda_j^0 + \mu_j^0) - n = \chi_{2m}^2 + o_p(1).$$

Для близкой гипотезы $H_1: \bar{\theta} = \bar{\theta}^1 = (\lambda_0^1, \lambda_1^1, \dots, \lambda_m^1, \mu_1^1, \dots, \mu_{m+1}^1)^T$ мощность критерия вычисляется на основе того, что

$$2 [L_t(\hat{\theta}) - L_t(\bar{\theta}^0)] = \chi_{2m, 0}^2 + o_p(1).$$

где параметр нецентральности

$$\delta^2 = \frac{\nu(t)}{2R} \left[\sum_{i=0}^{m-1} \frac{p_i (\lambda_i^0 - \lambda_i^1)^2}{\lambda_i^1} + \sum_{i=1}^m \frac{p_i (\mu_i^0 - \mu_i^1)^2}{\mu_i^1} \right]$$

Модель 2.

$$\begin{aligned} \lambda_j(\bar{\theta}) &= \lambda h(j), \quad j = 0, 1, \\ \mu_j(\bar{\theta}) &= \mu g(j), \quad j = 1, 2, \\ g(0) &= 0. \end{aligned}$$

Здесь $h(j)$ и $g(j)$ — произвольные известные функции, $\bar{\theta} = (\lambda, \mu)^T$. Как частные случаи эта модель описывает работу системы с m приборами и r ремонтными устройствами, в которых моменты отказов образуют пуассоновские потоки, а времена восстановления — показательные случайные величины. ОМП параметров λ и μ имеет вид

$$\hat{\lambda} = \sum_{j=0}^{\infty} u_j \Big/ \sum_{j=0}^{\infty} \gamma_j h(j);$$

$$\hat{\mu} = \sum_{j=1}^{\infty} d_j \Big/ \sum_{j=1}^{\infty} \gamma_j g(j).$$

$$\text{Поскольку } P \left\{ \frac{\lim_{t \rightarrow \infty} \sum_{j=0}^{\infty} u_j}{\sum_{j=1}^{\infty} d_j = 1} \right\} = 1,$$

то при больших t можно положить

$$\sum_{j=0}^{\infty} u_j \approx \sum_{j=0}^{\infty} d_j \approx \frac{1}{2} \nu(t).$$

Предельное распределение $\hat{\lambda}$ и $\hat{\mu}$ при этом не изменится. Далее

$$I^{-1}(\bar{\theta}) = \left\| \begin{array}{cc} 2\lambda^2 & 0 \\ 0 & 2\mu^2 \end{array} \right\|.$$

Это не означает, что асимптотические дисперсии $\hat{\lambda}$ и $\hat{\mu}$ не зависят от $h(j)$

и $g(j)$. От этих функций зависит дисперсия величины $\nu(t)$. В рассматриваемом случае параметр нецентральности

$$\begin{aligned} \delta^2 &= \frac{1}{2} \nu(t) \{[(\lambda^0 - \lambda')/\lambda']^2 + \\ &+ [(\mu^0 - \mu')/\mu']^2\}. \end{aligned}$$

5. АНАЛИЗ СИСТЕМ, ОПИСЫВАЕМЫХ ПОЛУМАРКОВСКИМ ПРОЦЕССОМ

Пусть $\{z(u), 0 \leq u < t\}$ — полумарковский процесс с $m < \infty$ состояниями; τ_1, τ_2, \dots — моменты скачков процесса $z(u)$; $Y_i = \tau_{i+1} - \tau_i$ — время между $(i+1)$ и i скачком; $z_i = z(\tau_i + 0), i > 0, z_0 = z(0)$; $P_{ij} = P\{z_{q+1} = j | z_q = i\}$ — переходные вероятности вложений марковской цепи, $P\{z_q = i, Y_q = y | z_0 = j_0, \dots, z_1 = j_1, \dots, z_{q-1} = j, Y_1 = y_1, \dots, Y_{q-1} = y_{q-1}\} = Q_{ij}(y)$. Если $P_{ij} > 0$, положим $F_{ij}(y) = P_{ij}^{-1} Q_{ij}(y)$. Как и прежде, $\nu(t)$ — число скачков процесса $z(u)$ на $[0, t)$; $\nu_i(t)$ — число попаданий в состояние E_i ; $\nu_{ij}(t)$ — число переходов из E_i в E_j на $[0, t)$. Обозначим X_{ij} — время пребывания $z(u)$ в E_i при j -м попадании.

Рассмотрим процесс $\{(z_n, Y_n); n \geq 0\}$. Легко показать, что это марковский процесс с дискретным временем. Более того, по траектории этого процесса на $(0, \nu(t)]$ восстанавливается траектория процесса $z(u)$ на $[0, t)$. Если величины P_{ij} и $Q_{ij}(y)$ зависят от параметра $\bar{\theta}$, а функциональная форма $P_{ij} = P_{ij}(\bar{\theta})$ и $Q_{ij}(y) = Q_{ij}(y, \bar{\theta})$ известна, но неизвестно значение параметра, то статистические выводы об этом параметре можно сделать методами на основе статистического анализа марковского процесса с дискретным временем $\{(z_n, Y_n); 0 \leq n \leq \nu(t)\}$. При $t \rightarrow \infty$ асимптотические свойства соответствующих оценок и статистик критерия аналогичны свойствам оценок критериев, основанных на методе максимума прав-

доподобия для $\{z(u), 0 \leq u < t\}$.
В частности, ЛФП принимает вид

$$L_t(\theta) = \ln P\{z_0, \tau_1; \theta\} + \\ + \sum_{i,j} v_{ij}(t) \ln P_{ij}(\theta) + \\ + \sum_{j=1}^m \sum_{q=0}^{v_j(t)} \ln Q_{jq}(X_{jq}; \theta).$$

Если рассматривается непараметрическая ситуация, то переходить к эквивалентному марковскому процессу неудобно. Предположим далее, что $F_{ij}(u) = H_i(u)$ для $1 \leq j \leq m$. Общность ситуации при этом не теряется. Положим

$$\hat{P}_{ij}(t) = v_{ij}(t)/v_i(t); \\ \hat{H}_i(u; t) = v_i^{-1}(t) \sum_{k=1}^{v_i(t)} \delta(u - X_{ik}); \\ \hat{Q}_{ij}(u; t) = \hat{P}_{ij}(t) \hat{H}_i(u; t),$$

где $\delta(u)$ — функция Хевисайда: $\delta(u) = 1$ для $u > 0$ и $\delta(u) = 0$ для $u \leq 0$. Если $Y_i(t) = 0$, положим $\hat{P}_{ij}(t) = 0$ и $\hat{H}_i(u; t) = 0$.

Далее будут предполагаться неразложимость и возвратность процесса $z(u)$. В этом случае

$$P\{\max_{i,j} \sup_u |\hat{Q}_{ij}(u; t) - Q_{ij}(u)| \rightarrow 0\} = 1 \\ \text{при } t \rightarrow \infty,$$

т. е. оценка $\hat{Q}_{ij}(u; t)$ строго состоятельная. Это верно и для $\hat{H}_i(u; t)$.

Для фиксированных i, j, u вектор

$$(\sqrt{t} [\hat{P}_{ij}(t) - P_{ij}], \\ \sqrt{t} [\hat{H}_i(u; t) - H_i(u)])^T = \\ = N_2(0, \Sigma) + O_p(1).$$

Здесь

$$\Sigma = (\sigma_{ij}), \quad \sigma_{11} = \mu_{ii} P_{ij} (1 - P_{ij}); \\ \sigma_{22} = \mu_{ii} H_i(u) [1 - H_i(u)], \quad \sigma_{12} = 0, \\ \sigma_{21} = 0, \quad \mu_{ij} = E\tau_{ij}; \\ \tau_{ij} = \inf\{n : z(u) = j, z(0) = i\}.$$

Таким образом, $\hat{P}_{ij}(t)$ и $\hat{H}_i(u, t)$ асимптотически независимы.

Пусть $\{W_{ijk}(t) = \hat{Q}_{ij}(u_k; t) - Q_{ij}(u_k), 1 \leq j, i \leq m \text{ и } 1 \leq k \leq s\}$. Для фиксированного s вектор \bar{W} с компонентами $W_{ijk}(t)$ сходится по распределению при $t \rightarrow \infty$ к $m^2 s$ -мерной нормальной случайной величине с нулевым средним и ковариационной матрицей $(\sigma_{ijk; uvw})$, где

$$\sigma_{ijk; uvw} = \mu_{ii} \sigma_{ik} P_{ij} [H_i(u_k \wedge u_w) P_{ij} + \\ + H_i(u_k) H_i(u_w) (\delta_{jv} - 2P_{iv})].$$

Здесь δ_{ij} — символ Кронекера, $a \wedge b = \min(a, b)$

6. АНАЛИЗ СИСТЕМ, ОПИСЫВАЕМЫХ ПРОЦЕССОМ ВОССТАНОВЛЕНИЯ

Реализация процесса восстановления определяется последовательностью $z = (x_1, x_2, \dots, x_\tau)$ величин интервалов между событиями процесса. Если $\tau \equiv n$, где n — фиксированное число, то z — независимая выборка.

Рассмотрим случай, когда процесс восстановления наблюдается на интервале $[0, T]$, т. е. $\tau = k_T$ — число восстановлений на $[0, T]$. Распределение процесса восстановления полностью определяется ф. р. $F(u) = P\{x_j < u\}$, где x_j — интервал между моментами осуществления j -го и $(j+1)$ -го событий процесса. С $F(u)$ тесно связана другая характеристика процесса восстановления — функция восстановления $H(u) = EK(u)$, где $K(u) = \max\{r : x + x_r \leq u\}$. Известно, что

$$H(u) = \sum_{i=1}^{\infty} F^{(i)}(u) \text{ и} \\ F(u) = \sum_{i=1}^{\infty} (-1)^{i+1} H^{(i)}(u),$$

где для функции $\Psi(u)$ под $\Psi^{(i)}(u)$ понимается i -кратная свертка $\Psi(u)$. Отсюда вытекает, что для процессов восстановления существуют две эквивалентные характеристики $F(u)$ и $H(u)$.

Далее будем предполагать, что результат наблюдений z состоит из N

независимых реализаций процесса восстановления на $[0, T]$, т. е. рассмотрен план $[N, B, T]$. Пусть $z = (x_{ij}, j = 1, \dots, k_{iT}, i = 1, \dots, N)$, где k_{iT} — число восстановлений в i -й реализации, x_{ij} — интервал между моментами j -го и $(j+1)$ -го восстановлений в i -й реализации.

Непараметрический случай. Пусть $\{F(u)\}$ — совокупность всех непрерывных функций на $[0, T]$, $0 < F(T) \leq 1$. Наилучшей несмещенной оценкой для $F(u)$ будет

$$F_{NT}(u) = \frac{1}{N} \sum_{i=1}^N \frac{v_{iu}(T)}{k_{iT}}, \quad (28)$$

где $v_{iu}(T)$ — число x_{ij} , меньших u , $j = 1, \dots, k_{iT}, i = 1, \dots, N$.
Далее

$$\hat{U}_{NT}(u_1, u_2) = \begin{cases} \frac{1}{N_2} \left[\sum_{i=1}^N \frac{v_{iu_1}(T) [v_{iu_2}(T-1)]}{k_{iT} [k_{iT} - 1]} + \sum_{i \neq j} \frac{v_{iu_1}(T) v_{iu_2}(T)}{k_{iT} k_{jT}} \right] & \text{при } u_1 + u_2 \leq T, \\ \frac{1}{N(N-1)} \sum_{i \neq j} \frac{v_{iu_1}(T) v_{iu_2}(T)}{k_{iT} k_{jT}} & \text{при } u_1 + u_2 > T. \end{cases} \quad (30)$$

$$\text{cov}[F_{NT}(u_1), F_{NT}(u_2)] =$$

$$= \begin{cases} \int_0^{u_1} \int_0^T \delta(T - s_1 - s_2) A(T - s_1 - s_2) dF(s_1) dF(s_2) & \text{при } u_1 + u_2 \leq T, \\ \int_0^{u_1} \int_0^T \delta(T - s_1 - s_2) A(T - s_1 - s_2) dF(s_1) dF(s_2) + F(u_2)F(T - u_2) + \\ + \int_{T-u_2}^{u_1} F(T - s) dF(s) - F(u_1)F(u_2) & \text{при } u_1 + u_2 > T, \end{cases} \quad (29)$$

где

$$A(u) = \sum_{l=0}^{\infty} \frac{\bar{F}(u) * F^{(l)}(u)}{l+2},$$

$\Psi_1 * \Psi_2$

— свертка Ψ_1 и Ψ_2 . Формула (29) для вычисления сложна. Часто можно ограничиться несмещенной оценкой для $U(u_1, u_2; T) = \text{cov}[F_{NT}(u_1), F_{NT}(u_2)]$, а именно

Наилучшей несмещенной оценкой для $H(u)$ является

$$H_{NT}(u) = \frac{1}{N} \sum_{i=1}^N \times \sum_{q=1}^{k_{iT}} q \frac{[v_{iu}(T, q) - v_{iu}(T, q+1)]}{k_{iT}^{[q]}}, \quad (31)$$

где $v_{iu}(T, q)$ — число различных под-последовательностей $(x_{ij_1}, \dots, x_{ij_q})$, образованных из $(x_{i1}, \dots, x_{ik_{iT}})$ так, что

$$\sum_{r=1}^q x_{ij_r} \leq u, \quad b^{[q]} = b(b-1)\dots(b-q+1).$$

При $T \rightarrow \infty$ вычислить оценку (31) становится сложно. Вместо нее можно рекомендовать оценку вида

$$\widehat{H}_{NT}(u) = \frac{1}{N} \sum_{i=1}^N \frac{1}{1+k_{i,T-u}} \times \sum_{j=0}^{k_{T-u}} [k_{i,u+s_{ij}} - j], \quad (32)$$

где

$$s_{ij} = \sum_{r=1}^j x_{ir}.$$

Оценки (28), (30) и (31) являются состоятельными и асимптотически нормальными при $NT \rightarrow \infty$ для всех $0 \leq u \leq T$. Далее

$$P \{ \lim_{N \rightarrow \infty} \sup_{0 \leq u \leq T} |F_{NT}(u) - F(u)| = 0 \} = 1$$

и

$$P \{ \lim_{N \rightarrow \infty} \sup_{0 \leq u \leq T} |H_{NT}(u) - H(u)| = 0 \} = 1.$$

Кроме того, при $T \rightarrow \infty$ при $u_1 \leq u_2$

$$U(u_1, u_2; T) = F(u_1) [1 - F(u_2)] \left\{ \frac{\alpha_1}{T} + \frac{7\alpha_1\alpha_2 + 5\alpha_2 - 8\alpha_1^2}{4T^2} + o\left(\frac{1}{T^2}\right) \right\}. \quad (33)$$

Из (33) вытекает, что при $T \rightarrow \infty$ $F_{NT}(u)$ ведет себя аналогично эмпирической функции распределения (э. ф. р.) в случае независимой выборки и для нее справедливы все утверждения о предельном распределении статистики от э. ф. р.

Параметрический случай. Пусть

$$F(u) = F(u, \bar{\theta}); \quad f(u, \bar{\theta}) = \frac{\partial}{\partial u} F(u, \bar{\theta});$$

$$H(u) = H(u, \bar{\theta}) \text{ и } \bar{\theta} = (\theta_1, \dots, \theta_m)^T \in \Theta \subset R^m$$

При использовании метода максимального правдоподобия полезно иметь в виду, что ЛФП

$$L(\bar{\theta}) = \sum_{i=1}^N \sum_{j=1}^{k_{iT}} f(x_{ij}, \bar{\theta}) + \sum_{i=1}^N \ln \left[1 - F\left(T - \sum_{j=1}^{k_{iT}} x_{ij}; \bar{\theta}\right) \right].$$

Элементы информационной матрицы Фишера $I(\bar{\theta}) = (\tilde{I}_{ij}(\bar{\theta}))$ при $N = 1$ в рассматриваемом случае принимают вид

$$\tilde{I}_{ij}(\bar{\theta}) = [1 + H(T, \bar{\theta})] I_{ij}(\bar{\theta}) - \int_0^T \bar{F}(T-s, \bar{\theta}) \times$$

$$\times \frac{\partial^2 \ln \bar{F}(T-s, \bar{\theta})}{\partial \theta_i \partial \theta_j} dH(s, \bar{\theta}) +$$

$$+ \int_0^T dH(s, \bar{\theta}) \times$$

$$\times f \int_0^\infty \frac{\partial^2 \ln f(T-s+u, \bar{\theta})}{\partial \theta_i \partial \theta_j} \times$$

$$\times f(T-s+u, \bar{\theta}) du +$$

$$+ \int_0^\infty \frac{\partial^2 \ln f(s, \bar{\theta})}{\partial \theta_i \partial \theta_j} f(s, \bar{\theta}) ds,$$

где

$$I_{ij}(\bar{\theta}) = \int_0^\infty \frac{\partial^2 \ln f(u, \bar{\theta})}{\partial \theta_i \partial \theta_j} f(u, \bar{\theta}) du.$$

Статистические методы находят широкое применение при контроле качества промышленной продукции. Используя эти методы, можно получить объективные данные о реально достигнутом уровне показателей качества, о тенденциях изменений уровней показателей качества (например, в связи со старением оборудования, обеспечивающего необходимый технологический процесс). Последующая обработка результатов статистического контроля является основой для принятия решений о необходимых изменениях параметров технологического процесса, его подналадке и т. п. Статистический контроль является эффективным средством поддержания или улучшения уровней показателей качества промышленной продукции.

Так как принимаемые по результатам контроля решения связаны с реальными затратами и затрагивают интересы различных производственных коллективов, то статистический контроль проводится на основе утвержденных нормативных документов (ГОСТов, ОСТов и др.).

1. ПЛАНЫ КОНТРОЛЯ

Статистический контроль партий (штучной) промышленной продукции проводят на основе системы планов контроля. План контроля определяется правилами: (1) отбора изделий на контроль, (2) классификации изделий по результатам их контроля, (3) прекращения контроля, (4) принятия необходимого решения. В зависимости от конкретных результатов контроля в систему планов контроля включают **правила корректировки**, например, правила перехода от нормального контроля к усиленному контролю. В си-

стему планов контроля могут быть включены правила перехода на сплошной контроль, а в особых случаях и остановка производства.

Если при контроле изделия классифицируют на годные и дефектные, то такой контроль называют **контролем по альтернативному признаку**.

Пример 1. При контроле измеряют l_i — длину изделия O_i , которая должна находиться в пределах от a до b . Если $a \leq l_i \leq b$, то изделие O_i классифицируют как годное; если $b < l_i$ либо $l_i < a$, то изделие O_i классифицируют как дефектное.

Если по результатам контроля изделие относят к одной из r выбранных групп, то такой контроль называют **контролем по качественному признаку**. Контроль по альтернативному признаку соответствует $r = 2$.

Пример 2. В результате контроля становится известным диаметр d_i изделия O_i . Если $h \leq d_i \leq H$, то изделие относят к классу K_1 годных изделий. Если $d_i < h$, то изделие относят к классу K_2 изделий, допускающих доработку, но непригодных к немедленной поставке потребителю. Если $d_i > H$, то изделие относят к классу K_3 дефектных изделий.

Если контролируемое изделие характеризуется параметром, принимающим вещественные или векторные значения, то такой контроль называют **контролем по количественному (векторному) признаку**.

Пример 3. Наполнение банки O_i жидкостью характеризуется объемом v_i налитой жидкости.

Замечание. При контроле по альтернативному признаку учитывают лишь числа изделий, классифицированных как годные и дефектные. При контроле по качественному признаку учитывают

количество изделий, включенных в каждый из классов K_1, \dots, K_r . При контроле по количественному признаку учитывают все результаты измерений параметров контролируемых изделий.

Если качество контролируемых изделий характеризуется совокупностью результатов измерений, составляющих функцию от некоторого аргумента (возможно, векторного), то такой контроль называют **контролем по функциональному признаку**.

Пример 4. При контроле участка микропровода длиной L результаты измерений толщины $d(l)$ микропровода на расстоянии l от его начала определяют значение функции $d(l)$, $0 \leq l \leq L$.

Если при контроле параметра, определяющего качество изделия, показатели качества изделия не меняются, то такой контроль называют **неразрушающим**. Если в результате контроля изделие оказывается полностью негодным для применения, то такой контроль называют **разрушающим**.

Пример 5. Емкость испытывают на предельное давление, вызывающее разрыв этой емкости. В результате измерений предельного давления емкость разрушается.

Если в контролируемой партии все изделия подвергаются контролю, то такой контроль называют **сплошным**. Если контролируется лишь часть изделий, то такой контроль называют **выборочным**.

Совокупность изделий, отобранных из партий для контроля, называют **выборкой**. Число изделий, составляющих выборку, называют **объемом выборки**. Число изделий в партии называют **объемом партии**.

Пример 6. Пусть $\mathfrak{F} = \{O_1, \dots, O_N\}$ — партия, предъявленная на контроль, а $\gamma_1, \gamma_2, \dots, \gamma_N$ — последовательность Бернулли (см. гл. 5), γ_i равны нулю или единице, $P\{\gamma_i = 1\} = v$, $i = 1, \dots, N$. В выборку включают изделия O_i , для которых $\gamma_i = 1$. Образованную таким образом выборку называют **биномиальной**. Объем такой выборки случаен и имеет биномиальное распределение.

Случайной выборкой объема n называют такой план контроля, при

котором все наборы из n различных изделий, отбираемых для контроля, имеют равные вероятности.

Случайным выбором без возвращения называют такой отбор изделий из партии на контроль, при котором изделия отбирают последовательно, отобранные и проконтролированные изделия в дальнейшем отборе не участвуют. Вероятность отбора для всех еще неотобранных изделий одна и та же.

По результатам выборочного контроля могут быть приняты различные решения. Типичными являются следующие решения:

D_1 : отвергнуть оставшуюся непроверенной часть партии без дальнейшего контроля,

D_2 : провести сплошную проверку всех изделий, не попавших в выборку,

D_3 : принять все изделия, не попавшие в выборку, без контроля.

Одноступенчатым планом контроля по альтернативному признаку с параметрами n, c называют план, при котором из партии берут случайную выборку фиксированного объема n . Все изделия из выборки проверяют. Если d — число обнаруженных в выборке дефектных изделий — не более **приемочного числа** c , то принимают решение D_j ; если же $d > c$, то принимают решение D_i , $i < j$, $i, j \in \{1, 2, 3\}$. Одноступенчатый план с решениями D_i, D_j обозначают $(n, c)_{i,j}$. Возможны следующие разновидности одноступенчатых планов: $(n, c)_{12}, (n, c)_{13}, (n, c)_{23}$.

Пример 7. При плане $(150, 2)_{13}$ из контролируемой партии берут случайную выборку объемом 150 изделий. Если среди изделий, отобранных в выборку, обнаружено 3 или более дефектных, то партию бракуют без дальнейшего контроля. Если $d \leq 2$, то партию принимают.

Пример 8. При плане $(80, 1)_{03}$ из партии берут случайную выборку объема $n = 80$; если в выборке оказалось $d = 2$ или более дефектных изделий, то проверяют все изделия, не попавшие в выборку. Если $d = 0$ или 1, то изделия, не попавшие в выборку, принимают без контроля.

Одноступенчатым планом контроля с двумя уровнями (n, c_1, c_2) называют

план, при котором на контроль отбирают случайную выборку объема n . Если число d дефектных изделий, обнаруженных в выборке, не более c_1 , то принимают решение D_3 ; если $c_1 < d \leq c_2$, то принимают решение D_2 ; если $d > c_2$, то принимают решение D_1 .

Двуступенчатый план контроля $(n_1, n_2, c_1, r_1, c_2, r_2)_{ij}$, $c_1 < r_1$, $c_1 < c_2$, $r_2 = c_2 + 1$ называют план, при котором из партии сначала берут случайную выборку объема n_1 . Если число d_1 дефектных изделий в этой выборке не более c_1 , то партию принимают без дальнейшего контроля с решением D_j ; если же $d_1 \geq r_1$, то бракуют без дальнейшего контроля с решением D_i . Если $c_1 < d_1 < r_1$, то берут вторую случайную выборку объема n_2 . Пусть d_2 — число дефектных изделий, обнаруженных во второй выборке, тогда, если $d_1 + d_2 \leq c_2$, то партию принимают с решением D_j . Если же $d_1 + d_2 \geq r_2$, то партию бракуют с решением D_i .

Планы последовательного контроля. Пусть изделия последовательно одно за другим отбирают на контроль. Результатам контроля сопоставляют последовательность $\varepsilon_1, \varepsilon_2, \dots$, в которой $\varepsilon_i = 1$, если i -е изделие, отобранное для контроля, оказалось дефектным, и $\varepsilon_i = 0$, если оно годное. Результаты последовательного контроля

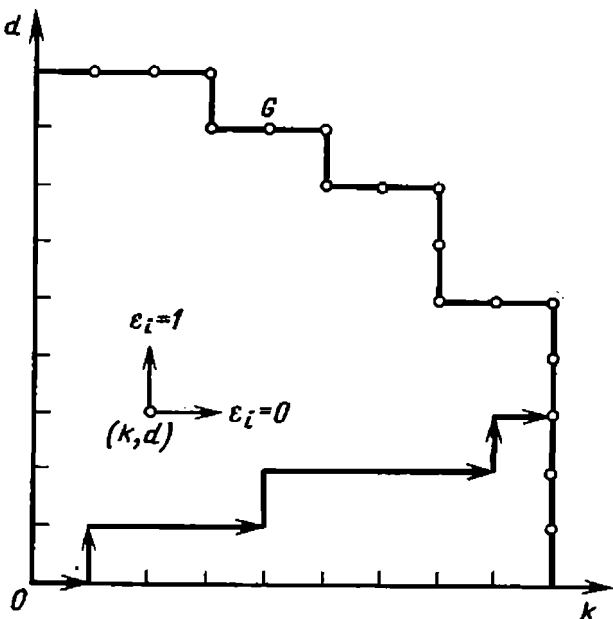


Рис. 1. Результаты последовательного контроля

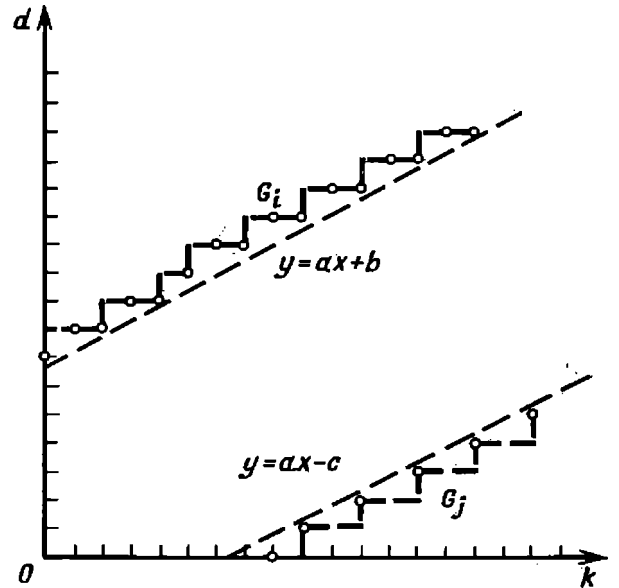


Рис. 2. План Вальда последовательного контроля

можно представить в виде движения по точкам решетки (k, d) , $k, d = 0, 1, 2, \dots$ (рис. 1). Начальной точкой является $(0, 0)$. В результате обнаружения годного изделия сдвигаемся из точки (k, d) в точку $(k + 1, d)$. В результате обнаружения дефектного изделия сдвигаемся из точки (k, d) в точку $(k, d + 1)$. В результате контроля n изделий достигается точка (k_n, d_n) , $k_n = \sum_{i=1}^n \bar{\varepsilon}_i$, $d_n = \sum_{i=1}^n \varepsilon_i$, $\bar{\varepsilon}_i = 1 - \varepsilon_i$, $k_n + d_n = n$.

Последовательность точек $(k_1, d_1), \dots, (k_n, d_n)$ называют путем из точки (k_1, d_1) в точку (k_n, d_n) , если такая последовательность может встретиться при контроле с положительной вероятностью. Точку (k, d) называют **достижимой**, если существует путь, ведущий из начальной точки $(0, 0)$ в точку (k, d) . Точку (k, d) называют **проходимой**, если существует путь, для которого точка (k, d) не является последней. Достижимую точку, не являющуюся проходимой, называют **граничной** точкой. Граничные точки образуют **границу**. При достижении границы контроль прекращают.

Планом последовательного контроля с границей G называют план с последовательной проверкой изделий, случайным образом отбираемых на контроль; контроль прекращают, когда путь при-

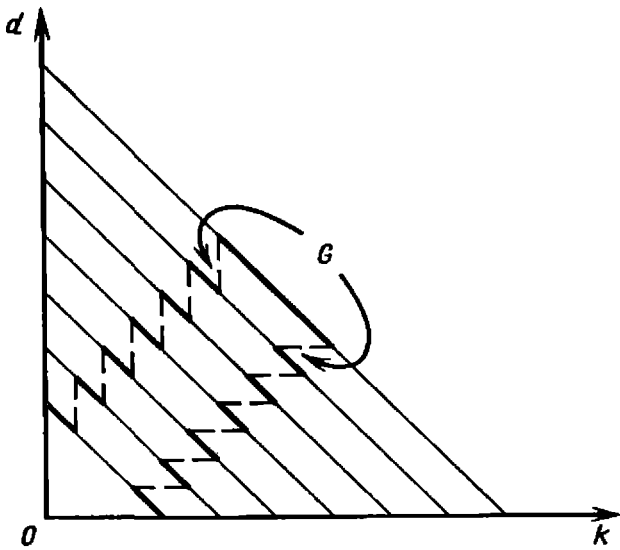


Рис. 3. Семиступенчатый план контроля

водит в точку из множества G (рис. 1). Множество G разбивают на непересекающиеся подмножества $G = G_1 + G_2 + G_3$. Если достигается точка $(k, d) \in G_j$, то принимают решение D_j .

Пример 9. Верхняя часть границы G_i плана контроля определяется ближайшими точками (k, d) , лежащими не ниже прямой $y = ax + b$ (рис. 2). Нижняя часть границы G_j определяется точками (k, d) , ближайшими к прямой $y = ax - c$, но лежащими ниже этой прямой. Последовательный план контроля с границей $G = G_i + G_j$ называют планом Вальда. Здесь $a > 0$, $b > 0$, $c > 0$.

В некоторые системы планов контроля, например *MIL STD 105D*, включены многоступенчатые планы, являющиеся разновидностью планов последовательного контроля. На рис. 3 показана граница G для семиступенчатого плана контроля. Точки границы G выделены в виде отрезков.

2. ЧИСЛОВЫЕ ПОКАЗАТЕЛИ ПЛАНОВ КОНТРОЛЯ

Вероятность забракования партии объема N , содержащей D дефектных изделий, рассматриваемую как функцию D или функцию доли дефектных изделий $q = D/N$, называют **оперативной характеристикой**.

Пример 1. Для одноступенчатых планов $(n, c)_{13}$ и $(n, c)_{23}$ оперативная

характеристика (т. е. вероятность забракования партии с решением D_i)

$$P(D) = \sum_{d=0}^c h_{N, D}^{n, d},$$

где $h_{N, D}^{n, d} = \frac{\binom{n}{d} \binom{N-n}{D-d}}{\binom{N}{D}}$ — гипергеометрические вероятности.

Пример 2. Для плана двухступенчатого контроля $(n_1, n_2, c_1, r_1, c_2, r_2)$ оперативная характеристика

$$P(D) = \sum_{d_1=0}^{c_1} h_{N, D}^{n_1, d_1} + \sum_{d_1=c_1+1}^{r_1-1} \sum_{d_2=0}^{c_2-d_1} h_{N, D}^{n_1, d_1} h_{N-n_1, D-d_1}^{n_2, d_2}$$

Для плана последовательного контроля с границей $G = G_i + G_j$, $i < j$, $i = 1, 2$ оперативная характеристика равна вероятности остановки в подмножестве точек G_i . Для вычисления оперативной характеристики можно использовать (обратные) рекуррентные соотношения

$$P(k, d) = \begin{cases} 1, & (k, d) \in G_i, \\ 0, & (k, d) \notin G_i, \end{cases} \quad (1)$$

$$P(k, d) = \frac{K-k}{N-(k+d)} P(k+1, d) + \frac{D-d}{N-(k+d)} P(k, d+1), \quad (2)$$

где $K = N - D$.

Формула (1) определяет начальные условия; формулы (2) определяют значения $P(k, d)$ для проходимых точек (k, d) через значения $P(k+1, d)$ и $P(k, d+1)$. Искомая оперативная характеристика

$$P(D) = P(0, 0). \quad (3)$$

Вычисления по формулам (1)–(3) ведут на ЭВМ.

Пусть по соглашению между поставщиком и заказчиком определены приемочная доля брака $q_1 = D_1/N$ и браковочная доля брака $q_2 = D_2/N$, $D_1 < D_2$. Вероятность забракования партии с долей брака q_1 , равную

$1 - P(D_1)$, называют вероятностью ошибки первого рода или **риском поставщика**. Вероятность приемки партии с долей брака q_2 , равную $P(D_2)$, называют вероятностью ошибки второго рода или **риском потребителя**. Здесь $P(D)$ — значение оперативной характеристики.

Риски поставщика и потребителя являются показателями эффективности используемого плана контроля.

Долю дефектных изделий $D_{0,5}/N$, при которой вероятность приемки партии предельно близка к 0,5, называют **безразличным уровнем доли дефектных изделий**.

Для безразличного уровня доли дефектных изделий выполнены неравенства

$$P(D_{0,5} - 1) < 0,5 < P(D_{0,5} + 1).$$

Пусть $L(D)$ — математическое ожидание доли дефектных изделий, принятых потребителем. $L = \max_{0 \leq D \leq N} L(D)$

называют **пределом среднего выходного качества**.

Пример 3. Для одноступенчатого плана $(n, c)_{23}$

$$L(D) = \frac{1}{N} \sum_{d=0}^c (D-d) h_{N,D}^{n,d}. \quad (4)$$

Если n/N и c/n малы, то

$$L \approx \rho_c \frac{N-n}{nN}, \quad (5)$$

где

$$\rho_c = \max_{\lambda > 0} \lambda \sum_{d=0}^c \frac{\lambda^d}{d!} e^{-\lambda},$$

$$\rho_0 = 0,368; \quad \rho_1 = 0,840; \quad \rho_2 = 1,371; \\ \rho_3 = 1,942.$$

Важной числовой характеристикой плана контроля является $m(D)$ — среднее число проконтролированных изделий. Для одноступенчатого плана (n, c) $m(D) = n$. Для двухступенчатого плана $(n_1, n_2, c_1, r_1, c_2, r_2)_{13}$

$$m(D) = n_1 + n_2 \sum_{d_1=c_1+1}^{r_1-1} h_{N,D}^{n_1,d_1}. \quad (6)$$

Для двухступенчатого плана $(n_1, n_2, c_1, r_1, c_2, r_2)_{23}$

$$m(D) = n_1 \sum_{d_1=0}^{c_1} h_{N,D}^{n_1,d_1} + (n_1 + n_2) \times \\ \times \sum_{d_1=c_1+1}^{r_1-1} \sum_{d_2=0}^{c_2-d_1} h_{N,D}^{n_1,d_1} h_{N-n_1, D-d_1}^{n_2,d_2} + \\ + N \left(\sum_{d_1=r_1}^{n_1} h_{N,D}^{n_1,d_1} + \sum_{d_1=c_1+1}^{r_1-1} \times \right. \\ \left. \times \sum_{d_2=c_2-d_1+1}^{n_2} h_{N,D}^{n_1,d_1} h_{N-n_1, D-d_1}^{n_2,d_2} \right). \quad (7)$$

Для последовательного плана с границей G и решениями D_1, D_3 расчет $m(D)$ ведут на основе (обратных) рекуррентных соотношений

$$m(k, d) = 0, \quad \text{если } (k, d) \in G, \quad (8)$$

$$m(k, d) = \\ = 1 + \frac{K-k}{N-(k+d)} m(k+1, d) + \\ + \frac{D-d}{N-(k+d)} m(k, d+1). \quad (9)$$

Уравнения (9) выполняются для проходных точек. Искомое среднее число проконтролированных изделий $m(D) = m(0, 0)$.

Если известны: e — стоимость контроля каждого изделия, a — ущерб от забракования без контроля годного изделия (входящего в непроконтролированную часть партии), то экономическим показателем является $\mathcal{E}(D)$ — среднее значение расходов на проведение контроля по заданному плану, когда число дефектных изделий в контролируемой партии равно D .

Пример 4. Для одноступенчатого плана $(n, c)_{13}$

$$\mathcal{E}(D) = \\ = cn + a \sum_{d=0}^c (N-n-D+d) h_{N,D}^{n,d}.$$

3. АПРИОРНЫЕ РАСПРЕДЕЛЕНИЯ ЧИСЛА ДЕФЕКТНЫХ ИЗДЕЛИЙ

В условиях стабильного производства число D_ω дефектных изделий в контролируемой партии можно рассматривать как случайную величину D_ω . Распределение, определяемое вероятностями

$$P\{D_\omega = D\} = f_N(D),$$

называют **априорным**.

Априорное распределение называют **биномиальным**, если

$$f_N(D) = \binom{N}{D} q^D (1-q)^{N-D}$$

Математическое ожидание доли брака $M \frac{D_\omega}{N} = q$. Априорное распределение называют **смесью биномиальных распределений**, если

$$f_N(D) = \int_0^1 \binom{N}{D} q^D (1-q)^{N-D} dG(q),$$

где $G(q)$ — **весовая функция распределения** доли дефектных изделий q .

Если априорное распределение D_ω является биномиальным, то результаты выборочного контроля $\varepsilon_1, \varepsilon_2$, (см. рис. 1) образуют последовательность Бернулли, $P(\varepsilon_i = 1) = q$. Если d — число дефектных изделий, обнаруженных в выборке объема n , то число $D_\omega - d$ дефектных изделий в непроконтролированной части партии имеет биномиальное распределение

$$P(D_\omega - d = D - d) = \binom{N-n}{D-d} q^{D-d} (1-q)^{N-n-D+d}$$

Если весовая функция распределения имеет плотность

$$g(q) = \frac{dG(q)}{dq} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} q^{a-1} (1-q)^{b-1},$$

$a > 0, b > 0,$

то априорное распределение задается вероятностями

$$f_N(D) = \binom{N}{D} \frac{a^{(D,1)} b^{(K,1)}}{(a+b)^{(N,1)}}, \quad (11)$$

где $h^{(x,1)} = h(h+1)\dots(h+x-1)$.

Распределение с вероятностями (11) называют **распределением Пуа**. При $a = b = 1$ имеем

$$f_N(D) = \frac{1}{N+1}, \quad 0 \leq D \leq N.$$

Пусть D_ω — случайная величина со средним $MD_\omega = Q_N$ и дисперсией $M(D_\omega - Q_N)^2 = V_N$. Если $V_N < Q_N(1-Q_N)N$, то априорное распределение D_ω называют **суббиномиальным**. Если $V_N = Q_N(1-Q_N)N$, то априорное распределение D_ω называют **эквibiномиальным**. Если $V_N > Q_N(1-Q_N)N$, то априорное распределение D_ω называют **супербиномиальным**.

Пусть d_ω — число дефектных изделий, обнаруженных в выборке объема n при контроле партии объема N , содержащей D_ω дефектных изделий. Случайные величины d_ω и $D_\omega - d_\omega$ отрицательно коррелированы, если распределение D_ω суббиномиальное; d_ω и $D_\omega - d_\omega$ положительно коррелированы, если априорное распределение супербиномиальное. Случайные величины d_ω и $D_\omega - d_\omega$ не коррелированы, если априорное распределение D_ω эквibiномиальное.

Пример 1. Гипергеометрическое распределение $f_N(D) = h_{N_0, D_0}^{N, D}$ является суббиномиальным.

Пример 2. Смесь биномиальных распределений задает априорное супербиномиальное распределение. При этом, если $d_\omega = d$, то условное математическое ожидание доли дефектных изделий среди изделий, не включенных в выборку, равно

$$q(d) = M\left(\frac{D_\omega - d}{N - n} \mid d_\omega = d\right),$$

является неубывающей функцией d . Таким образом, если априорное распределение числа дефектных изделий D_ω является смесью биномиальных распределений, то среднее значение

доли дефектных изделий в непроконтролированной части партии тем больше, чем больше число дефектных изделий, обнаруженных в выборке.

4. СИСТЕМЫ ПЛАНОВ КОНТРОЛЯ

В зависимости от конкретных условий производства могут быть использованы различные системы планов контроля. При выборе подходящей системы планов контроля нужно учитывать различие в интересах поставщика и потребителя. Надо учитывать также фактическую трудоемкость или более полно — стоимость контроля, а также степень защиты потребителя от приемки партий, сильно «засоренных» дефектными изделиями. Если принятие дефектных изделий ведет к недопустимо большому ущербу, то исходными являются показатели гарантии защиты потребителя от проникновения дефектных изделий. Если такие показатели выбраны, то поставщику следует использовать наиболее экономные планы контроля с заданными показателями защиты. Если же расходы на контроль и ущерб от принятия дефектных изделий соизмеримы, то при равных заданных расходах на контроль предпочтительней будет та система планов контроля, которая обеспечивает большую защиту потребителя от принятия плохих партий.

Примерами систем планов контроля по альтернативному признаку являются планы Доджа и Ромига, MILSTD-105D, ГОСТ 24660—81.

Система экономичных планов (СЭП) контроля, положенная в основу ГОСТ 24660—81, при заданных средних расходах на контроль в условиях стабильного производства обеспечивает приближенно равномерно близкую к максимально возможной вероятность забракования партии с большой долей брака.

Таблицы планов СЭП рассчитаны в предположении биномиальности априорного распределения для числа дефектных изделий в партиях, соответствующих стабильному ходу производства. Пусть $P(q_0)$ — вероятность приемки партии в условиях стабильного производства, q_0 — средняя доля

брака. Средние расходы на контроль партии при использовании одноступенчатого плана $(n, c)_{13}$

$$\mathcal{E} = a(N - n) [1 - P(q_0)] + en.$$

Относительный уровень расходов на проведение контроля

$$E = \frac{\mathcal{E}}{aN} = \left(1 - \frac{n}{N}\right) [1 - P(q_0)] + \frac{e}{a} \cdot \frac{n}{N}.$$

При заданном отношении e/a параметрами, определяющими план в таблицах СЭП, являются q_0 , aN/e и E . Параметр E является управляющим. Увеличение E усиливает жесткость контроля. Таблицы СЭП содержат одноступенчатые планы, обладающие хорошими свойствами защиты потребителя.

Система планов контроля должна обладать свойствами адаптации к изменениям уровня доли дефектных изделий. Для этого вводятся правила корректировки, связанные с изменением жесткости контроля. Эти правила предусматривают как переход от нормального контроля к усиленному, так и обратный переход от усиленного контроля к нормальному. В некоторых системах планов контроля допускается возможность перехода на **облегченный** контроль. Правила корректировки должны учитывать не только скачкообразное увеличение q — доли дефектных изделий, но и постепенное увеличение q .

В системах планов контроля могут быть использованы правила переключения с двумя уровнями (m, l_1, m_2, l_2) . При использовании таких правил переход с нормального контроля на усиленный происходит тогда, когда среди m_1 последних проконтролированных партий было забраковано l_1 партий, или среди m_2 последних проконтролированных партий было забраковано l_2 партий, $m_1 \ll m_2$, $l_1 \leq l_2$. Переход с усиленного контроля на нормальный может проводиться с использованием правил переключения с одним уровнем (m_3, l_3) , когда среди последних m_3 партий, проконтролированных в условиях усилен-

ного контроля, было забраковано не более l_3 партий.

Переключающие правила на усиленный контроль характеризуются двумя показателями: 1) $k(q)$ — средним числом партий, проверяемых до перехода на усиленный контроль, когда доля дефектных изделий равна q ; 2) $\frac{1}{k(q_0)}$ — частотой ложных переходов на усиленный контроль. Если q_1 — браковочный уровень качества, то $k(q_1)$ характеризует запаздывание в переходе от нормального контроля на усиленный контроль.

5. ПОСЛЕДУЮЩИЕ ОЦЕНКИ

При использовании любой системы планов контроля необходимо регистрировать результаты хода контроля, например, вид продукции, дату, объем партии, параметры использованного плана, число обнаруженных дефектных изделий, принятое решение и т. п. На основе накапливаемой информации о ходе контроля можно оценить его фактическую эффективность, можно следить за изменениями уровня качества продукции, предъявляемой на контроль, оценить количество дефектных изделий, проникших к потребителю. Такие расчеты проводят на основе последующих (статистических) оценок.

Пусть предъявляемая на контроль партия изделий объема N содержит D дефектных и K годных изделий, $K + D = N$. В зависимости от хода контроля и принятых по его результатам решений D_1, D_2 или D_3 имеем:

$$N = N' + N'' + N''';$$

$$D = D' + D'' + D''';$$

$$K = K' + K'' + K''',$$

где: 1) D' — число дефектных изделий, содержащихся среди N' изделий, отвергнутых без контроля, $K' = N' - D'$; 2) D'' — число дефектных изделий, обнаруженных при контроле N'' изделий из партии, $K'' = N'' - D''$; 3) D''' — число дефектных изделий среди N''' изделий, принятых потребителем, $K''' = N''' - D'''$. Кратко D''' называют принятым браком, а D — предъявленным браком.

Пример 1. Если контроль партии объема N , содержащей D дефектных изделий, проводят на основе одноступенчатого плана $(n, c)_{13}$, причем в выборке обнаружено d дефектных изделий, то $D'' = d, N'' = n$,

$$N''' = \begin{cases} N - n, & \text{если } d \leq c, \\ 0, & \text{если } d > c, \end{cases}$$

$$D''' = \begin{cases} D - d, & \text{если } d \leq c, \\ 0, & \text{если } d > c. \end{cases}$$

Пример 2. Если контроль партии объема N , содержащей D дефектных изделий, проводят на основе одноступенчатого плана $(n, c)_{23}$, то $N' = D' = K' = 0$,

$$N'' = \begin{cases} n, & \text{если } d \leq c, \\ N, & \text{если } d > c, \end{cases}$$

$$D'' = \begin{cases} d, & \text{если } d \leq c, \\ D, & \text{если } d > c, \end{cases}$$

$$D''' = \begin{cases} D - d, & \text{если } d \leq c, \\ 0, & \text{если } d > c. \end{cases}$$

Пусть проконтролировано m партий, объемы которых равны N_1, \dots, N_m . Если D_i — число дефектных изделий, содержащихся в i -й партии, а D_i''' — число принятых дефектных изделий, то $\sum_{i=1}^m D_i$ называют суммарным предъявленным браком, а $\sum_{i=1}^m D_i'''$ — сум-

марным принятым браком. Суммарный предъявленный брак и суммарный принятый брак могут быть использованы для характеристики эффективности контроля. Суммы $\sum_{i=1}^m D_i$ и $\sum_{i=1}^m D_i'''$ неизвестны, однако их можно оценить, используя последующие оценки.

Несмещенная оценка \hat{D} предъявленного брака D при последовательном плане контроля с границей G задается выражением

$$\hat{D} = N \frac{L((0, 1), (k, d))}{L((0, 0), (k, d))}, \quad (12)$$

где (k, d) — координаты точки остановки; d — число дефектных изделий,

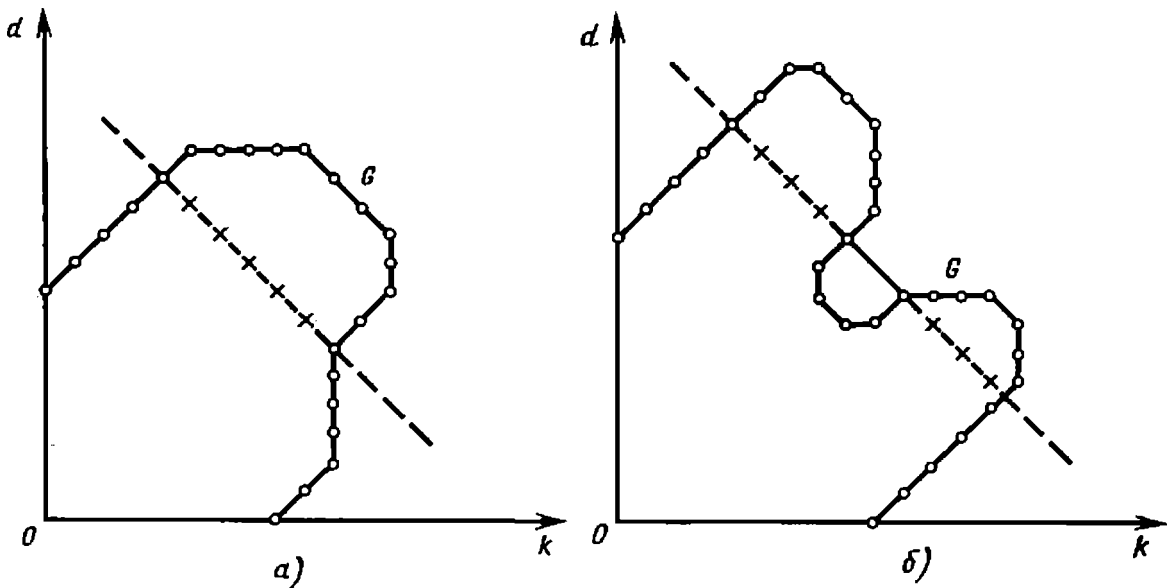


Рис. 4. Планы последовательного контроля:
 а — простой; б — непростой

обнаруженных при контроле; $(k, d) \in G$; $L((k_1, d_1), (k_2, d_2))$ — число путей, ведущих из точки (k_1, d_1) в точку (k_2, d_2) .

Несмещенная оценка \hat{V} дисперсии оценки \hat{D} , задаваемой выражением (12), имеет вид

$$\hat{V} = N(N-1) \frac{L((0, 2), (k, d))}{L((0, 0), (k, d))} + N \frac{L((0, 1), (k, d))}{L((0, 0), (k, d))} \quad (13)$$

Замечание. При вычислении оценок \hat{D} и \hat{V} на ЭВМ вместо чисел путей можно использовать вероятности путей для конкретного подобранного значения D_0 в качестве числа дефектных изделий. Расчет этих вероятностей ведется на основе обратных рекуррентных соотношений, аналогичных (1), (2), где $G_i = (k, d)$ состоит из одной точки — результатов контроля.

Пример 3. Для одноступенчатого плана контроля $(n, c)_{13}$

$$\hat{D} = \frac{dN}{n}, \quad \hat{V} = \frac{N(N-n)d(n-d)}{n^2(n-1)}$$

Пример 4. Для одноступенчатого плана контроля (n, c_1, c_2) с двумя уровнями несмещенная оценка для предъявленного брака имеет вид

$$\hat{D} = \begin{cases} \frac{N}{n}d, & \text{если } d \leq c_1 \text{ или } d \geq c_2, \\ \frac{N}{n} \frac{\sum_{d=c_1+1}^{c_2-1} dh_{N,D}^{n,d}}{\sum_{d=c_1+1}^{c_2-1} h_{N,D}^{n,d}}, & \text{если } c_1 < d < c_2. \end{cases}$$

План последовательного контроля с границей G называют **простым**, если все проходимые точки (k, d) с одинаковым значением суммы $k + d$ образуют отрезки, т. е. не разделены граничными или недостижимыми точками. На рис. 4, а показан простой план, на рис. 4, б непростой план.

План Вальда, одноступенчатые и двухступенчатые планы являются простыми.

Если план контроля является простым, то несмещенные оценки предъявленного брака в примерах 2, 3 имеют наименьшие дисперсии в классе несмещенных оценок.

Оценки \hat{D}_+^m и \hat{D}_-^m называют соответственно **верхней** и **нижней** оценками для принятого брака D^m , если для всех значений $D = 0, 1, \dots, N$

$$M\hat{D}_-^m \leq MD^m \leq M\hat{D}_+^m.$$

Для плана последовательного контроля с границей G , содержащей

точки (k, d) , $k + d < N$, не существует несмещенной оценки для принятого брака D'' . Однако можно построить верхнюю и нижнюю оценки \widehat{D}_+'' , \widehat{D}_-'' , для которых разность $M\widehat{D}_+'' - M\widehat{D}_-''$ будет малой величиной. Соответствующие расчеты \widehat{D}_-'' и \widehat{D}_+'' выполняются с помощью ЭВМ.

Если оценки \widehat{D}_-'' и \widehat{D}_+'' найдены, т. е. получены таблицы, в которых каждому возможному результату контроля (k, d) сопоставляются конкретные числовые значения, то можно аналогичным образом рассчитать на ЭВМ верхние и нижние оценки

Точность приближения определяется погрешностями, возникающими при замене точного распределения $\sum_{i=1}^m \widehat{D}_i$ нормальным распределением, а также при замене дисперсий V_i оценками \widehat{V}_i .

Пример 5. Если все контролируемые партии имеют одинаковый объем $N_i = N$ и контроль ведется с одноступенчатым планом $(n, c)_{13}$, то верхняя и нижняя границы γ -доверительного интервала для суммарного предъявленного брака определяются выражением

$$\frac{\frac{N}{n} d_{\Sigma} + \frac{u_{\alpha}^2}{2} \frac{N-n}{N-1} \pm u_{\alpha} \left[d_{\Sigma} \frac{N-n}{N-1} + \frac{u_{\alpha}^2}{4} \left(\frac{N-n}{N-1} \right)^2 - \frac{d_{\Sigma}}{mn} \left(\frac{N-n}{N-1} \right) \right]^{1/2}}{1 + \frac{u_{\alpha}^2}{nm} \frac{(N-n)}{N-1}}$$

\widehat{V}_{++} и \widehat{V}_{+-} для $M(\widehat{D}_+'' - M\widehat{D}_+'')^2 = V_+$ — дисперсии \widehat{D}_+'' , а также верхние и нижние оценки \widehat{V}_{-+} и \widehat{V}_{--} для $M(\widehat{D}_-'' - M\widehat{D}_-'')^2 = V_-$ — дисперсии \widehat{D}_-'' .

Приближенные, но асимптотически при $m \rightarrow \infty$ точные границы γ -доверительного интервала для суммарного предъявленного брака в результате контроля последовательности из m партий имеют вид

$$\sum_{i=1}^m \widehat{D}_i - u_{\frac{1-\gamma}{2}} \sqrt{\sum_{i=1}^m \widehat{V}_i}, \quad (14)$$

$$\sum_{i=1}^m \widehat{D}_i + u_{\frac{1-\gamma}{2}} \sqrt{\sum_{i=1}^m \widehat{V}_i}. \quad (15)$$

Формула (14) определяет нижнюю границу γ -доверительного интервала для $\sum_{i=1}^m D_i$; формула (15) определяет верхнюю границу γ -доверительного интервала; u_{α} — квантиль уровня $1 - \alpha$ стандартного нормального распределения.

Здесь $d_{\Sigma} = \sum_{i=1}^m d_i$, знак $+$ выбирается для верхней границы, знак $-$ для нижней границы приближенного γ -доверительного интервала для $\sum_{i=1}^m D_i$.

6. КОНТРОЛЬ НЕПРЕРЫВНОГО ПОТОКА ШТУЧНОЙ ПРОДУКЦИИ

Выборочный контроль непрерывного потока штучной продукции без разбиения на партии существенно отличается от выборочного контроля партий. Все изделия считают расположенными в порядке их производства.

Контроль по плану Доджа ведется по альтернативному признаку. Этот контроль имеет два режима: 1) случайного отбора изделий и 2) сплошной проверки. План Доджа определяется двумя параметрами: f — вероятностью выбора изделий на контроль и числом i — минимальным объемом сплошной проверки. В режиме случайного выбора каждое очередное изделие отбирают на контроль случайно с вероятностью f , $0 < f < 1$. Если ото-

бранное изделие оказалось годным, то режим случайного выбора сохраняется. Если отобранное изделие оказалось дефектным, то сплошному контролю подвергаются следующие за ним i изделий. Режим сплошного контроля сохраняется до тех пор, пока не будут обнаружены следующие друг за другом i годных изделий. Если i изделий, последовательно проконтролированные друг за другом, оказываются годными, то переходят с режима сплошного контроля на режим случайного отбора изделий на контроль. Все непроконтролированные изделия считают принятыми.

Различают две разновидности планов Доджа. Планы Доджа, в которых обнаруженные дефектные изделия заменяются (ремонтируются) годными, будем обозначать $D_1(f, i)$. Планы Доджа, в которых обнаруженные дефектные изделия удаляются из последовательности принимаемых изделий, обозначаем $D_2(f, i)$.

Основные характеристики планов Доджа рассчитывают в предположении, что в последовательности изделий $\{O_t, t = 1, 2, \dots\}$, подлежащих контролю, годные и дефектные изделия соответствуют последовательности Бернулли $\{\epsilon_t, t = 1, 2, \dots\}$, в которой вероятность встретить дефектное изделие $P\{\epsilon_t = 1\} = q$. При контроле по плану Доджа $D_j(f, i)$, $j = 1, 2$, средняя доля контролируемых изделий

$$f(q) = \frac{f}{f + (1-f)(1-q)^i}.$$

Средним уровнем выходного качества $L_j(q)$ называют математическое ожидание доли дефектных изделий среди принятых изделий при неограниченном во времени контроле по плану Доджа $D_j(f, i)$. Выполнены соотношения

$$L_1(q) = q[1 - f(q)]; \quad (16)$$

$$L_2(q) = \frac{q[1 - f(q)]}{1 - qf(q)}. \quad (17)$$

Уровнем предельного среднего выходного качества при контроле по плану Доджа $D_j(f, i)$ называют $L_j^* = \max_{0 \leq q \leq 1} L_j(q)$. Выполнены соотношения

$$L_1^* = \frac{(i+1)q_1^* - 1}{i}; \quad (18)$$

$$L_2^* = \frac{iq_2^* - 1}{i-1}, \quad (19)$$

величины q_j^* находят из уравнений

$$(1 - q_j^*)^{i-j+1} = \frac{f[(i-j+2)q_j^* - 1]}{(1-f)(1-q_j^*)},$$

где $j = 1$ для планов $D_1(f, i)$; $j = 2$ для планов $D_2(f, i)$.

Критичным планом контроля непрерывного потока продукции называют план $\Pi(f, i)$, при котором отбор изделий на контроль ведут в двух режимах: выборочного и сплошного отбора изделий на контроль. В режиме выборочного отбора каждое очередное изделие отбирают на контроль независимо с вероятностью f . При обнаружении дефектного изделия начинается режим сплошного контроля. Если i подряд проверенных изделий окажутся годными, то вновь начинается режим выборочного отбора. При обнаружении в режиме сплошного контроля каждого дефектного изделия объем контроля, необходимого для возвращения к режиму выборочного контроля, возрастает на i . Пусть L — число проконтролированных изделий от начала режима сплошного контроля до его окончания, а d — число обнаруженных дефектных изделий. Тогда $L = (d+1)i$.

Изделие, соответствующее режиму выборочного контроля, считают принятым, если оно не было проконтролировано. В случае неразрушающего контроля все проконтролированные изделия, оказавшиеся годными, считают принятыми.

Если последовательность изделий, подлежащих контролю, соответствует схеме Бернулли с долей дефектных изделий q , $iq < 1$, то средняя доля проконтролированных изделий

$$f(q) = \frac{f}{1 - (1-f)iq}.$$

Если $iq > 1$, то после контроля некоторого (случайного) числа изделий происходит переход на неограничен-

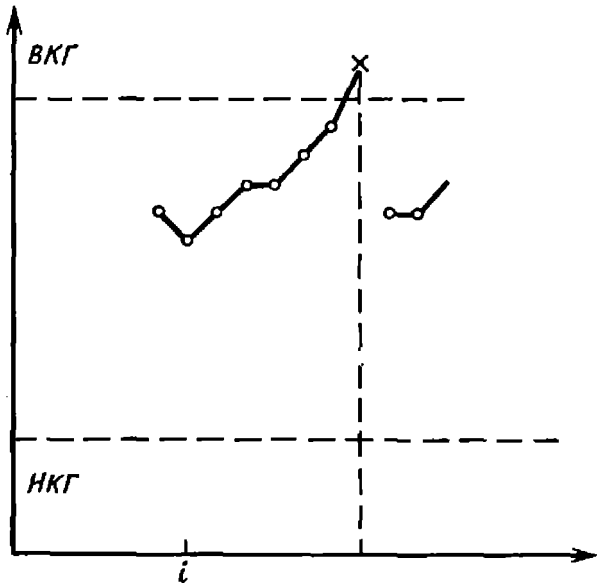


Рис. 5. Отображение результатов контроля на контрольной карте (x соответствует сигналу о необходимости наладки процесса производства)

ный сплошной контроль. В таком случае возвращение к режиму выборочного контроля возможно лишь после проведения мероприятий по уменьшению доли дефектных изделий q .

Важным показателем защиты потребителя от проникновения брака среди принятой без контроля части продукции при $q > \frac{1}{i}$ является $D(q)$ — среднее число принятых без контроля дефектных изделий до перехода на неограниченный сплошной контроль. Для критичного плана с параметрами f и i

$$D(q) = \frac{1}{1-c} \left(\frac{1}{f} - 1 \right), \quad (20)$$

где c — вероятность перехода на неограниченный сплошной контроль при условии, что первое проконтролированное изделие оказалось дефектным. Значение c находят как решение уравнения

$$c = (1 - q + qc)^i.$$

Пример 1. Пусть $q_H = 0,01$ — доля брака при стабильном ходе производства. Долю брака $q = 0,05$ считают для потребителя неприемлемой. В условиях стабильного производства решено контролировать 5% всех изготавливаемых изделий. Всего имеется

79 различных критичных планов $\Pi(f_j, i_j)$, $j = 1, 2, \dots, 79$, которые при $q = 0,01$ соответствуют контролю 5% продукции. Примерами таких планов являются $\Pi(0,0355; 30)$, $\Pi(0,0306; 40)$, $\Pi(0,0265; 50)$. Расчеты по формуле (20) показывают, что минимальное значение $D(0,05) < 40$ имеет план $\Pi(0,0306; 40)$. Таким образом, если в результате разладки доля брака увеличивается до значения $q > 0,05$, то при использовании плана $\Pi(0,0306, 40)$ с момента разладки в среднем будет принято не более 40 дефектных изделий.

При контроле непрерывного потока продукции широко используют планы контроля по количественному признаку. Каждое изделие O_i характеризуется значением параметра X_i , принимающего вещественные значения. При контроле по количественному признаку считаются выполненными следующие допущения:

а) значения X_i являются взаимно независимыми одинаково распределенными случайными величинами;

б) функция распределения X_i принадлежит к определенному параметрическому семейству.

Наиболее часто используют нормальное семейство, когда

$$P\{X_i \leq x\} = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy.$$

Часто используют семейства экспоненциальных распределений, Вейбулла — Гнеденко, логарифмически нормальных распределений.

Замечание. Допущения а), б) являются существенными. Если они не выполнены, то фактические показатели эффективности контроля будут существенно отличаться от теоретических (расчетных) значений. Неправильный выбор параметрического семейства может привести к тому, что показатели эффективности контроля по количественному признаку окажутся ниже показателей эффективности контроля по альтернативному при-

знаку. Только тщательно обоснованное применение статистического контроля по количественному признаку может служить эффективным средством управления качеством продукции.

При контроле по количественному признаку часто используют так называемые контрольные карты. В соответствии с планом контроля через случайные или детерминированные интервалы времени из контролируемой последовательности изделий берутся выборки фиксированного объема n . Обычно n невелико (3—20). В выборку отбирают либо подряд n изготовленных изделий либо отбирают случайную выборку объема n из изделий, изготовленных за определенный промежуток времени. Пусть $x_{i,1}, \dots, x_{i,n}$ — значения определяющего параметра для изделий i -й выборки. По этим значениям вычисляют базовые показатели. В ряде стандартов в качестве базовых показателей выбраны эмпи-

рические среднее $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{i,j}$ и

дисперсия $s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{i,j} - \bar{x}_i)^2$.

Если n велико, то с целью ускорения расчетов выбирают в качестве базовых показателей медиану $x_{0,5}$ или размах $R_i = \max_j x_{i,j} - \min_j x_{i,j}$. Для каж-

дого базового показателя ведется своя контрольная карта. На контрольной карте (рис. 5) по оси абсцисс откладывается номер выборки, по оси ординат — значение базового параметра. На контрольную карту наносятся **верхняя и нижняя контрольные границы** (ВКГ и НКГ). Если значение базового параметра выходит из полосы, определяемой ВКГ и НКГ, то производится подналадка технологического процесса производства. Базовые показатели \bar{x} и $x_{0,5}$ характеризуют систематическое изменение определяющего параметра, s^2 и R характеризуют разброс.

В практике контроля по количественному признаку используют и другие типы контрольных карт, учитывающие разладки различных типов. В качестве частных показателей эффективности контроля по количественному признаку учитывают среднюю частоту ложных остановок процесса на подладку и среднее время запаздывания с момента возникновения разладки до ее выявления по результатам контроля.

СПИСОК ЛИТЕРАТУРЫ

1. Беляев Ю. К. Вероятностные методы выборочного контроля. М.: Наука, 1975. 407 с.
2. Шиндовский Э., Шюрц О. Статистические методы управления качеством/Пер. с нем. М.: Мир, 1976. 597 с.

1. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ
ВЕЛИЧИН

Метод статистического моделирования состоит в использовании реализаций случайных величин и процессов для численного решения задач.

Пусть ξ — случайная величина с конечным математическим ожиданием a . Тогда, взяв достаточно большое число независимых реализаций ξ_1, \dots, ξ_n данной величины, можно оценить a средним арифметическим:

$$\bar{a} \approx \frac{1}{n} (\xi_1 + \dots + \xi_n).$$

Пусть $\xi(t)$ — любой случайный процесс, $f(\cdot)$ — функционал от него. По независимым реализациям $\xi_1(t), \dots, \xi_n(t)$ данного процесса можно оценить $a = Mf(\xi(\cdot))$ *, а именно

$$\bar{a} \approx \frac{1}{n} [f(\xi_1(\cdot)) + \dots + f(\xi_n(\cdot))].$$

Для получения реализации случайных величин и процессов необходим некоторый источник случайности. Поскольку практически реализация случайного процесса всегда выражается через соответствующие случайные величины, то источник случайности есть датчик случайных чисел.

Датчиком случайных чисел называют устройство, выдающее независимые реализации $\omega_1, \omega_2, \dots$ случайной величины ω с заданной функцией распределения $F_\omega(x)$. Данные реализации, выдаваемые в цифровом коде, называют случайными числами.

* Обозначение $f(\xi(\cdot))$ подчеркивает зависимость $f(\cdot)$ от реализации процесса $\xi(t)$, $t \geq 0$ в противоположность значению при фиксированном t .

Наиболее распространены равномерные и нормальные случайные числа. Для первых $F_\omega(x) = x$, $0 \leq x \leq 1$, т. е. ω распределено равномерно в интервале $(0, 1)$; для последних

$$F_\omega(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du,$$

т. е. ω имеет нормальное распределение с математическим ожиданием 0 и с дисперсией 1. В данной главе в дальнейшем предполагается, что ω_n — равномерные случайные числа. Поскольку ω_n выдаются в цифровом коде, то

$$\omega_n = 2^{m-1} \varepsilon_{n1} + \dots + 2^{m-2} \varepsilon_{n2} + \dots + \varepsilon_{nm},$$

где $\varepsilon_{n1}, \dots, \varepsilon_{nm}$ — двоичные разряды случайного числа. Для равномерных случайных чисел $\varepsilon_{n1}, \dots, \varepsilon_{nm}$ независимы и принимают значения 0 и 1 с вероятностью 1/2. Таким образом распределение ω_n является дискретным приближением к равномерному распределению; приближение тем точнее, чем больше m .

В реальных датчиках случайных чисел наблюдается некоторое отклонение распределения знаков ε_{ni} от равномерного закона, равно как и отклонение от независимости различных ω_n , а также знаков $\varepsilon_{n1}, \dots, \varepsilon_{nm}$ для данного n . На практике подобные отклонения при расчетах не учитываются, но принимаются специальные меры для обеспечения заданных статистических свойств случайных чисел — статистический контроль последовательности случайных чисел и преобразования, увеличивающие равномерность.

Пример. Устройство выдает независимые двоичные символы η_1, η_2 , для которых

$$P(\eta_n = 0) = \frac{1}{2}(1 + \Delta);$$

$$P(\eta_n = 1) = \frac{1}{2}(1 - \Delta),$$

где Δ — независимый параметр, колеблющийся в пределах $\pm 0,5$. Требуется построить реализацию независимых двоичных случайных чисел $\varepsilon_1, \varepsilon_2$, с соблюдением условия

$$|P(\varepsilon_n = 0) - P(\varepsilon_n = 1)| \leq 10^{-3}.$$

Решение. Будем суммировать исходные величины η_n по модулю 2, используя m величин для образования одного члена последовательности $\{\varepsilon_n\}$:

$$\varepsilon_1 = \eta_1 \oplus \dots \oplus \eta_m,$$

$$\varepsilon_2 = \eta_{m+1} \oplus \dots \oplus \eta_{2m},$$

где \oplus — символ суммы по модулю 2.

Справедливы формулы

$$P(\varepsilon_n = 0) = \frac{1}{2}(1 + \Delta^m),$$

$$P(\varepsilon_n = 1) = \frac{1}{2}(1 - \Delta^m),$$

откуда

$$|P(\varepsilon_n = 0) - P(\varepsilon_n = 1)| =$$

$$= |\Delta|^m \leq \frac{1}{2^m}.$$

Взяв $m = 10$, получим

$$|P(\varepsilon_n = 0) - P(\varepsilon_n = 1)| \leq$$

$$\leq \frac{1}{2^{10}} < 10^{-3},$$

т. е. поставленное требование удовлетворяется.

Статистической моделью оценки функционала $f(\cdot)$ называют алгоритм получения реализации $f(\cdot)$ с использованием данного источника случайности.

Статистическое моделирование часто производится по следующей схеме. Имеется математическая модель исследуемой физической системы в виде некоторого случайного процесса $\xi(t)$,

а функционал $f(\cdot)$ — тот или иной показатель качества работы системы в данной реализации процесса. Тогда построение статистической модели оценки $f(\cdot)$ сводится к построению реализации процесса $\xi(t)$ с помощью датчика случайных величин.

Ниже приведен ряд достаточно общих рекомендаций по построению реализаций случайных величин и процессов.

Реализация дискретной случайной величины (случайного выбора). Пусть имеется случайное испытание с возможными исходами $A_n, n \geq 0$, причем $P(A_n) = p_n, \sum_{n=0}^{\infty} p_n = 1$. Тре-

буется построить реализацию данного испытания*

Разобьем отрезок $[0, 1]$ на непересекающиеся множества $\Delta_n, n \geq 0$, причем так, чтобы мера Δ_n была равна p_n , например, можно положить

$$\Delta_0 = [0, p_0),$$

$$\Delta_1 = [p_0, p_0 + p_1),$$

$$\Delta_n = [p_0 + \dots + p_{n-1},$$

$$p_0 + \dots + p_n),$$

Если $\omega \in \Delta_n$, полагают, что исход испытания есть A_n .

Пусть ξ — дискретная случайная величина, принимающая значения x_n с вероятностями $p_n, n \geq 0$. Тогда реализуют аналогично предыдущему, случайное испытание с исходами $A_n, P(A_n) = p_n$. При исходе A_n полагают $\xi = x_n$. Получена реализация случайной величины ξ .

Упрощение алгоритма моделирования часто достигается использованием двоичных знаков $\varepsilon_1, \dots, \varepsilon_m$ случайного числа ω , учитывая, что любое возможное значение вектора из k знаков ($1 \leq k \leq m$) имеет вероятность 2^{-k} .

* При конечном числе исходов можно положить $p_n = 0, n > N$, поэтому в данном и подобных ему случаях особых оговорок о возможности конечного множества исходов делать не будем.

Пример. Пусть распределение $\{p_n\}$ задается таблицей:

n	0	1	2	3	4	5
p_n	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

При $\varepsilon_1 = 1$ реализуется исход A_0 . Если $\varepsilon_1 = 0$, вызывают значение ε_2 . При $\varepsilon_2 = 1$ реализуется один из исходов A_1, A_2 ; при $\varepsilon_2 = 0$ — один из исходов A_3, A_4, A_5 . Пусть $\varepsilon_2 = 1$. Тогда полагают, что A_1 реализуется при $\varepsilon_3 = 1$, A_2 — при $\varepsilon_3 = 0$. Если $\varepsilon_2 = 0$, то A_3 реализуется при $\varepsilon_3 = 1$; если же $\varepsilon_3 = 0$, вызывают значение ε_4 : имеем A_4 при $\varepsilon_4 = 1$, A_5 при $\varepsilon_4 = 0$.

Если возможные исходы испытания представлены окончными точками дерева с двоичным разветвлением, причем движение по любому пути разветвления имеет вероятность $1/2$, то выбор очередного пути осуществляется с помощью очередного, не использованного ранее знака случайного числа.

Реализация дискретной цепи Маркова. Пусть $\{\xi_n, n \geq 0\}$ — цепь Маркова с начальными вероятностями $P(\xi_0 = i) = p_i^{(0)}$ и вероятностями перехода

$$P(\xi_n = j | \xi_{n-1} = i) = p_{ij},$$

$$n \geq 1, i \geq 0, j \geq 0.$$

Построим, аналогично предыдущему, набор функций $f_i(\omega)$, таких, что $P(f_i(\omega) = j) = p_{ij}$. (Например, можно положить $f_i(\omega) = 0$ при $0 \leq \omega \leq p_{i0}$, $f_i(\omega) = 1$ при $p_{i0} \leq \omega < p_{i0} + p_{i1}$, ..., $f_i(\omega) = n$ при $p_{i0} + \dots + p_{in-1} \leq \omega \leq p_{i0} + \dots + p_{in}$).

Построим также функцию $f^{(0)}(\omega)$, равную $i \geq 0$ с вероятностью $p_i^{(0)}$. Искомый алгоритм задается рекуррентной формулой

$$\xi_0 = f^{(0)}(\omega_0), \quad \xi_n = f_{\xi_{n-1}}(\omega_n), \quad n \geq 1.$$

Пример. Требуется построить реализацию цепи Маркова $\{\xi_n, n \geq 0\}$ с возможными состояниями 0, 1, начальными вероятностями $p_0^{(0)} = 0,57$,

$p_1^{(0)} = 0,43$ и вероятностями перехода, задаваемыми таблицей:

i	j	
	0	1
0	0,88	0,12
1	0,36	0,64

Положим $f_0(\omega) = 0$ при $0 \leq \omega < 0,88$, $f_0(\omega) = 1$ при $0,88 \leq \omega < 1$; $f_1(\omega) = 0$ при $0 \leq \omega < 0,36$, $f_1(\omega) = 1$ при $0,36 \leq \omega \leq 1$; $f^{(0)}(\omega) = 0$ при $0 \leq \omega < 0,57$, $f^{(0)}(\omega) = 1$ при $0,57 \leq \omega \leq 1$. Пусть первые шесть последовательных случайных чисел задаются таблицей:

n	0	1	2
ω_n	0,855	0,714	0,291
n	3	4	5
ω_n	0,505	0,909	0,784

Тогда в данной реализации получают значения $\{\xi_n\}$, представленные в таблице:

n	0	1	2	3	4	5
ξ_n	1	1	0	0	1	1

Реализация одномерной случайной величины с заданным распределением. Пусть $F(x)$ — функция распределения. Требуется построить реализацию случайной величины ξ с данной функцией распределения.

Универсальные алгоритмы моделирования.

Метод 1. Построим график функции $y = F(x)$, дополнив его отрезками, параллельными оси Oy , до непрерывной кривой в точках разрыва (рис. 1).

Случайную величину ξ получают как абсциссу точки пересечения прямой $y = \omega$ с построенной кривой. В частности, если $F(x)$ — непрерывная функция, $F^{-1}(x)$ — обратная к ней функция, то $\xi = F^{-1}(\omega)$. Тот же результат получится, если положить $\xi = F^{-1}(\varphi(\omega))$, где $\varphi(\omega)$ — любая случайная величина, равномерно распределенная в интервале $(0, 1)$. Так, можно положить $\varphi(\omega) = 1 - \omega$.

Пример 1 (равномерное распределение в интервале (a, b)). Имеем

$$F(x) = \frac{x-a}{b-a}, \quad a \leq x \leq b;$$

$$\xi = a + (b-a)\omega;$$

можно также положить $\xi = b - (b-a)\omega$.

Пример 2 (экспоненциальное распределение с параметром λ). Имеем $F(x) = 1 - e^{-\lambda x}, \quad x \geq 0;$

$$\xi = \frac{1}{\lambda} \ln \frac{1}{1-\omega};$$

можно также положить $\xi = \frac{1}{\lambda} \ln \frac{1}{\omega}$.

Пример 3 ($F(x)$ — непрерывная кусочно-линейная функция). При $x_n \leq x \leq x_{n+1}$ имеем

$$F(x) = a_n + (a_{n+1} - a_n) \frac{x - x_n}{x_{n+1} - x_n}.$$

Полагаем

$$\xi = x_n + \frac{x_{n+1} - x_n}{a_{n+1} - a_n} (\omega - a_n)$$

при $a_n \leq \omega \leq a_{n+1}$.

Пример 4 ($F(x)$ — степенная функция). Имеем: $F(x) = x^\rho / x_0^\rho, \quad 0 \leq x \leq x_0$, где $\rho > 0$ — параметр. Полагаем

$$\xi = x_0 \omega^{1/\rho}.$$

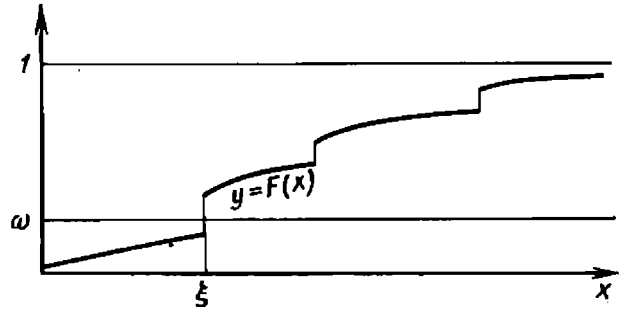


Рис. 1. Реализация случайной величины с заданной функцией распределения

Пример 5 (распределение Вейбулла). Имеем:

$$F(x) = 1 - \exp\{- (\lambda x)^\alpha\}, \quad x \geq 0,$$

где $\lambda > 0, \alpha > 0$. Полагаем

$$\xi = \frac{1}{\lambda} \left(\ln \frac{1}{1-\omega} \right)^{1/\alpha}$$

В данной формуле можно также взять ω вместо $1 - \omega$.

Пример 6 (непрерывная функция $F(x)$ — сплайн функций $F_k(x), x_k \leq x < x_{k+1}$, т. е. в каждом таком интервале $F(x)$ задается своим аналитическим выражением). Обозначим $a_k = F(x_k)$ и положим

$$\xi = F_k^{-1}(\omega), \quad a_k \leq \omega < a_{k+1}.$$

Так, для двоянного экспоненциального распределения с плотностью $p(x) = \frac{1}{2} \lambda e^{-\lambda |x|}$ можно положить

$$\xi = \begin{cases} \frac{1}{\lambda} \ln(1-2\omega) & \text{при } 0 \leq \omega < \frac{1}{2}, \\ \frac{1}{\lambda} \ln \frac{1}{2\omega-1} & \text{при } \frac{1}{2} < \omega \leq 1. \end{cases}$$

Метод 2. Пусть случайная величина ξ имеет плотность $p(x)$, причем $p(x) \leq c$,

$$\int_a^b p(x) dx = 1. \quad \text{Пусть } \omega_1,$$

ω_2, \dots — последовательные случайные числа, ν — наименьшее $n \geq 1$, для которого

$$p(a + (b-a)\omega_{2\nu}) > c\omega_{2\nu-1}.$$

Тогда полагаем

$$\xi = a + (b-a)\omega_{2\nu}.$$

Для одной реализации требуется в среднем $2(b - a)$ случайных чисел.

Обобщенный метод 2. Пусть η — случайная величина с плотностью $p_0(x)$, ξ — случайная величина с плотностью $p(x)$. Требуется получить реализацию ξ в ситуации, когда реализации η получаются достаточно просто. Допустим, что при некотором $c > 1$

$$p(x) \leq cp_0(x)$$

для всех x .

Пусть η_1, η_2, \dots — независимые реализации η , а $\omega_1, \omega_2, \dots$ — независимые случайные числа, независимые также от $\{\eta_n\}$.

Обозначим через ν наименьшее n , для которого $p(\eta_n) \geq cp_0(\eta_n)\omega_n$, и положим $\xi = \eta_\nu$. Среднее число реализаций пар (η_n, ω_n) в расчете на одну реализацию ξ равно c .

Специализированные алгоритмы моделирования для конкретных типов распределений. По центральной предельной теореме при $n \rightarrow \infty$ случайная величина

$$s_n = \frac{1}{\sqrt{n/12}} \left(\omega_1 + \dots + \omega_n - \frac{n}{2} \right)$$

сходится по распределению к стандартной нормальной случайной величине η . Отсюда возник практический способ реализации нормальной случайной величины ξ , $M\xi = a$, $D\xi = \sigma^2$, по формуле

$$\xi = a + \sigma s_n,$$

где n — достаточно большое число.

Вместо ω_n можно взять последовательные двоичные знаки ε_n ; в этом случае

$$\xi = a + \sigma s_n^*,$$

где

$$s_n^* = \frac{2}{\sqrt{n}} \left(\varepsilon_1 + \dots + \varepsilon_n - \frac{n}{2} \right).$$

Расположим $\omega_1, \dots, \omega_n$ в возрастающем порядке: $x_{(1)} < x_{(2)} < \dots < x_{(n)}$. Положим $\zeta_{k,n} = \frac{n}{\lambda} x_{(k)}$. При $n \rightarrow \infty$ случайная величина $\zeta_{k,n}$ схо-

дится по распределению к эрланговской случайной величине ξ с плотностью

$$\frac{\lambda (\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}, \quad x \geq 0.$$

Следовательно, при достаточно большом n для реализации эрланговской случайной величины ξ можно поло-

$$\xi = \frac{n}{\lambda} x_{(k)}.$$

Моделирование многомерных случайных величин. Реализация многомерной случайной величины $\xi = (\xi_1, \xi_2, \dots, \xi_m)$ сводится к реализации m одномерных величин. Для этого нужно знать условные распределения

$$F_k(x | x_1, \dots, x_{k-1}) = P(\xi_k < x | \xi_1 = x_1, \dots, \xi_{k-1} = x_{k-1}).$$

Если по правилам, относящимся к одномерным случайным величинам, построить функции $f_1(\omega), f_2(\omega; x_1), \dots, f_m(\omega; x_1, \dots, x_{m-1})$ так, что $f_1(\omega)$ имеет распределение $F_1(x)$, $f_2(\omega; x_1)$ — распределение $F_2(x | x_1)$, \dots , $f_m(\omega; x_1, \dots, x_{m-1})$ — распределение $F_m(x | x_1, \dots, x_{m-1})$, то алгоритм моделирования ξ формулируется следующим образом:

$$\xi_1 = f_1(\omega_1), \quad \xi_2 = f_2(\omega_2; \xi_1), \quad \dots, \quad \xi_m = f_m(\omega_m; \xi_1, \dots, \xi_{m-1}).$$

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — многомерная нормальная случайная величина. Без ограничения общности можно считать $M\xi_k = 0, 1 \leq k \leq m$. Обозначим $b_{ij} = M\xi_i \xi_j, 1 \leq i, j \leq m$. Для моделирования ξ используем величину $\eta = (\eta_1, \dots, \eta_m)$, где η_i — независимые нормальные случайные величины с математическим ожиданием 0 и дисперсией 1. Полагаем

$$\xi_k = a_{k1}\eta_1 + \dots + a_{kk}\eta_k.$$

Для определения a_{k1}, \dots, a_{kk} при известных $\{a_{ij}, i < k\}$ имеем линейную систему уравнений

$$\sum_{j=1}^i a_{kj} a_{ij} = b_{ik}, \quad 1 \leq i \leq k-1,$$

позволяющую последовательно определить $a_{k1}, \dots, a_{k, k-1}$.

Затем a_{kk} определяется из соотношения

$$a_{kk} = \pm \left(b_{kk} - \sum_{i=1}^{k-1} a_{ki}^2 \right)^{1/2}$$

где любой знак \pm удовлетворяет условию задачи.

Метод 2, описанный для случая одномерных величин, справедлив и для многомерных; лишь отрезок $[a, b]$ заменяется параллелепипедом $[a_1, b_1] \times \dots \times [a_m, b_m]$, $a + (b - a) \omega_{2n}$ — вектором $(a_1 + (b_1 - a_1) \omega_n^{(1)}, \dots, a_m + (b_m - a_m) \omega_n^{(m)})$, где $\omega_n^{(1)}, \dots, \omega_n^{(m)}$ — независимые случайные числа.

2. МОДЕЛИРОВАНИЕ ПУАССОНОВСКОГО ПОТОКА С ПАРАМЕТРОМ $\lambda(t)$,

$$\int_0^{\infty} \lambda(t) dt = \infty$$

Метод 1. Пусть $\gamma_1, \gamma_2, \dots$ — независимые экспоненциально распределенные случайные величины с параметром 1, $t(\gamma, t_0)$ — корень уравнения

$$\int_{t_0}^{t(\gamma, t_0)} \lambda(x) dx = \gamma.$$

Тогда моменты t_n событий потока определяются рекуррентным образом:

$$t_1 = t(\gamma_1, 0); \quad t_n = t(\gamma_n, t_{n-1}),$$

$$n \geq 2.$$

Метод 2. Реализуется пуассоновская случайная величина ν с параметром

$$\Lambda = \int_0^T \lambda(t) dt. \text{ Если } \nu \text{ приняла значение } n, \text{ то реализуется } n \text{ независимых случайных величин } \xi_1, \dots, \xi_n$$

с плотностью $\lambda(t)/\Lambda, 0 \leq t \leq T$. Расположив ξ_1, \dots, ξ_n в порядке возрастания, получим моменты событий потока в интервале $(0, T)$.

Метод 2'. При достаточно большом Λ в методе 2 вместо пуассоновской

случайной величины можно взять целое число, ближайшее к Λ .

Данный метод опасно применять в ситуации, когда оцениваются вероятности, связанные с большими скоплениями событий на данном интервале.

Метод 3. Пусть $\lambda(t) \leq \lambda$. Моделируется (например, методом 1) пуассоновский поток с постоянной интенсивностью λ . Если некоторое его событие произошло в момент t , то это событие «оставляется» с вероятностью $\lambda(t)/\lambda$, независимо от остальных событий. Тогда множество моментов «оставленных» событий и принимается в качестве искомой реализации.

Метод 4. Производится последовательность независимых испытаний, для n -го из которых вероятность успеха

$$p_n = \int_{(n-1)\Delta}^{n\Delta} \lambda(t) dt. \text{ Если в } n\text{-м испытании произошел успех, полагают, что}$$

в интервале $((n-1)\Delta, n\Delta)$ произошло событие потока. Метод 4 адекватен лишь при $\Delta \rightarrow 0$.

3. МОДЕЛИРОВАНИЕ ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Пусть цепь Маркова с непрерывным временем $\xi(t), t \geq 0$, задается начальными вероятностями состояний $p_i^{(0)} = P(\xi(0) = i), i = 0, 1, 2,$

интенсивностями перехода $\lambda_{ij}, i \neq j$, где $\lambda_{ij} dt = P(\xi(t+dt) = j | \xi(t) = i)$, и интенсивностями выхода $\lambda_i = \sum_{j \neq i} \lambda_{ij}$.

Построение реализации непрерывного справа процесса $\xi(t)$ равносильно заданию последовательности $\{(i_n, \tau_n), n \geq 0\}$, где i_n — n -е значение процесса, τ_n — длина интервала, в котором это значение остается неизменным. Траектория $\xi(t)$ реконструируется по данной последовательности таким образом: $\xi(t) = i_n, s_n \leq t < s_{n+1}$, где $s_0 = 0, s_n = \tau_1 + \dots + \tau_n, n \geq 0$. Опишем основные способы построения последовательности $\{(i_n, \tau_n), n \geq 0\}$.

Метод 1. Значение i_0 реализуется в соответствии с распределением $P(i_0 = i) = p_i^{(0)}$.

Если для некоторого $n \geq 0$ получена реализация $i_n = i$, то τ_n реализуется как независимая от всего предыдущего экспоненциально распределенная случайная величина с параметром λ_i , а i_{n+1} реализуется как величина, принимающая значение j с вероятностью

$$p_{ij} = \begin{cases} \lambda_{ij}/\lambda_i & \text{при } j \neq i, \\ 0 & \text{при } j = i. \end{cases}$$

Метод 2. Пусть $\lambda_i \leq \lambda$. Моделируется начальное состояние i_0 и пуассоновский поток однородных событий с параметром λ . Обозначим $t_1 < t_2 < \dots$ моменты событий этого потока. Траектория $\xi(t)$ последовательно определяется в интервалах $[0, t_1)$, $[t_1, t_2)$, ..., а именно:

$$\xi(t) = i_0, \quad 0 \leq t < t_1.$$

При $\xi(t_n - 0) = i$ производится случайное испытание с вероятностью успеха λ_i/λ . В случае неудачи имеем $\xi(t) = i, t_n \leq t < t_{n+1}$; в случае успеха реализуется испытание, исходом которого может быть любое j с вероятностью p_{ij} ; в таком случае

$$\xi(t) = j, \quad t_n \leq t < t_{n+1}.$$

Метод 3. Положим

$$p_{ij} = \begin{cases} \lambda_{ij} \Delta, & j \neq i, \\ 1 - \lambda_i \Delta, & j = i. \end{cases}$$

Строится реализация цепи Маркова $\{v_n\}$ с начальными вероятностями состояний $p_i^{(0)}$ и вероятностями перехода p_{ij} . Если $v_{n-1} = v_n$, полагают $\xi(t) = v_n$ при $(n-1)\Delta \leq t < n\Delta$; если $v_{n-1} = i \neq j = v_n$, полагают

$$\xi(t) = i, \quad (n-1)\Delta \leq t < (n-\omega)\Delta;$$

$$\xi(t) = j, \quad (n-\omega)\Delta \leq t < n\Delta,$$

где ω — равномерно распределенная в интервале $(0,1)$ случайная величина.

Метод адекватен при $\Delta \rightarrow 0$. Если не опасен эффект решетчатости моментов скачков процесса $\xi(t)$, можно обойтись без ω , полагая

$$\xi(t) = j, \quad (n-1)\Delta \leq t < n\Delta.$$

4. МОДЕЛИРОВАНИЕ ПОЛУМАРКОВСКИХ ПРОЦЕССОВ

Пусть $v(t)$ — полумарковский процесс. Два эквивалентных способа его определения порождают соответствующие способы непосредственного моделирования реализации процесса.

Алгоритм 1. Моделируется цепь Маркова $\{v_n\}$, описывающая последовательные состояния процесса $v(t)$ и последовательность $\{\xi_n\}$ длительностей его пребывания в различных состояниях: если $v_n = i, v_{n+1} = j$, то ξ_n есть случайная величина с функцией распределения $F_{ij}(x)$. Процесс $v(t)$ восстанавливается по $\{v_n\}$ и $\{\xi_n\}$ следующим образом:

$$v(t) = v_n, \quad s_n \leq t < s_{n+1}, \quad n \geq 0,$$

$$\text{где } s_0 = 0, s_n = \xi_1 + \dots + \xi_n, \quad n \geq 1.$$

Модификация на случай, когда начальное состояние и время пребывания в нем имеют особые распределения, очевидна.

Алгоритм 2. Пусть реализация полумарковского процесса $v(t)$ построена до момента t' , причем в момент t' произошел переход $v(t)$ в состояние i . Реализуются случайные величины η_{ij} с функциями распределения $\Phi_{ij}(x)$. Пусть $\zeta_i = \min_j \eta_{ij} = \eta_{ik}$. Полагаем

$$v(t) = i, \quad t' \leq t < t' + \zeta_i,$$

$$v(t' + \zeta_i) = k.$$

Так как ситуация возобновляется с заменой i на k и t' на $t' + \zeta_i$, имеем циклический алгоритм, позволяющий «протягивать» реализацию полумарковского процесса вдоль интервалов его постоянства. Так как между характеристиками $\{p_{ij}, F_{ij}(x)\}$ и $\{\Phi_{ij}(x)\}$ существует взаимно однозначное соответствие (по крайней мере в случае, когда функции $\Phi_{ij}(x)$ непрерывны), то один и тот же процесс можно моделировать как первым, так и вторым способами. Второй способ более нагляден. Так, при моделировании явлений надежности случайные величины $\{\eta_{ij}\}$ — длительности безотказной работы, восстановления элементов, интервалы контроля и другие величины, определяемые очевидным образом. В то же время при данном

способе моделирования при каждом переходе процесса необходим просмотр всех возможных вариантов следующего состояния.

При принципиальной простоте обоих способов моделирования возникает сложность в их реализации при большом числе состояний.

Способ 1 предполагает, что в памяти ЭВМ размещен массив информации о распределениях $F_{ij}(x)$, по которому для любых заданных i, j реализуется соответствующая величина ξ_n . При большом числе состояний рекомендуется следующий комбинированный способ реализации полумарковского процесса.

Множество пар (i, j) разбивается на два подмножества J_1 и J_2 . Для $(i, j) \in J_1$ табличным способом задается алгоритм моделирования случайной величины с функцией распределения $F_{ij}(x)$. При $(i, j) \in J_2$ реализация получается с помощью каких-либо вычислений, более сложных по сравнению с табличным способом. Предпочтением для включения в множество J_1 обладают те пары (i, j) , для которых переход процесса из i в j более вероятен.

Замена реализаций случайных величин вычислением суммы вероятностей. Пусть применяется способ моделирования реализации полумарковского процесса $v(t)$. Для каждого i образуем множества A_i, B_i состояний процесса, назвав переходы $i \rightarrow j$ высоковероятными при $j \in A_i$ и маловероятными при $j \in B_i$. Практически приемлемая точность вычислений обеспечивается условием

$$\sum_{j \in B_i} p_{ij} \leq 0,1.$$

Случайные величины η_{ij} с функциями распределения $\Phi_{ij}(x)$ моделируются лишь для $j \in A_i$, таким образом, моделируется вспомогательный полумарковский процесс $v_0(t)$ с запретом переходов из i в B_i . Кроме того, моделируется экспоненциально распределенная с параметром 1 случайная величина θ . Для любого перехода $v_0(t)$ из i в j вычисляется «вес» $g = \sum_{k \in B_i} \Phi_{ik}(\zeta_i)$, где ζ_i — время пре-

бывания $v_0(t)$ в состоянии i . «Вес» g суммируется по последовательным переходам $v_0(t)$. Когда суммарный «вес» превосходит θ , на данном шаге реализации $v(t)$ отклоняется от реализации $v_0(t)$, а именно: из состояния i происходит переход в $k \in B_i$ с вероятностью $\Phi_{ik}(\zeta)/g$, после чего снова реализуется процесс $v_0(t)$, и процедура циклически повторяется.

5. ТОЧНОСТЬ ОЦЕНОК И КОЛИЧЕСТВО РЕАЛИЗАЦИЙ

Пусть \bar{a}_n — оценка для среднего $M\xi = a < \infty$, вычисленная по n независимым реализациям ξ_1, \dots, ξ_n случайной величины ξ по формуле

$$\bar{a}_n = \frac{1}{n} (\xi_1 + \dots + \xi_n).$$

Оценка \bar{a}_n построена с абсолютной точностью $\varepsilon > 0$ и достоверностью $1 - \alpha$, если для нее справедливо соотношение

$$P(|\bar{a}_n - a| < \varepsilon) = 1 - \alpha.$$

Рассмотрим задачу определения числа n независимых реализаций случайной величины ξ , необходимого для получения оценки \bar{a}_n с абсолютной точностью ε и достоверностью $1 - \alpha$.

Если дисперсия $\sigma^2 < \infty$ случайной величины ξ известна, то число n определяется по формуле

$$n = \left[\left(\frac{x_{\alpha/2}}{\varepsilon} \right)^2 \sigma^2 \right] + 1,$$

где $[y]$ — целая часть числа y , а $x_{\alpha/2}$ — квантиль стандартизованного нормального распределения, т. е. $x_{\alpha/2}$ — решение уравнения

$$\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\{-u^2/2\} du = \frac{\alpha}{2}.$$

Чаще других используется достоверность $1 - \alpha = 0,997$, которой отвечает $x_{\alpha/2} = 3$ или $1 - \alpha = 0,95$, которой отвечает $x_{\alpha/2} = 1,96$.

Точность оценки \bar{a}_n при фиксированном n определяется по формуле

$$\varepsilon = \frac{x_{\alpha/2} \sigma}{\sqrt{n}}.$$

Следовательно, для увеличения точности оценки на порядок необходимо на два порядка увеличить число реализаций n .

Пример. Имеется случайная последовательность знаков, состоящих из 0 и 1. Знаки в каждой позиции появляются независимо с одинаковой вероятностью $p = 1/2$. Такую последовательность можно разбить на группы знаков с одинаковыми значениями в каждой группе (например, 1110001111...). Вводится случайная величина v_i , равная числу знаков в i -й группе. Требуется определить число групп n , необходимое для вычисления оценки \bar{a}_n для Mv_i с абсолютной точностью $\varepsilon = 0,1$ и достоверностью 0,95.

Решение. Случайная величина v_i имеет геометрическое распределение, начинающееся с 1. Следовательно,

$$a = Mv_i = \frac{1}{p} = 2;$$

$$\sigma^2 = Dv_i = \frac{1-p}{p^2} = 2.$$

Поскольку $\bar{a}_n = \frac{1}{n} \sum_{i=1}^n v_i$, то при больших n распределение \bar{a}_n является близким к нормальному. Поэтому можно воспользоваться формулой

$$n = \left[\left(\frac{x_{\alpha/2}}{\varepsilon} \right)^2 \sigma^2 \right] + 1 = \left[\frac{3,84 \cdot 2}{0,01} \right] + 1 = 769.$$

Таким образом, справедливо соотношение

$$P(|\bar{a}_{769} - 2| < 0,1) = 0,95.$$

Если дисперсия $\sigma^2 < \infty$ неизвестна, то для вычисления оценки \bar{a}_n существуют три метода.

Метод двойной выборки [9]. 1. По предварительным n независимым реализациям ξ_1, \dots, ξ_n случайной величины ξ вычисляют выборочные значения для $M\xi$ и $D\xi$ по формулам

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i;$$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi}_n)^2.$$

2. Определяют

$$N = \left(\frac{t_{\alpha/2, n-1} S_n}{\varepsilon} \right)^2,$$

где $t_{\alpha/2, n-1}$ — квантиль распределения Стьюдента с $n-1$ степенями свободы, т. е. $t_{\alpha/2, n-1}$ — решение уравнения

$$\int_x^\infty S_{n-1}(y) dy = \alpha/2; \quad S_{n-1}(y) =$$

$$= \frac{1}{\sqrt{\pi(n-1)}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \times$$

$$\times \left(1 + \frac{y^2}{n-1}\right)^{-\frac{n}{2}},$$

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1} e^{-x} dx, \quad \beta \geq 0.$$

3. Число реализаций, необходимое для получения оценки $\bar{\xi}$, для $M\xi$ с абсолютной точностью ε и достоверностью $1 - \alpha$ будет

$$K = [N] + 1.$$

Эффективность метода зависит от удачного выбора n . Большие значения n снижают $t_{\alpha/2, n-1}$ и, следовательно, уменьшают число $[N - n]$ дополнительных реализаций, однако случай $N < n$ приводит к потере информации по первым n реализациям.

Последовательный метод [9].

1. Строят первую реализацию ξ_1 случайной величины ξ .

2. Строят реализацию ξ_n , $n \geq 2$, независимую от предыдущих ξ_1, \dots, ξ_{n-1} реализаций случайной величины ξ ,

и вычисляют оценку S_n^2 для дисперсии $D\xi$:

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i;$$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi})^2 = \\ = \frac{1}{n-1} \left[\sum_{i=1}^n \xi_i^2 - \frac{1}{n} \left(\sum_{i=1}^n \xi_i \right)^2 \right].$$

3. Вычисляют точность δ оценки $\bar{\xi}$ по формуле

$$\delta = \frac{t_{\alpha/2, n-1} S_n}{\sqrt{n}}.$$

4. Проверяют условие $\delta \leq \varepsilon$. Если условие не выполнено, то осуществляют переход на шаг 2 с заменой n на $n+1$; в противном случае в качестве числа реализаций, необходимого для вычисления $\bar{\xi}$ с абсолютной точностью ε и достоверностью $1-\alpha$, берется число n .

Недостаток метода состоит в большом объеме вычислений. Правило остановки вычислений можно проверять не на каждом шаге, а через $K > 1$ шагов. При этом общая эффективность метода повысится.

Многошаговый метод [9]. В отличие от последовательного метода число реализаций на $(l+1)$ -м шаге вычисляют по формуле

$$K_{l+1} = [(N_l - M_l) \theta],$$

где N_l — число реализаций, необходимое для получения оценки с абсолютной точностью ε и достоверностью $1-\alpha$, вычисленное после l -го шага; M_l — суммарное число реализаций, полученных за l шагов, $0 < \theta < 1$.

Алгоритм, реализующий данный метод, состоит в следующем.

1. Выбирают начальное значение K_1 ($l=1$, $M_l = K_l$).

2. Осуществляют $M_l - M_{l-1}$, $M_0 = 0$, независимых реализаций $\xi_{M_{l-1}+1}, \dots, \xi_{M_l}$ случайной величины ξ .

3. Аналогично последовательному методу вычисляют величины

$$\bar{\xi} = \frac{1}{M_l} \sum_{i=1}^{M_l} \xi_i;$$

$$S_{M_l}^2 = \frac{1}{M_l-1} \sum_{i=1}^{M_l} (\xi_i - \bar{\xi})^2;$$

$$\delta = \frac{t_{\alpha/2, M_l-1} S_{M_l}}{\sqrt{M_l}},$$

и проверяют условие $\delta < \varepsilon$. Если условие выполнено, то число реализаций, необходимых для получения оценки $\bar{\xi}$ с абсолютной точностью ε и достоверностью $1-\alpha$ равно M_l , и вычисления прекращают; в противном случае переходят на следующий шаг.

4. Определяют величины N_l и K_{l+1} по формулам

$$N_l = \left(\frac{t_{\alpha/2, M_l-1} S_{M_l}}{\delta} \right)^2$$

$$K_{l+1} = [(N_l - M_l) \theta].$$

5. Число шагов l увеличивают на 1 ($l = l+1$) и после вычислений величины

$$M_l = \sum_{i=1}^l K_i$$

осуществляют переход на шаг 2.

Оценка \bar{a}_n построена с относительной точностью ε и достоверностью $1-\alpha$, если выполнено соотношение

$$P \left(\left| \frac{\bar{a}_n - a}{a} \right| < \varepsilon \right) = 1 - \alpha.$$

Для определения числа реализаций n , необходимого для достижения заданной относительной точности ε и достоверности $1-\alpha$, на практике часто используют следующую формулу:

$$n = \inf \left\{ N \geq L : N > \right. \\ \left. > \left(\frac{x_{\alpha/2}}{\varepsilon} \right)^2 \frac{S_N^2}{\bar{a}_N} \right\},$$

где L — некоторое наперед заданное число реализаций.

Существуют и другие практические подходы к оценке числа реализаций n (в частности, когда дисперсия σ^2 бесконечна) [6].

Оценка точности и число реализаций для регенерирующих процессов. Пусть задана совокупность независимых одинаково распределенных пар $\{(\beta_i, X_i(t)), 0 \leq t < \beta_i\}$, первый элемент которых является неотрицательной случайной величиной, а второй — случайным процессом, определенным на $0 \leq t < \beta_i$. Обозначим:

$$\tau_i = \sum_{j=1}^i \beta_j,$$

$$\mu(t) = \max(i \mid \tau_i \leq t).$$

Регенерирующим процессом $\{X(t), t \geq 0\}$ называют случайный процесс вида

$$X(t) = X_{\mu(t)+1}(t - \tau_{\mu(t)}), \quad t \geq 0,$$

где τ_1, τ_2, \dots — последовательность моментов регенерации, а β_1, β_2, \dots — последовательность длительностей периодов регенерации.

Для регенерирующих процессов число реализаций принимает двойной смысл в зависимости от того, подразумевается под ними число периодов регенерации n или системное время моделирования T . Однако к дополнительным трудностям это обстоятельство не приводит, поскольку при больших значениях T и n справедливо соотношение

$$T \approx nM\beta_1.$$

Обозначим через $\omega(t)$ процесс накопления вида

$$\omega(t) = \int_0^t X(u) du;$$

$$y_i = \omega(\tau_{i+1}) - \omega(\tau_i),$$

а через $\sigma_{y_1}^2, \sigma_{\beta_1}^2$ дисперсии случайных величин y_1 и β_1 . Тогда число реализаций n при известных $\sigma_{y_1}^2 < \infty, M y_1, \sigma_{\beta_1}^2 < \infty, M\beta_1$ (хотя наличие подобной информации о процессе маловероятно) в случае абсолютной по-

грешности определяется соотношением

$$n = \left[\left(\frac{x_{\alpha/2} \sigma_{\beta_1}}{\epsilon M \beta_1} \right)^2 \right] + 1,$$

а в случае относительной погрешности — соотношением

$$n = \left[\left(\frac{x_{\alpha/2} \sigma_{y_1}}{\epsilon M y_1} \right)^2 \right] + 1.$$

При неизвестных величинах, входящих в предыдущие два соотношения, поступают следующим образом [14]. Пусть

$$Z = \{a_{10}, a_{01}, m_{20}, m_{02}, m_{11}\},$$

где a_{ij} и m_{ij} — соответственно начальные и центральные выборочные моменты двумерной случайной величины $x_k = (\beta_k, y_k)$, вычисленные по n периодам регенерации:

$$a_{ij} = \frac{1}{n} \sum_{k=1}^n \beta_k^i y_k^j;$$

$$m_{ij} = \frac{1}{n} \sum_{k=1}^n (\beta_k - a_{10})^i (y_k - a_{01})^j.$$

Пусть T_1 — число периодов регенерации, необходимое для построения оценки \bar{a} для $a = M y_1 / M \beta_1$ с абсолютной погрешностью ϵ , T_2 — с относительной погрешностью ϵ , T_3 — системное время моделирования, обеспечивающее получение абсолютной, а T_4 — относительной погрешности оценки \bar{a} с достоверностью $1 - \alpha$. Тогда

$$T_i = \left(\frac{x_{\alpha/2}}{\epsilon} \right)^2 H_i(Z), \quad i = \overline{1, 4},$$

где

$$H_1(Z) = \frac{m_{02}}{a_{10}^2} - 2m_{11} \frac{a_{01}}{a_{10}^3} + \frac{a_{01}^2}{a_{10}^4} m_{20};$$

$$H_2(Z) = \frac{m_{02}}{a_{01}^2} - 2 \frac{m_{11}}{a_{10} a_{01}} + \frac{m_{20}}{a_{10}^2};$$

$$H_3(Z) = a_{01} H_1(Z);$$

$$H_4(Z) = a_{10} H_2(Z).$$

6. НЕКОТОРЫЕ ОБЩИЕ МЕТОДЫ УМЕНЬШЕНИЯ ДИСПЕРСИИ

Метод Монте-Карло — наиболее универсальный метод определения характеристик сложных систем. В то же время для построения оценки \bar{a} искомой величины $a = M\xi < \infty$ с относительной погрешностью ε и достоверностью $1 - \alpha$ требуется число испытаний

$$n \sim c(\varepsilon, \alpha) \frac{1}{a}.$$

Таким образом, при малых значениях a этот метод обладает весьма существенным недостатком — большой трудоемкостью, что в ряде случаев делает практически невозможным получение каких-либо нетривиальных оценок для a .

Пусть $\sigma^2 = D\xi < \infty$. Предположим, что несмещенная оценка для a может быть построена N различными методами, причем σ_i^2 и τ_i — соответственно дисперсия оценки в одной реализации и среднее время одной реализации при вычислении i -м методом ($1 \leq i \leq N$). В качестве критерия сравнения различных методов естественно выбрать величину $(\tau_i \sigma_i^2)^{-1}$, называемую эффективностью i -го метода (см., например, [6]). Метод i , для которого $\tau_i \sigma_i^2 \leq \tau_j \sigma_j^2$, $j \neq i$, позволяет строить наиболее точные оценки искомой величины a при заданном времени моделирования. Поскольку, как правило, величины σ_i^2 и τ_i неизвестны, то на практике используют их выборочные значения S_{ni}^2 и $\bar{\tau}_i$. Рассмотрим некоторые общие методы уменьшения дисперсии, позволяющие сократить объем вычислений, необходимый для достижения заданной точности и достоверности оценки [4, 6, 22].

Метод выделения главной части. Предположим, что искомую характеристику системы можно представить в виде

$$a = \int_{-\infty}^{\infty} f(x) P(dx),$$

где $P(\cdot)$ — вероятностное распределение на $R = (-\infty, +\infty)$, а $f(x)$, $x \in R$ — функция, такая, что

$\int_{-\infty}^{\infty} |f(x)| P(dx) < \infty$. Пусть удалось подобрать достаточно близкую к $f(x)$ функцию $g(x)$, такую, что $b = \int_{-\infty}^{\infty} g(x) P(dx)$ вычисляется в явном виде. Тогда

$$a = b + \int_{-\infty}^{\infty} [f(x) - g(x)] P(dx).$$

В качестве оценки для a , построенной в n реализациях, выбирают

$$\bar{a} = b + \frac{1}{n} \sum_{i=1}^n [f(\xi_i) - g(\xi_i)],$$

где ξ_i , $i \geq 1$ — независимые случайные величины с распределением $P(\cdot)$. Обобщение данного метода на случай многомерного пространства R^m очевидно.

Метод существенной выборки. Предположим, что искомая характеристика системы может быть представлена в виде

$$a = M\varphi(\xi),$$

где ξ — многомерная случайная величина с плотностью $p(x)$, $x \in R^m$. Пусть η — многомерная случайная величина с плотностью $q(x)$, $x \in R^m$, причем $q(x) \neq 0$, если $p(x) > 0$, $x \in R^m$. Тогда

$$M\varphi(\xi) = \int [\varphi(x) p(x)/q(x)] q(x) dx = M\varphi(\eta) p(\eta)/q(\eta).$$

Таким образом, для вычисления $M\varphi(\xi)$ можно вместо реализаций случайной величины ξ использовать независимые реализации η_1, η_2, \dots случайной величины η , т. е. в качестве несмещенной оценки \bar{a} выбирают

$$\bar{a} = \frac{1}{n} \sum_{k=1}^n \varphi(\eta_k) p(\eta_k)/q(\eta_k).$$

причем

$$D\bar{a} = \frac{1}{n} \sigma^2,$$

где

$$\sigma^2 = D \{ \varphi(\eta) p(\eta) / q(\eta) \} =$$

$$= \int [\varphi^2(x) p^2(x) / q(x)] dx - a^2.$$

Функцию $q(x)$ можно выбрать так, чтобы дисперсия σ^2 приняла минимальное значение, а именно, если

$$q(x) = |\varphi(x)| p(x) / \int |\varphi(t)| p(t) dt,$$

то

$$\sigma^2 = \left[\int |\varphi(x)| p(x) dx \right]^2 - a^2.$$

Если же $\varphi(t)$ — знакопостоянная функция, то при таком выборе $q(x)$ $\sigma^2 = 0$.

Поскольку в выражение для плотности $q(x)$ входит неизвестная величина $\int |\varphi(t)| p(t) dt$, то использовать оптимальную плотность $q(x)$ не представляется возможным. В то же время, если плотность $q(x)$ выбрать пропорциональной $|\varphi(x)| p(x)$, то можно добиться существенного уменьшения дисперсии оценки.

Пример [22]. Пусть имеется система, состоящая из невозстанавливаемых элементов, разделенных на m групп. Заданы вероятности $q_i, i = 1, \dots, m$, отказа элементов i -й группы в течение заданного интервала времени $(0, T)$. Число элементов в i -й группе равно $N_i, N = N_1 + \dots + N_m$ — общее число элементов. Состояние системы в момент T задается N -мерным вектором

$$\xi = (\xi_1, \dots, \xi_{N_1}, \xi_{N_1+1}, \dots, \xi_{N_1+N_2}, \dots, \xi_{N_1+\dots+N_{m-1}+1}, \dots, \xi_N),$$

компоненты которого независимы и равны единице в случае неисправности соответствующего элемента и нулю в противном случае. Распределение вероятностей вектора ξ имеет вид:

$$p(x; q) = q_1^{x_1} (1 - q_1)^{1-x_1} \dots q_1^{x_{N_1}} \times \\ \times (1 - q_1)^{1-x_{N_1}} q_2^{x_{N_1+1}} \times$$

$$\times (1 - q_2)^{1-x_{N_1+1}} \dots q_2^{x_{N_1+N_2}} \times$$

$$\times (1 - q_2)^{1-x_{N_1+N_2}} \dots q_m^{x_N} \times$$

$$\times (1 - q_m)^{1-x_N},$$

где $q = (q_1, \dots, q_m); x = (x_1, \dots, x_N), x_i, i = 1, \dots, N$, принимают лишь два значения: 0 или 1.

Требуется определить вероятность того, что в момент T система неисправна:

$$a = \sum_x \varphi(x) p(x; q),$$

где $\varphi(x) = 1$, если при состоянии x система неисправна, и $\varphi(x) = 0$ в противном случае.

Пусть $q^* = (q_1^*, \dots, q_m^*)$, где $q_i^* > q_i, i = 1, \dots, m$. В качестве оценки для a выберем

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n \varphi(x^{*(i)}) \frac{p(x^{*(i)}; q)}{p(x^{*(i)}; q^*)},$$

где $x^{*(i)}, i = 1, 2, \dots$ — независимые реализации случайного вектора ξ^* с распределением $p(x; q^*)$. При соответствующем выборе значений $\{q_j^*\}$ дисперсия оценки \bar{a} существенно меньше дисперсии соответствующей оценки, полученной методом непосредственного моделирования. Действительно, пусть исследуется система, состоящая из N последовательно соединенных элементов $m = 1, q = q_1, q^* = q_1^*, N = N_1$. Обозначим через $\theta(q^*)$ относительную среднюю квадратическую погрешность:

$$\theta(q^*) = \sqrt{n D \bar{a} / a}.$$

Точность оценки \bar{a} тем выше, чем меньше $\theta(q^*)$. Следующая таблица иллюстрирует изменение $\theta(q^*)$ в зависимости от q^* :

$$q = 10^{-2}$$

N	2				10			
	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$
10^{-2}	0,16	7,01	10^{-2}	0,04	3,07	0,04	0,08	0,16
0,32	1,63	1,12	0,64	1,07	1,52	1,19	1,32	

$$q = 10^{-4}$$

N	2				10			
	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$	q^*	$\theta(q^*)$
10^{-4}	0,2048	70,7	10^{-4}	0,0512	31,6	0,1024	1,25	0,2048
0,4096	1,42	1,02	0,8192	1,46	1,65	1,25	1,66	

Метод расслоенной выборки. Пусть искомая характеристика системы может быть представлена в виде

$$a = \int_X f(x) P(dx),$$

где $P(\cdot)$ — вероятностная мера на X , $\int_X f^2(x) P(dx) < \infty$.

Предположим, что $X = A_1 \cup \dots \cup A_m$, $A_i \cap A_j = \emptyset$, $i \neq j$. Тогда

$$a = \sum_{i=1}^m P(A_i) \int_{A_i} f(x) \frac{P(dx)}{P(A_i)}.$$

Каждое слагаемое суммы, стоящей в правой части последнего соотношения, может быть вычислено методом Монте-Карло. А именно, если $\xi_j^{(i)}$, $j = 1, \dots, n_i$ — независимые реализации случайной величины $\xi^{(i)}$ с распределением $P(dx)/P(A_i)$, $x \in A_i$, то в качестве оценки \bar{a} выбирают

$$\bar{a} = \sum_{i=1}^m \frac{1}{n_i} P(A_i) \sum_{j=1}^{n_i} f(\xi_j^{(i)}).$$

Очевидно, $M\bar{a} = a$ и

$$D\bar{a} = \sum_{i=1}^m \frac{1}{n_i} P(A_i) \int_{A_i} f^2(x) P(dx) - \left[\sum_{i=1}^m \frac{1}{n_i} \left[\int_{A_i} f(x) P(dx) \right]^2 \right].$$

Пусть $n = n_1 + \dots + n_m$. Если разбиение $X = A_1 \cup \dots \cup A_m$ и число n фиксированы, то минимум $D\bar{a}$ достигается при

$$n_j = \frac{P(A_j) \sigma_j}{\sum_{i=1}^m P(A_i) \sigma_i} n$$

и равен

$$D\bar{a} = \frac{1}{n} \left[\sum_{i=1}^m P(A_i) \sigma_i \right]^2$$

Здесь

$$\begin{aligned} \sigma_j^2 &= Df(\xi^{(j)}) = \\ &= \frac{1}{P(A_j)} \int_{A_j} f^2(x) P(dx) - \\ &- \left[\frac{1}{P(A_j)} \int_{A_j} f(x) P(dx) \right]^2 \end{aligned}$$

В реальных задачах дисперсии $\{\sigma_j^2\}$, как правило, заранее неизвестны, а известны лишь вероятности $P(A_j)$. Если разбиение $X = A_1 \cup \dots \cup A_m$ и число n фиксированы, то при $n_j = nP(A_j)$

$$D\bar{a} \leq \frac{1}{n} \left[\int_X f^2(x) P(dx) - a^2 \right].$$

Таким образом, любое разбиение и выбор n_1, \dots, n_m , пропорциональных мере множеств A_1, A_2, \dots, A_m , ведет к уменьшению дисперсии по сравнению с обычным методом Монте-Карло. Если же число точек выбирают не пропорционально мерам множеств A_1, A_2, \dots, A_m , то этот факт в общем случае не имеет места. В некоторых задачах для построения достаточно точных оценок целесообразно применять сочетание метода существенной выборки с методом расслоенной выборки [20].

Пример [20]. При анализе надежности сетей связи возникает такая задача. Пусть связный граф имеет R ребер. Заданы вероятности $q_i, i = 1, \dots, R$, разрыва ребер в фиксированном интервале времени $(0, T)$. Требуется найти вероятность a того, что в момент T между двумя заданными вершинами отсутствует путь вследствие разрыва некоторых ребер. Пусть

$$\begin{aligned} x_i &= \begin{cases} 0, & \text{если } i\text{-е ребро есть,} \\ 1 & \text{в противном случае,} \end{cases} \\ x &= (x_1, \dots, x_R), \quad w(x) = \sum_{i=1}^R x_i, \end{aligned}$$

$$y(x) = \begin{cases} 0, & \text{если при состоянии } x \\ & \text{путь имеется,} \\ 1 & \text{в противном случае,} \end{cases}$$

$$\begin{aligned} p(x) &= \prod_{i=1}^R q_i^{x_i} (1 - q_i)^{1-x_i}, \quad a = \\ &= \sum_{x: y(x)=1} p(x). \end{aligned}$$

Пусть при разрыве любых d ребер путь между заданными вершинами сохраняется, т. е. $y(x) = 0$ при $w(x) \leq d$. При решении данной задачи вначале можно использовать метод существенной выборки, а затем — метод расслоенной выборки. В качестве функции $q(x)$ выберем

$$q(x) = \begin{cases} 0, & \text{если } w(x) \leq d, \\ \frac{1}{c} q^{w(x)} (1 - q)^{R-w(x)}, & \text{если } w(x) > d, \end{cases}$$

$$c = \sum_{j=d+1}^R C_R^j q^j (1 - q)^{R-j},$$

где q — некоторая вероятность (например, $q = (q_1 \dots q_R)^{1/R}$).

Далее воспользуемся методом расслоенной выборки. Обозначим: $\xi = (\xi_1, \dots, \xi_R)$ — случайный вектор, имеющий распределение $q(\cdot)$;

$$\begin{aligned} X &= \{x : w(x) \geq d + 1\}, \quad A_i = \\ &= \{x : w(x) = i\}, \quad i = d + 1, \quad R, \\ P_i &= P(\xi \in A_i) = \\ &= \frac{1}{c} C_R^i q^i (1 - q)^{R-i}, \\ & \quad i = d + 1, \quad R. \end{aligned}$$

Пусть $\eta^{(i)}, i = d + 1, \dots, R$, — случайные векторы, имеющие распределение $q(x)/P_i, x \in A_i$, а $x_j^{(i)}, j = 1, \dots, n_i$, — независимые реализации вектора $\eta^{(i)}$. Тогда

$$\begin{aligned}
 a &= \sum_x y(x) p(x) = \\
 &= \sum_{x: \omega(x) > d+1} \frac{y(x) p(x)}{q(x)} q(x) = \\
 &= \sum_{\substack{i=d+1 \\ x: \omega(x)=i}}^R P_i \frac{y(x) p(x) q(x)}{P_i} = \\
 &= \sum_{\substack{i=d+1 \\ x: \omega(x)=i}}^R C_{Ri}^i y(x) p(x) \frac{q(x)}{P_i}.
 \end{aligned}$$

Окончательная оценка при $n_i = nP_i$ имеет вид

$$\bar{a} = \sum_{i=d+1}^R C_{Ri}^i \frac{1}{n_i} \sum_{j=1}^{n_i} y(x_j^{(i)}) p(x_j^{(i)}).$$

Метод зависимых испытаний. Точность моделирования методом Монте-Карло можно увеличить, если использовать зависимые отрицательно коррелированные случайные величины. А именно, пусть искомая характеристика системы может быть представлена в виде

$$a = M\varphi(\eta_1, \dots, \eta_N),$$

где η_1, \dots, η_N — независимые случайные величины с функциями распределения $F_1(x), \dots, F_N(x)$.

Предположим далее, что функция $\varphi(\cdot)$ монотонно не убывает (или монотонно не возрастает) по каждому из аргументов при фиксированных остальных. Пусть для каждого $i = 1, \dots, N$ существует обратная функция $F_i^{-1}(x)$. Положим

$$\begin{aligned}
 \xi_1 &= \varphi(F_1^{-1}(\omega_1), \dots, F_N^{-1}(\omega_N)); \\
 \xi_2 &= \varphi(F_1^{-1}(\omega'_1), \dots, F_N^{-1}(\omega'_N)),
 \end{aligned}$$

где $(\omega_1, \omega'_1), (\omega_N, \omega'_N)$ — независимые двумерные случайные величины, причем ω_i и ω'_i в общем случае зависимые равномерно распределенные в $(0, 1)$ случайные величины.

Тогда

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 + 2K(\xi_1, \xi_2),$$

где

$$\begin{aligned}
 K(\xi_1, \xi_2) &= M(\xi_1 - a)(\xi_2 - a) = \\
 &= M\xi_1\xi_2 - a^2.
 \end{aligned}$$

Если ω_i и ω'_i независимы, то

$$K(\xi_1, \xi_2) = 0 \text{ и } D(\xi_1 + \xi_2) = 2D\xi_1.$$

Если же выбрать $\omega'_i = 1 - \omega_i$, то $K(\xi_1, \xi_2) \leq 0$ и

$$D(\xi_1 + \xi_2) \leq 2D\xi_1.$$

Таким образом, для уменьшения дисперсии оценки можно использовать следующий прием: на всех реализациях с нечетными номерами использовать независимые числа ω_i , а на всех реализациях с четными номерами использовать числа $1 - \omega_i$. Тогда в качестве оценки за $2n$ реализаций получают

$$\bar{a} = \frac{1}{2n} \sum_{i=1}^{2n} \xi_i,$$

где

$$\begin{aligned}
 \xi_{2j-1} &= \varphi(F_1^{-1}(\omega_{(j-1)N+1}), \\
 &F_N^{-1}(\omega_{jN}));
 \end{aligned}$$

$$\begin{aligned}
 \xi_{2j} &= \varphi(F_1^{-1}(1 - \omega_{(j-1)N+1}), \\
 &F_N^{-1}(1 - \omega_{jN})), \quad j = 1, \dots, n;
 \end{aligned}$$

$$D\bar{a} = \frac{D\xi_1}{2n} + \frac{K(\xi_1, \xi_2)}{2n} \leq \frac{D\xi_1}{2n}.$$

7. АНАЛИТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ

Аналитико-статистические методы — совокупность методов ускоренного моделирования, основанных на сочетании аналитических и статистических методов. Аналитико-статистические методы основаны на методе малого параметра: среди исходных характеристик системы выделяется та, которую можно выбрать в качестве малого параметра ϵ . Затем, используя различные аналитические методы, искомая характеристика системы представляется в виде ряда по степеням ϵ , причем коэффициенты этого ряда интерпретируются как математические ожи-

дания функционалов от некоторых вспомогательных случайных процессов и могут быть найдены с помощью метода статистического моделирования. С практической точки зрения во многих случаях достаточно построить статистическую оценку для коэффициента при главном члене ряда по степеням ε , оценив остаток этого ряда. Применительно к оценке характеристик высоконадежных систем данный метод развивался в работах [7, 10—12, 16, 17, 26].

Рассмотрим применение аналитико-статистического метода для оценки характеристик цепи Маркова, зависящей от малого параметра [11].

Пусть имеется однородная цепь Маркова $\{v_n, n \geq 0\}$ с конечным или счетным множеством состояний X и характеристиками, зависящими от малого параметра $\varepsilon > 0$:

$$P(v_0 = i) = p_i^{(0)} = \sum_{m=0}^{\infty} \varepsilon^m p_i^{(0)[m]},$$

$$P(v_{n+1} = j | v_n = i) = p_{ij} = \sum_{m=0}^{\infty} \varepsilon^m p_{ij}^{[m]},$$

где степенные ряды сходятся при $0 \leq \varepsilon \leq \varepsilon_0$. Задано множество $A \subset X$, в которое цепь Маркова попадает с вероятностью единица за конечное число шагов, исходя из любого состояния. На этом множестве задана ограниченная числовая функция $f(j)$, $|f(j)| \leq 1$, $j \in A$. Пусть $\gamma \geq 0$ — первый момент попадания v_n в множество A , $\zeta = f(v_\gamma)$. Требуется построить статистические оценки коэффициентов разложения $M\zeta$ по степеням ε . Решение данной задачи позволяет построить приближенную зависимость $M\zeta$ от ε для всех значений ε из интервала $(0, \varepsilon_0)$.

Пусть для некоторого $c < \frac{1}{\varepsilon_0}$ выполнены неравенства

$$\sum_{i \in X} |p_j^{(0)[m]}| \leq c^m, \quad \sum_{j \in X} |p_{ij}^{[m]}| \leq c^m,$$

$$i \in X, m \geq 0, \text{ и } p_{ij}^{[0]} = \delta_{ij},$$

$$p_{ij}^{[m]} = 0, m \geq 1 \text{ при } i \in A, j \in X,$$

где δ_{ij} — символ Кронекера.

Введем случайные величины $v_n^{[m]}$, $\theta_n^{[m]}$, $n \geq 0$, $\alpha_n^{[m]}$, $n \geq 1$. Пусть $v_0^{[m]}$ — случайная величина с распределением

$$P(v_0^{[m]} = j) = \frac{|p_j^{(0)[m]}|}{\sum_{k \in X} |p_k^{(0)[m]}|}, \quad j \in X,$$

определенная для тех $m \geq 0$, для которых делитель положителен;

$$\theta_0^{[m]} = \begin{cases} 0, & \text{если } \sum_{k \in X} p_k^{(0)[m]} = 0, \\ \text{sign} \left(p_{v_0^{[m]}}^{(0)[m]} \right) \sum_{k \in X} |p_k^{(0)[m]}| & \\ \text{в противном случае.} \end{cases}$$

Предположим, что значения $v_n^{[m]}$ и $\theta_n^{[m]}$ построены для всех $m \geq 0$. Пусть $U_{n+1}^{[m]}$ — множество целых неотрицательных l , таких, что $v_n^{[l]} \in \bar{A}$, $\theta_n^{[l]} \neq 0$ и $\sum_{k \in X} |p_{v_n^{[l]}k}^{[m-l]}| > 0$.

Число элементов в множестве $U_{n+1}^{[m]}$ обозначим $|U_{n+1}^{[m]}|$. Если $|U_{n+1}^{[m]}| = 0$, то полагают $\theta_{n+1}^{[m]} = 0$. В противном случае строят случайную величину $\alpha_{n+1}^{[m]}$, принимающую любое значение из $U_{n+1}^{[m]}$ с вероятностью $|U_{n+1}^{[m]}|^{-1}$. Если $v_n^{[l]} = i$, $\alpha_{n+1}^{[m]} = l$, то $v_{n+1}^{[m]}$ — случайная величина, равная $j \in X$ с вероятностью $|p_{ij}^{[m-l]}| / \sum_k |p_{ik}^{[m-l]}|$.

Если $v_{n+1}^{[m]} = j$, то

$$\theta_{n+1}^{[m]} = \text{sign} (p_{ij}^{[m-l]}) \times \\ \times |U_{n+1}^{[m]}| \sum_k |p_{ik}^{[m-l]}| \theta_n^{[l]}.$$

Пусть $\gamma^{[m]}$ — наименьшее $n \geq 0$, для которого $v_n^{[m]} \in A$. Положим

$$\zeta^{[m]} = \theta_{\gamma^{[m]}}^{[m]} f(v_{\gamma^{[m]}}^{[m]}).$$

Пусть существует $N < \infty$, такое, что с вероятностью 1 $\gamma^{[m]} \leq N$, $m \geq$

≥ 0 . Тогда $M\xi$ разлагается в абсолютно сходящийся ряд вида

$$M\xi = \sum_{m=0}^{\infty} \varepsilon^m M\xi^{[m]},$$

т. е. $\xi^{[m]}$ служит несмещенной оценкой коэффициента при ε^m разложения $M\xi$ по степеням ε . Для практических расчетов можно воспользоваться соотношением

$$\left| M\xi - \sum_{n=0}^N \sum_{m=0}^{s-1} \varepsilon^m b^{(n)} [m] \right| \leq \sum_{n=0}^N \frac{1}{n!} \sum_{m=s}^{\infty} (m+1) \dots (m+n) (\varepsilon)^m,$$

где

$$b^{(n)} [m] = M\theta_n^{[m]} f(v_n^{[m]}) \times \times I(v_n^{[m]} \in A);$$

$I(B)$ — индикатор события B . Число s выбирают исходя из требований к точности вычислений.

Если предположение относительно ограниченности $\gamma^{[m]}$, $m \geq 0$, не выполнено, то необходимо ввести дополнительные предположения относительно «скорости» попадания в множество A (см. [11]).

8. ОЦЕНКА РАСПРЕДЕЛЕНИЯ ВРЕМЕНИ БЕЗОТКАЗНОЙ РАБОТЫ СИСТЕМ, ОПИСЫВАЕМЫХ РЕГЕНЕРИРУЮЩИМИ ПРОЦЕССАМИ

Функционирование многих технических систем имеет циклический характер, и по истечении каждого цикла ситуация, которая может привести или не привести к наступлению редкого события, возобновляется. Математической моделью для описания систем такого типа является схема регенерирующих процессов.

Пусть $\xi(t)$, $t \geq 0$, — регенерирующий процесс, ξ_1, ξ_2, \dots — независимые одинаково распределенные случайные величины (длительности периодов регенерации), $s_i = \xi_1 + \dots + \xi_i$, $i \geq 1$, $s_0 = 0$, — моменты регенерации. Предположим, что в полуинтервале $(s_n, s_{n+1}]$ с вероятностью $q > 0$ может произойти некоторое событие A . Обозначим:

ν — наименьшее значение n , при котором в полуинтервале $(s_n, s_{n+1}]$ произошло событие A ;

$$\gamma = s_\nu = \sum_{i=1}^{\nu} \xi_i;$$

$F_0(x)$ — условная функция распределения ξ_i при условии, что в полуинтервале $(s_{i-1}, s_{i-1} + \xi_i]$ событие A не произошло.

Функция распределения случайной величины γ определяется по формуле полной вероятности

$$\Phi(x) = P(\gamma < x) = q \sum_{n=0}^{\infty} (1-q)^n P(s_n^0 < x),$$

где $s_n^0 = \xi_1^0 + \dots + \xi_n^0$, $s_0^0 = 0$, $\{\xi_i^0\}$ — независимы и имеют распределение $F_0(x)$.

В качестве примера описанной схемы рассмотрим процесс функционирования периодически сменяемого элемента. А именно, в моменты s_n , $n \geq 0$, включается новый элемент. Величина ξ_i — минимум двух величин: η_i и τ_i , где η_i — время безотказной работы элемента, τ_i — время, после которого неотказавший элемент заменяют. Если

$$G(x) = P(\eta_i < x), \quad T(x) = P(\tau_i < x),$$

то

$$P(\xi_i < x) = F(x) = 1 - [1 - G(x)][1 - T(x)].$$

Редким событием является отказ элемента до его замены. Данное событие может произойти в полуинтервале $(s_n, s_{n+1}]$ с вероятностью

$$q = P(\eta_i < \tau_i) = \int_0^{\infty} G(x) dT(x).$$

Если $G(x) = 1 - e^{-\lambda x}$, $x \geq 0$, и $M\tau_i^2 < \infty$, то справедливо неравенство

$$\lambda \int_0^\infty x dT(x) - \frac{\lambda^2}{2} \int_0^\infty x^2 dT(x) \leq q \leq \leq \lambda \int_0^\infty x dT(x),$$

позволяющее строить достаточно точные оценки для q при малых значениях λ .

Некоторые оценки и асимптотические результаты для функции $\Phi(x)$ [11]:

1. Если $F_0(x)$ — фиксированное распределение и

$$\int_0^\infty x dF_0(x) = a < \infty, \quad \text{то } q \rightarrow 0,$$

$$\lim_{q \rightarrow 0} \Phi(ax/q) = 1 - e^{-x}, \quad x \geq 0.$$

Предельные теоремы в схеме серий получены в работе [24].

2. При любых фиксированных q и a для $\Phi(x)$ нельзя указать никакой нетривиальной верхней границы. Другими словами, при любых фиксированных $a > 0$, $T > 0$, $q > 0$, $\varepsilon > 0$ существует такая функция распределения $F_0(x)$, удовлетворяющая условию

$$\int_0^\infty x dF_0(x) = a, \quad \text{что}$$

$$\Phi(T) > 1 - \varepsilon.$$

3. Если $\sigma^2 = \int_0^\infty (x-a)^2 dF_0(x) < \infty$, то

$$\Phi(ax/q) \leq 1 - e^{-x} \{1 - 3(qx)^{1/3} \times \times (\sigma/2a)^{2/3} [1 + o(1)]\}, \quad q \rightarrow 0.$$

4. Пусть для фиксированного $T > 0$ $\rho = 1 - F_0(T)$,

$$a = \int_0^T x dG(x), \quad G(x) = \begin{cases} F_0(x)/F_0(T) & \text{при } x \leq T, \\ 1 & \text{при } x > T. \end{cases}$$

Тогда для любого $\varepsilon \in (0, 1)$ справедливы оценки

$$\Phi(ax/q) \geq \left\{ 1 - \frac{Tq(1+\varepsilon)}{\varepsilon^2 a [x - q(1+\varepsilon)]} \right\} \times \times \frac{q}{q + \rho(1-q)} \left[1 - (1-q)^{\frac{x}{q(1+\varepsilon)}} \right];$$

$$\Phi(ax/q) \leq 1 - (1-q)^{\frac{x}{q(1-\varepsilon)}} \times \times \left[1 - \frac{Tq(1-\varepsilon)}{\varepsilon^2 ax} \right].$$

5. Пусть $F_0(x)$ зависит от малого параметра $q > 0$:

$$F_0(x) = F_0^{(q)}(x)$$

и существует функция $\rho = \rho(q) = = 0(q)$, $T = T(q)$ и $a = a(q)$, для которых $Tqa = 0(1)$, $q \rightarrow 0$,

$$\int_0^T x dF_0(x) = a, \quad F_0(T) = 1 - \rho. \quad \text{Тог-}$$

да равномерно по $x \geq 0$

$$\Phi\left(\frac{ax}{q}\right)_{q \rightarrow 0} \rightarrow 1 - e^{-x}.$$

6. Если $H_0(t)$ — функция восстановления процесса восстановления $\{s_n^p, n \geq 1\}$, то

$$\Phi(x) \leq q [1 + H_0(x)], \quad x \geq 0.$$

7. Справедливо равенство (см. [26])

$$\Phi(x) = q [1 + W(x)], \quad x \geq 0,$$

где $W(x)$ — среднее число восстановлений в интервале $(0, x)$ у обрывающегося процесса восстановления $\{s_n, n \geq 1\}$, у которого длительности между восстановлениями — случайные величины ξ_1, ξ_2, \dots , и в момент $s_n = = \xi_1 + \dots + \xi_n, n \geq 1$ может произойти обрыв с вероятностью q .

Свойства 6 и 7 позволяют строить весьма точные оценки для $\Phi(x)$ методом статистического моделирования. Для этого достаточно аналитически вычислить вероятность q , а затем методом Монте-Карло построить оценку для $H(x)$ или $W(x)$ в точке x .

Пример [26]. Пусть $\nu(t)$ — полумарковский процесс с конечным множеством состояний X и $A \subset X$ — фиксированное подмножество его со-

стояний, $v(+0) = i \in A$ — с вероятностью единица. Обозначим через $Q_{kj}(x)$ вероятность того, что процесс $v(t)$ будет непрерывно находиться в состоянии k время меньше, чем x , и следующим его состоянием будет j , $Q_{kj} = Q_{kj}(+\infty)$, $Q_k = \sum_{j \in A} Q_{kj}$.

Предполагается, что вложенная цепь Маркова, определяемая переходными вероятностями $\{Q_{kj}\}$, неприводима. Требуется найти вероятность $P(x) = P(\tau_i < x)$, где τ_i — время непрерывного пребывания процесса $v(t)$ в подмножестве A .

Процесс $v(t)$ является регенерирующим, моментами регенерации которого являются моменты его возвращения в состояние i . Тогда

$$q = Q_i + \sum_{k \in A \setminus \{i\}} Q_{ik} q_{ki},$$

где вероятности q_{ki} удовлетворяют системе линейных алгебраических уравнений

$$q_{ki} = Q_k + \sum_{r \in A \setminus \{i\}} Q_{kr} q_{ri}, \quad k \in A \setminus \{i\}.$$

Пусть процесс $v(t)$ в момент выхода из подмножества A обрывается. Тогда для вероятности $P(x)$ справедлива оценка

$$P(x) \leq \Phi(x) = q [1 + W(x)],$$

где $W(x)$ — число возвращений процесса $v(t)$ в состояние i в интервале $(0, x)$.

9. ОЦЕНКА НЕСТАЦИОНАРНЫХ ХАРАКТЕРИСТИК НАДЕЖНОСТИ СИСТЕМ АНАЛИТИКО-СТАТИСТИЧЕСКИМ МЕТОДОМ

Основными показателями надежности, характеризующими работу высоконадежных систем, являются нестационарные характеристики — вероятность безотказной работы системы в заданном интервале времени $(0, T)$ и коэффициент готовности системы в фиксированный момент t . Ведущая роль нестационарных характеристик объясняется тем, что в реальных системах длительность безотказной работы от-

дельных элементов значительно больше длительности интервала, в котором исследуется надежность системы, т. е. система не успевает войти в стационарный режим. При достаточно простых и естественных условиях [13, 24] для широкого класса систем доказана асимптотическая экспоненциальность времени ζ до первого отказа, т. е. для любого фиксированного $x > 0$

$$P\left(\frac{\zeta}{a(\varepsilon)} < x\right) \xrightarrow{\varepsilon \rightarrow 0} 1 - e^{-x},$$

где $\varepsilon > 0$ — некоторый малый параметр, характеризующий соотношение между исходными характеристиками системы (например, ε — отношение среднего времени восстановления элемента к среднему времени его безотказной работы), а $a(\varepsilon)$ — нормирующий множитель, который в некоторых случаях может быть вычислен через стационарные характеристики системы [13]. На практике последнее соотношение используют следующим образом: при малых значениях $\varepsilon > 0$

$$P(x) = P(\zeta < x) \approx P_{ac}(x) = 1 - e^{-\frac{x}{a(\varepsilon)}}.$$

В то же время существуют многочисленные примеры систем, для которых

$$P(x)/P_{ac}(x) \xrightarrow[\varepsilon \rightarrow 0]{x \rightarrow 0} \infty,$$

т. е. при малых значениях $x > 0$ аппроксимация вероятности $P(x)$ отказа системы в $(0, x)$ асимптотической функцией $P_{ac}(x)$ может привести к весьма существенному искажению результата. Этот же эффект имеет место и для нестационарного коэффициента готовности, если пытаться аппроксимировать его стационарным (см. приведенные ниже примеры).

Вычисление нестационарных вероятностей состояний у альтернирующего процесса восстановления. Пусть имеется некоторый элемент, поведение которого описывается альтернирующим процессом восстановления, т. е. элемент попеременно находится в рабочем и отказовом состояниях. Пусть

далее $F(x)$ и $G(x)$ — функции распределения соответственно длительности безотказной работы и восстановления элемента, ξ и η — независимые случайные величины с функциями распределения $F(x)$ и $G(x)$,

$$v(t) = \begin{cases} 0, & \text{если в момент } t \\ & \text{элемент исправен} \\ 1 & \text{в противном случае,} \end{cases}$$

$$v(0) = 0 \text{ с вероятностью } 1, P(t) = \mathbf{P}(v(t) = 1).$$

Алгоритм построения несмещенной оценки $\bar{P}(t)$ для $P(t)$.

1. Реализуют случайную величину η . Пусть $\eta = x$. При отказе элемента длительность его восстановления будет равна x . Поэтому, если $\xi \in (t - x, t)$, то в момент t элемент неисправен. Вероятность этого события есть

$$A_1 = F(t) - F(t - x), \text{ причем } F(t - x) = 0 \text{ при } x > t.$$

2. Реализуют случайную величину ξ^* с функцией распределения

$$F^*(y) = \mathbf{P}\{\xi < y \mid \xi \in (t - x, t)\}.$$

Если $\xi^* = z$, то по формуле полной вероятности оценку $\bar{P}_1(t)$ в одной реализации определяют согласно рекуррентной формуле

$$\bar{P}_1(t) = A_1 + (1 - A_1) I(z + x < t) \bar{P}_1(t - z - x),$$

где $I(B)$ — индикатор события B . Далее описанную процедуру повторяют с заменой t на $t - z - x$.

Сформулированный алгоритм может быть описан следующим эквивалентным образом.

Пусть x_1, x_2, \dots — независимые реализации случайной величины η , $A_1 = F(t) - F(t - x_1)$, z_1 — реализация случайной величины ξ^* с функцией распределения $F_1^*(y) = \mathbf{P}\{\xi < y \mid \xi \in (t - x_1, t)\}$. Пусть далее для некоторого $n \geq 2$ построены величины $A_1, z_1, \dots, A_{n-1}, z_{n-1}$ и $\gamma_{n-1} = z_1 + x_1 + \dots + z_{n-1} + x_{n-1} < t$. Тогда вычисляют

$$A_n = F(t - \gamma_{n-1}) - F(t - \gamma_{n-1} - x_n)$$

и реализуют случайную величину ξ_n^* с функцией распределения

$$F_n^*(y) = \mathbf{P}\{\xi < y \mid \xi \in (t - \gamma_{n-1} - x_n, t - \gamma_{n-1})\}.$$

Пусть $\xi_n^* = z_n$. Если $z_n + x_n < t - \gamma_{n-1}$, то вычисляют A_{n+1} и z_{n+1} . Если же $z_n + x_n > t - \gamma_{n-1}$, то в качестве оценки $\bar{P}_1(t)$ выбирают

$$\begin{aligned} \bar{P}_1(t) &= A_1 + A_2(1 - A_1) + \dots \\ &\dots + A_n(1 - A_1) \dots (1 - A_{n-1}) = \\ &= 1 - \prod_{i=1}^n (1 - A_i), \mathbf{M}\bar{P}_1(t) = P(t). \end{aligned}$$

Два способа реализации случайных величин ξ_n^* с функциями распределения $F_n^*(x)$. 1. Реализуют последовательность ξ_1, ξ_2, \dots независимых одинаково распределенных случайных величин с функцией распределения $F(x)$. Пусть k_1, k_2, \dots — номера тех случайных величин, для которых выполнено соотношение $\xi \in (t - \gamma_{n-1} - x_n, t - \gamma_{n-1})$. Тогда в качестве независимых реализаций случайной величины ξ_n^* выбирают ξ_{k_1}, ξ_{k_2} .

Данный способ целесообразно применять, если $M\eta$ — малая величина.

2. Реализуют две случайные величины ω_1 и ω_2 . Если

$$\omega_1 < \frac{[1 - F(t - \gamma_{n-1})] / [1 - F(t - \gamma_{n-1}) + F(t - \gamma_{n-1} - x_n)]},$$

то $\xi_n^* > t - \gamma_{n-1}$ и ее значение нет смысла находить, так как реализация окончена. В противном случае ξ_n^* имеет распределение $F(y)/F(t - \gamma_{n-1} - x_n)$, $y \in (0, t - \gamma_{n-1} - x_n)$. Данный способ целесообразно применять в том случае, когда построение реализаций величины ξ достаточно трудоемко, а для построения реализаций ξ_n^* может быть использован метод Неймана или одна из его модификаций.

Пример 1. Пусть $F(x) = 1 - e^{-\lambda x}$, $G(x) = 1 - e^{-\mu x}$. Тогда $P(t)$ вычисляют по точной формуле

$$P(t) = \frac{\lambda}{\lambda + \mu} [1 - e^{-(\lambda + \mu)t}];$$

$$P_{ст} = \lim_{t \rightarrow \infty} P(t) = \frac{\lambda}{\lambda + \mu}.$$

При $\lambda = 1$, $\mu = 50$ $P_{ст} = 0,01961$. Значения оценки $\bar{P}(t)$, выборочной дисперсии $D(t)$ и $P(t)$, полученные при различных t , приведены в следующей таблице (для построения оценок при каждом t было выполнено 1000 реализаций):

Замечание. Приведенный метод может быть использован для определения нестационарного коэффициента готовности $K(t)$ систем, поведение которых описывается m независимыми альтернирующими процессами восстановления. Пусть $v(t) = (v_1(t), \dots, v_m(t))$ — вектор, характеризующий состояние системы в момент t , где

t	10^{-4}	10^{-3}	10^{-2}	10^{-1}	$5 \cdot 10^{-1}$
$P(t)$	$9,974 \cdot 10^{-9}$	$9,749 \cdot 10^{-5}$	$7,833 \cdot 10^{-3}$	$1,949 \cdot 10^{-2}$	$1,961 \cdot 10^{-2}$
$\bar{P}(t)$	$9,971 \cdot 10^{-3}$	$9,761 \cdot 10^{-5}$	$7,866 \cdot 10^{-3}$	$2,005 \cdot 10^{-2}$	$1,947 \cdot 10^{-2}$
$D(t)$	$2,2 \cdot 10^{-11}$	$1,5 \cdot 10^{-8}$	$1,0 \cdot 10^{-5}$	$3,4 \cdot 10^{-4}$	$3,6 \cdot 10^{-4}$

Пример 2. Пусть $F(x) = 1 - e^{-\sqrt{x}}$, $x > 0$, $G(x) = 100x$, $x \in (0, 0, 01)$, $G(x) = 1$, $x \geq 0,01$, $G(x) = 0$, $x \leq 0$.

В этом случае для нахождения точного значения $P(t)$ формул в явном виде не существует. Стационарная вероятность $P_{ст}$ имеет вид

$$P_{ст} = \frac{M\eta}{M\xi + M\eta} = 0,002494.$$

Значения оценки $\bar{P}(t)$ и выборочной дисперсии $D(t)$ при различных t приведены в следующей таблице, из которой следует, что $P(t)$ резко возрастает на начальном интервале, а затем монотонно и довольно медленно убывает к своему стационарному значению (при каждом t было проведено 1000 реализаций):

$$v_i(t) = \begin{cases} 0, & \text{если в момент } t \text{ } i\text{-й} \\ & \text{элемент исправен,} \\ 1 & \text{в противном случае,} \\ i = 1, & m; \end{cases}$$

$$\varphi(v(t)) = \begin{cases} 0, & \text{если при состоянии} \\ & v(t) \text{ система исправна,} \\ 1 & \text{в противном случае.} \end{cases}$$

Тогда

$$P(t) = 1 - K(t) = \sum_v \varphi(v) \prod_{i=1}^m P_i^{v_i}(t) [1 - P_i(t)]^{1-v_i},$$

где $v = (v_1, \dots, v_m)$, и сумма берется по всем возможным состояниям системы v ; $P_i(t)$ — вероятность того, что i -й элемент неисправен в момент t .

t	10^{-6}	10^{-4}	10^{-3}	$5 \cdot 10^{-3}$
$\bar{P}(t)$	$2,995 \cdot 10^{-4}$	$9,875 \cdot 10^{-3}$	$2,894 \cdot 10^{-2}$	$4,422 \cdot 10^{-2}$
$D(t)$	$1,6 \cdot 10^{-16}$	$5,5 \cdot 10^{-7}$	$4,6 \cdot 10^{-5}$	$6,5 \cdot 10^{-4}$

t	10^{-2}	10^{-1}	1	20
$\bar{P}(t)$	$3,225 \cdot 10^{-2}$	$9,211 \cdot 10^{-3}$	$3,915 \cdot 10^{-3}$	$2,619 \cdot 10^{-3}$
$D(t)$	$5,4 \cdot 10^{-4}$	$1,2 \cdot 10^{-4}$	$4,0 \cdot 10^{-5}$	$3,4 \cdot 10^{-5}$

Если система состоит из одинаковых элементов и $\varphi(v) = 1$ при $v_1 + \dots + v_m \geq k$, $\varphi(v) = 0$ при $v_1 + \dots + v_m < k$, то последняя формула имеет более простой вид

$$P(t) = \sum_{i=k}^m C_m^i P_1^i(t) [1 - P_1(t)]^{m-i}$$

Вычисление нестационарного коэффициента готовности восстанавливаемых систем [17]. Суть метода состоит в следующем. Пусть $\zeta(u)$, $u \geq 0$, — непрерывный справа марковский процесс, описывающий поведение системы, а C — некоторое подмножество его состояний, включающее в себя множество отказов состояний системы. При $\zeta(u) = \zeta \in C$ вводится событие $B(u, \zeta)$, связанное с поведением процесса $\zeta(v)$ в интервале (u, t) , при наступлении которого система в момент t неисправна. Марковский процесс $\zeta(u)$, $u \geq 0$, множество C и событие $B(u, \zeta)$ выбирают так, чтобы вероятность $P(B(u, \zeta))$ вычислялась в явном виде для любых $u \in [0, t]$ и $\zeta \in C$. Алгоритм построения оценки $\bar{P}_1(t)$ в одной реализации формулируют следующим образом.

1. Методом имитационного моделирования строят траекторию процесса $\zeta(u)$ до первого момента u_1 попадания $\zeta(u)$ в множество C . Если $u_1 > t$, то полагают $\bar{P}_1(t) = 0$. Пусть $u_1 \leq t$.

2. Вычисляют вероятность $Q_1 = P(B(u_1, \zeta(u_1)))$.

3. Строят условную траекторию процесса $\zeta(u)$ в полуинтервале $(u_1, t]$ при условии, что событие $B(u_1, \zeta(u_1))$ не произошло, а именно: методом непосредственного моделирования строят траектории процесса $\zeta(u)$ в интервале (u_1, t) , отбрасывая те из них, на кото-

рых произошло событие $B(u_1, \zeta(u_1))$. В качестве условной траектории выбирают первую траекторию, на которой событие $B(u_1, \zeta(u_1))$ не произошло. Пусть $\tau_1 \in [u_1, t]$ — первый момент, когда стало известно, что событие $B(u_1, \zeta(u_1))$ не произошло. Если $\tau_1 = t$, то реализация окончена и $\bar{P}_1(t) = Q_1$. Пусть $\tau_1 < t$.

4. Если для любых $u \in [\tau_1, t]$ $\zeta(u) \notin C$, то полагают $\bar{P}_1(t) = Q_1$. В противном случае определяют $u_2 = \inf \{u \mid u \geq \tau_1, \zeta(u) \in C\}$.

5. Вычисляют $Q_2 = P(B(u_2, \zeta(u_2)))$.

Далее повторяют описанную процедуру с заменой индекса 1 на 2. Пусть для некоторого $l \geq 1$ $u_{l+1} > t$. Тогда в качестве оценки $\bar{P}_1(t)$ выбирают

$$\begin{aligned} \bar{P}_1(t) &= Q_1 + Q_2(1 - Q_1) + \dots + \\ &+ Q_l(1 - Q_1) \dots (1 - Q_{l-1}) = \\ &= 1 - \prod_{j=1}^l (1 - Q_j). \end{aligned}$$

Эффективность описанного алгоритма существенно зависит от выбора процесса $\zeta(u)$, множества C и события $B(u, \zeta)$. Проиллюстрируем данный метод на примере дублированной системы.

Система состоит из двух одинаковых элементов, находящихся в нагруженном резерве. Имеется одно ремонтное устройство. Длительности безотказной работы и восстановления элементов имеют соответственно функции распределения $F(x)$ и $G(x)$. Требуется найти вероятность $P(t)$ того, что в момент t неисправны оба элемента.

В качестве марковского процесса $\zeta(u)$ выбирают

$$\zeta(u) = (v_1(u), v_2(u);$$

$$\gamma_1(u), \gamma_2(u)), u \geq 0,$$

где

$$v_i(u) = \begin{cases} 0, & \text{если в момент } u \text{ } i\text{-й} \\ & \text{элемент исправен,} \\ -1, & \text{если элемент находится} \\ & \text{на восстановлении,} \\ 1, & \text{если элемент находится} \\ & \text{в очереди на восстановление;} \end{cases}$$

$$\gamma_i(u) = \begin{cases} \sup \{x: v_i(u-x) = 0\}, & \text{если } v_i(u) = 0, \\ \sup \{x: v_i(u-x) = -1\}, & \text{если } v_i(u) = -1, \\ 0, & \\ \text{если } v_i(u) = 1, \quad i = 1, 2. \end{cases}$$

Пусть

$$C = \{\zeta = (v_1, v_2; \gamma_1, \gamma_2): v_1 = -1 \text{ или } v_2 = -1\},$$

$$B(u, \zeta) = \{\text{до момента } t \text{ не окончится восстановление неисправного в момент } u \text{ элемента и в интервале } (u, t) \text{ откажет исправный элемент}\},$$

$$P(B(u, \zeta)) = \prod_{i: v_i' = -1} \frac{1 - G(t - u + \gamma_i)}{1 - G(\gamma_i)} \times$$

$$\times \prod_{i: v_i = 0} \frac{F(t - u + \gamma_i) - F(\gamma_i)}{1 - F(\gamma_i)}.$$

Далее оценку $\bar{P}_1(t)$ для $P(t)$ строят согласно описанному выше алгоритму. В качестве моментов $\{\tau_i\}$ выбирают моменты окончаний восстановлений элементов, а в качестве $\{u_i\}$ — последовательность моментов изменений состояний системы, таких, что $|v(u_i - 0)| \neq 1, |v(u_i + 0)| = 1$, где $|v(u)|$ — число неисправных элементов в момент u . Другой способ выбора $\zeta(u), C$ и $B(u, \zeta)$ приведен в работе [17].

Пример. Пусть $F(x) = 1 - e^{-x}, G(x) = 1 - e^{-20x}$. Точное значение $P(t)$ вычисляют по формуле

$$P(t) = \frac{1}{221} \left(1 + \frac{17}{9} e^{-26t} - \frac{26}{9} e^{-17t} \right), \quad P_{ст} = \frac{1}{221}.$$

Значения оценки $\bar{P}(t)$, выборочной дисперсии $D(t)$ и $P(t)$, полученные при различных t , приведены в следующей таблице (было проведено по 10 000 реализаций при каждом значении t):

t	$5 \cdot 10^{-4}$	10^{-3}	10^{-2}	10^{-1}	$5 \cdot 10^{-1}$
$P(t)$	$2,482 \cdot 10^{-7}$	$9,858 \cdot 10^{-7}$	$8,677 \cdot 10^{-5}$	$2,772 \cdot 10^{-3}$	$4,522 \cdot 10^{-3}$
$\bar{P}(t)$	$2,140 \cdot 10^{-7}$	$8,477 \cdot 10^{-7}$	$8,361 \cdot 10^{-5}$	$2,799 \cdot 10^{-3}$	$4,585 \cdot 10^{-3}$
$D(t)$	$6,6 \cdot 10^{-11}$	$5,6 \cdot 10^{-10}$	$4,6 \cdot 10^{-7}$	$4,0 \cdot 10^{-5}$	$5,0 \cdot 10^{-5}$

Улучшение сходимости метода Монте-Карло при вычислении характеристик цепей Маркова [18]. Пусть $\{\xi(n), n = 0, 1, \dots\}$ — неоднородная цепь Маркова с конечным множеством состояний $N = \{0, 1, \dots, K\}$. Заданы следующие величины:

$$p_i^{(0)} = P \{ \xi(0) = i \},$$

$$p_{ij}^{(n)} = P \{ \xi(n) = j | \xi(n-1) = i \},$$

$$i, j \in N, n \geq 0.$$

Требуется вычислить функционал

$$c = \sum_{i=0}^K c_i P_i(T),$$

где $P_i(T) = P \{ \xi(T) = i \}, c > 0, c_i \geq 0$.

Пусть $D(T)$ — множество цепей Маркова $\{\xi^*(n), n \geq 0\}$, определяемых характеристиками $\{p_i^{*(0)}, p_{ij}^{*(n)}\}$ и удовлетворяющих условию: если для некоторой последовательности

$$i_0, i_1, \dots, i_T, i_j \in N, j=1, \dots, T,$$

$$p_{i_0}^{(0)} p_{i_0 i_1}^{(1)} \dots p_{i_{T-1} i_T}^{(T)} c_{i_T} < 0, \text{ то } p_{i_0}^{*(0)} \times$$

$$\times p_{i_0 i_1}^{*(1)} \dots p_{i_{T-1} i_T}^{*(T)} > 0.$$

Пусть построена траектория i_0, i_1, \dots, i_T цепи Маркова $\{\xi^*(n), n \geq 0\}$ в $[0, T]$ и

$$\bar{\theta} = p_{i_0}^{(0)} p_{i_0 i_1}^{(1)} \dots p_{i_{T-1} i_T}^{(T)} c_{i_T} / (p_{i_0}^{*(0)} \times$$

$$\times p_{i_0 i_1}^{*(1)} \dots p_{i_{T-1} i_T}^{*(T)}).$$

Тогда $M\bar{\theta} = c$, т. е. оценка $\bar{\theta}$ является несмещенной. Задача состоит в таком выборе характеристик $\{p_i^{*(0)}, p_{ij}^{*(n)}\}$, чтобы $D\bar{\theta}$ была минимальной. Пусть

$$\psi_i(t) = \sum_{j=0}^K P_{ij}^{[t, T]} c_j, \text{ где } P_{ij}^{[t, T]} =$$

$$= P \{ \xi(T) = j | \xi(t) = i \}.$$

Если

$$p_i^{*(0)} = \frac{1}{c} p_i^{(0)} \psi_i(0),$$

$$p_{ij}^{*(n)} = p_{ij}^{(n)} \psi_j(n) / \psi_i(n-1),$$

то

$$M\bar{\theta} = c, \quad D\bar{\theta} = 0.$$

Пусть $\bar{\psi}_i(t)$ — некоторое приближение для $\psi_i(t)$, причем $\bar{\psi}_i(T) = c_i, \bar{\psi}_i(t) > 0$ для любого $t = 0, 1, \dots, T-1$. Если в качестве $\{p_i^{*(0)}, p_{ij}^{*(n)}\}$ выбрать

$$p_i^{*(0)} = \frac{p_i^{(0)} \bar{\psi}_i(0)}{\sum_{k \in N} p_k^{(0)} \bar{\psi}_k(0)},$$

$$p_{ij}^{*(n)} = \frac{p_{ij}^{(n)} \bar{\psi}_j(n)}{\sum_{k \in N} p_{ik}^{(n)} \bar{\psi}_k(n)},$$

то $D\bar{\theta}$ будет тем меньше, чем более точно подобраны функции $\{\bar{\psi}_i(t)\}$. Таким образом, алгоритм построения статистической оценки для c состоит из двух этапов. На первом этапе строят приближения для функций $\{\psi_i(t)\}$ (один из возможных способов построения статистической оценки $\bar{\psi}_i(t)$ для $\psi_i(t)$ приведен в [18]), а на втором — используя цепь Маркова с переходными характеристиками $\{p_i^{*(0)}, p_{ij}^{*(n)}\}$, строят статистическую оценку для c .

Нахождение вероятности безотказной работы сложных систем [12, 16]. Опишем аналитико-статистический метод построения оценки вероятности $P(T)$ отказа в $(0, T)$ резервированной восстанавливаемой системы. Несмещенность этой оценки доказана в [16]. Соответствующие методы исследования надежности систем более сложной структуры изложены в [16, 17].

Система состоит из m независимых элементов и n ремонтных устройств ($1 \leq n < m$). В момент $t = 0$ система включается в рабочее состояние и все ее элементы исправны. Длительности безотказной работы элементов имеют функции распределения $F_i(x), i = 1, \dots, m$. Пусть $v(t)$ — число неисправных элементов в момент t . Если

в момент t отказал i -й элемент ($1 \leq i \leq m$) и $v(t-0) < n$, то немедленно начинается его восстановление с функцией распределения $G_i(x)$. Если же $v(t-0) \geq n$, то элемент становится в очередь. Первый момент отказа системы определим следующим образом:

$$\zeta = \inf \{t \geq 0: v(t) \geq r\},$$

где $n + 1 \leq r \leq m$.

Требуется найти вероятность отказа системы в промежутке $[0, T]$, т. е. нужно определить $P(T) = P(\zeta < T)$. Пусть функции $F_i(x)$, $1 \leq i \leq m$ являются абсолютно непрерывными с плотностями $f_i(x) < \infty$ для любого $x > 0$:

$$F_i(x) = \int_0^x f_i(u) du, \quad 1 \leq i \leq m, \quad x \geq 0.$$

Обозначим:

$$\alpha_i(u) = \frac{f_i(u)}{1 - F_i(u)}, \quad u > 0,$$

$$v_i(t) = \begin{cases} 0, & \text{если в момент } t \text{ } i\text{-й элемент исправен,} \\ -1, & \text{если элемент находится на восстановлении,} \\ k, & \text{если элемент находится } k\text{-м по счету в очереди на восстановление,} \\ k = 1, & m - n; \end{cases}$$

$$\gamma_i(t) = \begin{cases} 0; & \\ \sup \{x \in [0, t]: v_i(t-x) = 0\}, & \text{если } v_i(t) > 0, \\ \sup \{x \in [0, t]: v_i(t-x) = -1\}, & \text{если } v_i(t) = 0, \\ -1, & \end{cases}$$

$$i = 1, \dots, m, \quad t \geq 0,$$

— время пребывания i -го элемента в состоянии $v_i(t) \leq 0$ к моменту t . Величины $v_i(t)$ и $\gamma_i(t)$, $1 \leq i \leq m$, полностью определяют состояние системы в момент t , и дальнейшее ее поведение, т. е. процесс

$$\xi(t) = (v_1(t), \dots, v_m(t);$$

$$\gamma_1(t), \dots, \gamma_m(t)), \quad t \geq 0,$$

является марковским. Пусть z — произвольное натуральное число, $1 \leq$

$z \leq r$, и каждому состоянию системы $(v_1, \dots, v_m; \gamma_1, \dots, \gamma_m)$ соответствует число неисправных в этом состоянии элементов

$$|v| = \sum_{i: v_i \neq 0} 1.$$

Пусть

$$A_i = \{(v_1, \dots, v_m; \gamma_1, \dots, \gamma_m):$$

$$|v| \geq r - z + i\},$$

$$\tau_i = \inf \{t \geq 0: \xi(t) \in A_i\},$$

$$1 \leq i \leq z,$$

— первый момент времени, когда в системе неисправно $r - z + i$ элементов, $\zeta = \tau_z$.

Суть сформулированного ниже алгоритма состоит в следующем.

1. Специальным образом моделируя систему, находят интенсивность

$$\lambda_1(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P\{\tau_1 \in (t, t + \Delta t)\},$$

$$t \in (0, T).$$

2. Согласно распределению

$$\frac{1}{J_1} \left[1 - \exp \left\{ - \int_0^x \lambda_1(u) du \right\} \right],$$

$$J_1 = 1 - \exp \left\{ - \int_0^T \lambda_1(u) du \right\},$$

находят первый момент τ_1 попадания системы в множество A_1 .

3. Находят состояние системы в момент $\tau_1 + 0$.

4. Специальным образом моделируя систему, находят интенсивность

$$\lambda_2(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P\{\tau_2 \in (t, t + \Delta t)\},$$

$$t \in (\tau_1, T).$$

5. Согласно распределению

$$\frac{1}{J_2} \left[1 - \exp \left\{ - \int_{\tau_1}^x \lambda_2(u) du \right\} \right],$$

$$J_2 = 1 - \exp \left\{ - \int_{\tau_1}^T \lambda_2(u) du \right\},$$

находят первый момент τ_2 попадания системы в множество A_2 .

6. Находят состояние системы в момент $\tau_2 + 0$.

Далее описанная процедура повторяется до того момента, когда будут найдены интенсивность $\lambda_z(t)$ и J_z . В качестве оценки в одной реализации выбирают

$$\bar{P}_1(T) = J_1 \quad J_z.$$

Более подробно алгоритм построения оценки $\bar{P}_1(T)$ выглядит следующим образом.

1. Методом статистического моделирования строят траекторию процесса $\xi(t)$ до первого момента $t^{(1)}$, в который $|v(t^{(1)} + 0)| = r - z$. Если $t^{(1)} \geq T$, то выборочная траектория $z_1(t)$, $t \in [0, T]$, построена и $\bar{P}_1(T) = 0$. Пусть $t^{(1)} < T$.

2. Находят минимальное время $\omega^{(1)}$ до окончания восстановления одного из неисправных в момент $t^{(1)} + 0$ элементов и номер i_1 этого элемента. Тогда

$$\lambda_1(t) = \begin{cases} 0, & \text{если } t < t^{(1)} \\ \sum_{i: v_i(t^{(1)}+0)=0} \alpha_i (\gamma_i(t^{(1)} + \\ + t - t^{(1)}), & \text{если } t \in [t^{(1)}, \\ t^{(1)} + \omega^{(1)}]. \end{cases}$$

3. Предполагая, что в интервале $(t^{(1)}, t^{(1)} + \omega^{(1)})$ не произойдет отказа ни одного из исправных элементов и в момент $t^{(1)} + \omega^{(1)}$ окончится восстановление i -го элемента, строят траекторию $z_1(t)$ до первого момента $t^{(2)} > t^{(1)} + \omega^{(1)}$, такого, что $|v(t^{(2)} + 0)| = r - z$. Далее находят $\omega^{(2)}$ и i_2 — соответственно минимальное время до окончания восстановления одного из неисправных в момент $t^{(2)} + 0$ элементов и номер этого элемента. При этом

$$\lambda_1(t) = \begin{cases} 0, & \text{если } t \in [t^{(1)} + \omega^{(1)}, t^{(2)}), \\ \sum_{i: v_i(t^{(2)}+0)=0} \alpha_i (\gamma_i(t^{(2)} + \\ + t - t^{(2)}), & \text{если } t \in [t^{(2)}, \\ t^{(2)} + \omega^{(2)}]. \end{cases}$$

Затем находят $t^{(3)}$, $\omega^{(3)}$, i_3 и т. д.

4. Пусть выборочная траектория $z_1(t)$, $t \in [0, T]$, построена и найдены следующие величины:

K — число интервалов времени из $[0, T]$, в течение которых неисправно ровно $r - z$ элементов (предполагается, что $K > 0$);

$t^{(k)}$, $k = 1, \dots, K$, — момент начала k -го интервала;

$\omega^{(k)}$, $k = 1, \dots, K$, — длительность k -го интервала ($t^{(K)} + \omega^{(K)} \leq T$);

$$\xi^{(k)} = \xi(t^{(k)} + 0) = \\ = (v_1^{(k)}, \dots, v_m^{(k)}; \gamma_1^{(k)}, \dots, \gamma_m^{(k)}),$$

$k = 1, \dots, K$.

5. Вычисляют

$$J_1 = 1 - \\ - \exp \left[- \sum_{k=1}^K \sum_{i: v_i^{(k)}=0} \int_{t^{(k)}}^{t^{(k)} + \omega^{(k)}} \alpha_i \times \right. \\ \left. \times (\gamma_i^{(k)} + u - t^{(k)}) du \right] = \\ = 1 - \prod_{k=1}^K \prod_{i: v_i^{(k)}=0} \frac{1 - F_i(\gamma_i^{(k)} + \omega^{(k)})}{1 - F_i(\gamma_i^{(k)})}.$$

6. Реализуют случайную величину x_1 , распределенную в промежутке $[0, T]$ с функцией распределения

$$H(y) = \frac{1}{J_1} \times \\ \times \left\{ 1 - \prod_{k=1}^K \prod_{i: v_i^{(k)}=0} \frac{1 - F_i(\gamma_i^{(k)} + \min(y, t^{(k)} + \omega^{(k)}) - t^{(k)})}{1 - F_i(\gamma_i^{(k)})} \right\}.$$

Функцию $H(y)$ можно представить в виде

$$H(y) = \sum_{k=1}^K p_k H_k(y), \quad y \geq 0,$$

где

$$p_1 = \frac{1}{J_1} u_1,$$

$$p_k = \frac{1}{J_1} u_k \prod_{i=1}^{k-1} (1 - u_i),$$

$$k = 2, \quad K,$$

$$u_k = 1 - \prod_{i: v_i^{(k)}=0} [1 - F_i(\gamma_i^{(k)} +$$

$$+ \omega^{(k)})] / [1 - F_i(\gamma_i^{(k)})],$$

$$H_k(y) = \frac{1}{u_k} \left\{ 1 - \prod_{i: v_i^{(k)}=0} [1 -$$

$$- F_i(y)] / [1 - F_i(\gamma_i^{(k)})] \right\},$$

$$t^{(k)} \leq y < t^{(k)} + \omega^{(k)},$$

$$H_k(y) = 0, \quad y < t^{(k)}, \quad H_k(y) = 1,$$

$$y \geq t^{(k)} + \omega^{(k)}.$$

Далее реализуют случайную величину σ , равную k с вероятностью p_k . Если $\sigma = k$, то κ_1 распределена в $(t^{(k)}, t^{(k)} + \omega^{(k)})$ с функцией распределения $H_k(y)$. Пусть θ_i — случайная величина с функцией распределения

$$D_i(x) = [F_i(x - t^{(k)} + \gamma_i^{(k)}) - F_i(\gamma_i^{(k)})] / [1 - F_i(\gamma_i^{(k)})], \quad i: v_i^{(k)} = 0.$$

Тогда

$$H_k(y) = \mathbf{P} \left\{ \min_{i: v_i^{(k)}=0} \theta_i < y \mid \min_{i: v_i^{(k)}=0} \theta_i \in (t^{(k)}, t^{(k)} + \omega^{(k)}) \right\}.$$

Номер j_1 , на котором достигается $\min_{i: v_i^{(k)}=0} \theta_i$ при указанном условии, и значение этого минимума определяют методом статистического моделирования. Пусть

$$v_i = \begin{cases} [F_i(\gamma_i^{(k)} + \omega^{(k)}) - F_i(\gamma_i^{(k)})] / [1 - F_i(\gamma_i^{(k)})], \\ \text{если } v_i^{(k)} = 0, \\ 0 \text{ в противном случае,} \end{cases}$$

$$u_k = 1 - \prod_{i=1}^m (1 - v_i), \quad q_1 = \frac{v_1}{u_k},$$

$$q_i = v_i \prod_{s=1}^{i-1} (1 - v_s) / u_k, \quad i = 1, \dots, m.$$

Далее реализуют случайную величину β , равную i с вероятностью q_i . Пусть $\beta = i$. Это означает, что $\theta_s > t^{(k)} + \omega^{(k)}$, $s = 1, \dots, i - 1$, $s: v_s^{(k)} = 0$, $\theta_i \in [t^{(k)}, t^{(k)} + \omega^{(k)})$ и имеет функцию распределения

$$H_{ki}(y) = \frac{F_i(\gamma_i^{(k)} + y - t^{(k)}) - F_i(\gamma_i^{(k)})}{F_i(\gamma_i^{(k)} + \omega^{(k)}) - F_i(\gamma_i^{(k)})},$$

$$y \in [t^{(k)}, t^{(k)} + \omega^{(k)}),$$

$$H_{ki}(y) = 0, \quad y < t^{(k)}, \quad H_{ki}(y) = 1,$$

$$y \geq t^{(k)} + \omega^{(k)}.$$

В качестве κ_1 выбирают

$$\kappa_1 = \min \left(\theta_i^*, \min_{\substack{s \geq i+1 \\ s: v_s^{(k)}=0}} \theta_s \right),$$

где θ_i^* имеет функцию распределения $H_{ki}(y)$, а θ_s , $s = i + 1, \dots, m$, $s: v_s^{(k)} = 0$, — функцию распределения $D_s(x)$. Номер j_1 , на котором достигается данный минимум, указывает номер элемента, отказ которого произойдет в момент κ_1 .

7. Пусть в момент $\kappa_1 \in [t^{(k)}, t^{(k)} + \omega^{(k)})$ отказал j -й элемент. Тогда вы-

числяют состояние системы в момент $x_1 + 0$:

$$\xi(x_1 + 0) = (v'; \gamma') = (v'_1, v'_m; \gamma'_1, \gamma'_m),$$

где

$$v'_s = \begin{cases} -1, & \text{если } s = j_1 \\ \text{и } r - z + 1 \leq n, \\ r - z + 1 - n, & \text{если } s = j_1 \\ \text{и } r - z + 1 > n, \\ v_s & \text{в противном случае,} \end{cases}$$

$$\gamma'_s = \begin{cases} 0, & \text{если } s = j_1 \text{ или } v_s^{(k)} > 0, \\ \gamma_s^{(k)} + x_1 - t^{(k)} & \\ \text{в противном случае} \end{cases}$$

$s = 1, \dots, m.$

8. В интервале (x_1, T) методом статистического моделирования строят траекторию процесса $\xi(t)$, $t \geq x_1$, с «запретом» попадать в множество состояний $A_2 = \{(v; \gamma) \mid |v| \geq r - z + 2\}$ и вычисляют интенсивность $\lambda_2(t)$, $t \in (x_1, T]$, попадания в множество A_2 при условии $\xi(x_1 + 0) = (v'; \gamma')$, $|v'| = r - z + 1$. Для этого вычисляют новые значения K , $\{t^{(k)}, \omega^{(k)}, \xi^{(k)}, k = 1, \dots, K\}$ и находят J_2 . Затем описанным выше способом реализуют случайную величину x_2 , и алгоритм полностью повторяется. В качестве оценки $\bar{P}_1(T)$ выбирают

$$\bar{P}_1(T) = J_1 J_2 \quad J_z.$$

Сформулированный алгоритм зависит от параметра z . С ростом z выборочная дисперсия убывает, хотя при этом возрастает время одной реализации. Пусть $T(i)$ и $D(i)$, $i = 1, \dots, r$, — соответственно оценка средней продолжительности одной реализации и выборочная дисперсия, полученная при $z = i$ за достаточно большое число реализаций. В качестве оптимального z_0 естественно выбрать то значение z , при котором произведение $T(z) D(z)$ принимает минимальное значение (см. [6]), т. е. $z_0 = \arg \min_{i=1, \dots, r} T(i) D(i)$.

Примеры. Пусть система состоит из

одинаковых элементов, у которых длительности исправной работы имеют распределение Вейбулла

$$F(x) = P\{\xi < x\} = 1 - \exp(-x^2),$$

$$x \geq 0,$$

а длительность восстановления η имеет равномерное распределение в $[0, \varepsilon]$.

Пример 1. Пусть $m = 2$, $n = 1$, $T = 20$, $\varepsilon = 5 \cdot 10^{-5}$, $z = 1$. Для построения оценки с достоверностью 0,95 и абсолютной погрешностью 10^{-4} потребовалось всего 17 реализаций. При этом

$$\bar{P}(T) = 0,00122, \quad D = 4,4 \cdot 10^{-8},$$

где D — выборочная дисперсия, построенная по 17 реализациям.

Для данной системы вероятность отказа может быть найдена и с помощью асимптотических формул:

$$P(T) \approx P_{ас}(T) = 1 - \exp\left\{-\frac{2M\eta}{(M\xi)^2} T\right\} \approx 0,00127.$$

Таким образом, приближенные значения для $P(T)$, найденные двумя совершенно различными способами, довольно близки.

Пример 2. Пусть в отличие от предыдущего примера $m = 3$, $r = 2$, а все остальные начальные данные остаются без изменения. Для построения оценки с достоверностью 0,95 и абсолютной погрешностью $2 \cdot 10^{-9}$ потребовалось 6932 реализации, причем

$$\bar{P}(T) = 7,1 \cdot 10^{-8}, \quad D = 7,2 \cdot 10^{-15}.$$

Согласно асимптотической формуле приближенное значение имеет вид

$$P(T) \approx P_{ас}(T) = 1 - \exp\left\{-\frac{3M\eta^2}{(M\xi)^3} T\right\} \approx 7,2 \cdot 10^{-8}.$$

Практическое совпадение $\bar{P}(T)$ и $P_{ас}(T)$ получилось благодаря тому, что T достаточно велико по сравнению с $M\xi$, и поэтому до момента T система «успевает войти» в стационарный режим. Если $M\xi > T$, то использование приведенных выше асимптотических формул может существенно исказить истинное значение $P(T)$.

Пример 3. Пусть $m = 5, n = 2, r = 4, \varepsilon = 0,014, T = 100, z = 1$. Для построения оценки с достоверностью 0,95 и абсолютной точностью $2 \cdot 10^{-4}$ потребовалось 2270 реализаций. При этом

$$\bar{P}(T) = 0,00147, D = 2,3 \cdot 10^{-5}.$$

Пример 4. Пусть $m = 4, n = 4, r = 4, \varepsilon = 0,01, T = 0,2$. При каждом $z = 1, 2, 3, 4$ было сделано по 500 реализаций. Соответствующие оценки имеют вид

$$\begin{aligned} z = 1 \quad P(T) = 0, \quad D = 0, \\ z = 2 \quad P(T) = 0, \quad D = 0, \\ z = 3 \quad P(T) = 3,14 \cdot 10^{-10}, \\ D = 2,5 \cdot 10^{-18}, \\ z = 4 \quad P(T) = 3,88 \cdot 10^{-10}, \\ D = 4,2 \cdot 10^{-19}. \end{aligned}$$

Следовательно, расчет надежности данной системы следует вести при $z = 3$ или $z = 4$ (по-видимому, случай $z = 4$ более предпочтителен).

10. КУСОЧНО-ЛИНЕЙНЫЙ АГРЕГАТ

Кусочно-линейный агрегат (КЛА) — это объект, в любой момент t характеризующийся (внутренним) состоянием $z(t)$, способный принимать входные сигналы x и посылать выходные сигналы y . Возможные состояния агрегата, а также входные и выходные сигналы имеют вид вектора (i, x) , где i называют дискретной (основной) компонентой из конечного или счетного множества I , а x — вектором (дополнительных) вещественных переменных, размерность $|i|$ которого зависит от состояния i .

Функционирование агрегата [3, 10].
1. Спонтанное изменение состояния.

Пусть $z(t) = (v(t), \xi(t))$ — состояние агрегата в момент $t, v(t) = k, \xi(t) = (\xi_1(t), \dots, \xi_{|v(t)|}(t)) = (z_1, \dots, z_{|k|})$ и за время dt на вход агрегата не поступил сигнал. Тогда в интервале $(t, t + dt)$ может произойти так называемый спонтанный переход $v(t)$ в состояние l с вероятностью

$$P(v(t + dt) = l/v(t) = k; \xi_j(t) > 0, 1 \leq j \leq |k|) = \lambda_{kl} dt,$$

где $p_{kl} = \frac{\lambda_{kl}}{\lambda_k}, \lambda_k = \sum_{l \neq k} \lambda_{kl}$ — вероятность перехода $v(t)$ в состояние l при условии, что произошел выход $v(t)$ из состояния k .

После изменения $v(t)$ определяют случайный вектор $\gamma = (\gamma_1, \gamma_2, \dots)$ с размерностью, зависящей от k, l , и многомерной функцией распределения $H_{kl}^{(0)}(x) = H_{kl}^{(0)}(x_1, x_2, \dots)$. Далее формируют вектор

$$(\xi\gamma) = (z_1, \dots, z_{|k|}, \gamma_1, \gamma_2, \dots)$$

и

$$\xi(t + 0) = (\xi_1(t + 0),$$

$$\xi_2(t + 0), \dots) = (\xi\eta) A_{kl}^{(0)},$$

где $A_{kl}^{(0)}$ — фиксированная матрица.

Вероятность того, что спонтанное изменение состояния агрегата не происходит за время dt , равна $1 - \lambda_k dt$. В этом случае

$$v(t + dt) = v(t);$$

$$\xi_j(t + dt) = z_j - \alpha_{kj} dt, \quad j \geq 1,$$

где α_{kj} — скорость изменения j -й компоненты вектора дополнительных переменных при $v(t) = k$.

2. Обращение дополнительной переменной в нуль.

Если в некоторый момент времени t $\xi_j(t)$ обратилась в нуль ($\xi_j(t - 0) = z_j = 0$), то $v(t + 0)$ определяется по вероятностному закону

$$P(v(t + 0) = l/v(t - 0) = k,$$

$$\xi_j(t - 0) = 0;$$

$$\exists \varepsilon > 0, \xi_j(u) \neq 0,$$

$$t - \varepsilon < u < t) = p_{kl}^{(j)},$$

где $\varepsilon > 0$ — сколь угодно малая величина. Затем реализуют случайный вектор $\gamma = (\gamma_1, \gamma_2, \dots)$ с размерностью, зависящей от k, l, j , и многомерной функцией распределения $H_{kl}^{(j)}(x) = H_{kl}^{(j)}(x_1, x_2, \dots)$, после чего вектор дополнительных переменных в момент

времени $t + 0$ определяют из соотношения

$$\xi(t + 0) = (z_1, \quad z_{1k}),$$

$$\gamma_1, \gamma_2, \dots) A_{kl}^{(j)},$$

где $A_{kl}^{(j)}$, $j \geq 1$ — фиксированные матрицы.

Предполагается, что в любой момент t может обратиться в нуль лишь одна дополнительная компонента.

3. Поступление входного сигнала.

Если в момент t на вход агрегата поступил сигнал $x = (i, x_1, \dots, x_{|i|})$, причем $z(t - 0) = (k, z_1, \dots, z_{1k})$, то $v(t + 0)$ определяют согласно закону распределения

$$P(v(t + 0) = l/v(t - 0) = i) = p_{kil}.$$

Далее реализуют случайный вектор $\gamma = (\gamma_1, \gamma_2, \dots)$ с размерностью, зависящей от k, i, l , и многомерной функцией распределения $H_{kil}(x) = H_{kil}(x_1, x_2, \dots)$ и затем определяют вектор дополнительных переменных в момент $t + 0$

$$\xi(t + 0) = (z_1, \quad z_{1k},$$

$$x_1, \quad x_{|i|}, \gamma_1, \gamma_2, \dots) A_{kil},$$

где A_{kil} — фиксированные матрицы.

Выходные сигналы кусочно-линейного агрегата определяют следующим образом. Если произошел переход $v(t)$ из k в l , то условие посылки выходного сигнала (j, y) определяется соотношением $(k, l) \in Z_j$, где Z_j — фиксированное подмножество множества Z , причем $Z_i \cap Z_j = \emptyset$, $i \neq j$. Множество Z называют множеством, определяющим условие посылки выходных сигналов.

Схема сопряжения. Для описания сложных систем кусочно-линейными агрегатами вводится схема сопряжения, позволяющая объединить их в единую модель системы посредством установления соответствия между входами и выходами таких агрегатов. Полученная модель является автономным кусочно-линейным агрегатом, функционирование которого описывается кусочно-линейным марковским процессом

$$z(t) = (z_1(t), \quad z_N(t)),$$

где $z_n(t)$ — условный кусочно-линейный марковский процесс, описывающий функционирование n -го агрегата.

Принципы построения моделирующих алгоритмов. Моделирование случайного процесса $z(t)$ можно представить как построение значений $z(t)$ в некоторые моменты t_n , $n = 1, 2, \dots$, его развития. В зависимости от способа выбора моментов t_n различают два основных принципа построения моделирующих алгоритмов [2].

Принцип Δt . Он предполагает определение значений $z(t)$ через заранее фиксированные промежутки времени Δt в моменты $t_0 + n \Delta t$, $n = 1, 2, \dots$. Принцип Δt является универсальным, поскольку применим для моделирования широкого класса сложных систем дискретного и непрерывного характера. Однако он весьма неэффективен с вычислительной точки зрения.

Принцип «узловых» моментов. Данный принцип позволяет моделировать значения $z(t)$ только в так называемые «узловые» моменты времени t_n^* , в которых процесс принимает особые состояния. Такие состояния характерны тем, что именно по ним оцениваются свойства систем, описываемых $z(t)$. В отличие от первого принципа, моменты t_n^* определяются в процессе развития траектории $z(t)$ по некоторому алгоритму.

11. МОДЕЛИРОВАНИЕ КУСОЧНО-ЛИНЕЙНЫХ АГРЕГАТОВ

При моделировании кусочно-линейных агрегатов используется принцип узловых моментов [10]. Узловыми являются моменты t_n поступления входного сигнала, обращения непрерывной компоненты в нуль, спонтанное изменение состояния.

Процесс моделирования сводится к определению значений $z(t_n - 0)$, $z(t_n + 0)$ и очередного узлового момента t_{n+1} . Пусть известны значения t_n и $z(t_n + 0) = (k, z_1, \dots, z_{1k})$. Тогда t_{n+1} определяется по следующему алгоритму.

1. Находят минимальное время γ_k , через которое одна из дополнительных переменных обратится в нуль

$$\gamma_k = \min_{j \geq 1: \alpha_{kj} > 0} \{z_j / \alpha_{kj}\},$$

и полагают $t_{n+1} = t_n + \gamma_k$.

2. Моделируют случайную величину v , принимающую значение 0 с вероятностью $p = e^{-\lambda_k v_k}$ и 1 с вероятностью $q = 1 - p$. Если v приняла значение 1, то реализуют случайную величину τ с плотностью

$$p(x) = \frac{\lambda_k e^{-\lambda_k x}}{k!}, \quad 0 \leq x < \gamma_k,$$

и полагают $t_{n+1} = t_n + \tau$.

3. Если в момент $t \in (t_n, t_{n+1})$ поступил входной сигнал, то $t_{n+1} = t$.

Переход от $z(t_n + 0)$ к $z(t_{n+1} - 0)$ осуществляют на основании формул

$$v(t_{n+1} - 0) = v(t_n + 0),$$

$$\xi_j(t_{n+1} - 0) = \xi_j(t_n + 0) -$$

$$- \alpha_{vj}(t_{n+1} - t_n), \quad j \geq 1,$$

а переход от $z(t_n - 0)$ к $z(t_n + 0)$ реализуют согласно определению кусочно-линейного агрегата.

Пример. Имеется рабочий элемент, который может находиться в двух состояниях — отказовом (1) и исправном (0). Время пребывания в i -м состоянии — случайная величина с функцией распределения $H_i(t)$. Функционирование элемента описывается кусочно-линейным агрегатом с состоянием в момент t вида

$$z(t) = (v(t), \xi_1(t)),$$

где

$$v(t) = \begin{cases} 0, & \text{если в момент } t \\ & \text{элемент исправен,} \\ 1 & \text{в противном случае,} \end{cases}$$

а $\xi_1(t)$ — вектор дополнительных переменных, состоящий из одной компоненты: $\xi_1(t)$ — время от момента t до смены значения $v(t)$, убывающее с единичной скоростью ($\alpha_{k1} = 1, k = 0, 1$).

При обращении $\xi_1(t)$ в нуль $v(t + 0)$ принимает значение 0 либо 1 согласно распределению

$$p_{00} = 0, \quad p_{01} = 1,$$

если $v(t - 0) = 0$, и те же значения согласно распределению

$$p_{10} = 1, \quad p_{11} = 0,$$

если $v(t - 0) = 1$. Затем реализуют случайную величину с функцией распределения $H_{v(t+0)}(x)$ и новое значение $\xi_1(t + 0)$ полагают равным γ .

Для установления различий между моментами отказов и восстановлений элемента вводится два типа выходных сигналов агрегата x_0 и x_1 , которым соответствуют два непересекающихся множества

$$Z_0 = \{(1, 0)\}, \quad Z_1 = \{(0, 1)\}.$$

Сигнал x_0 является следствием смены в момент t значения $v(t)$ ($v(t - 0) = 1, v(t + 0) = 0$) и свидетельствует о восстановлении элемента в момент t , а x_1 образуется в момент t отказа элемента ($v(t - 0) = 0, v(t + 0) = 1$).

12. ОСНОВНЫЕ СИСТЕМЫ МОДЕЛИРОВАНИЯ

На базе моделирующих алгоритмов созданы и разрабатываются универсальные программные средства, позволяющие автоматизировать процесс построения имитационных моделей. В настоящее время широкое распространение получили такие системы моделирования, как GPSS [27], СИМСКРИПТ [19], CSL [28], SOL [29], СИМУЛА [5], НЕДИС [23].

Используемый в них принцип проводки событий [2] приводит к значительным трудностям, возникающим при любом изменении структуры исследуемого объекта.

Этого недостатка лишены программные средства, в основе которых лежит формализованная модель кусочно-линейных агрегатов. К таким моделирующим системам относятся: автоматизированная имитационная модель (АИМ) [1], агрегативная имитационная система (АИС) [8], пакет программ для агрегатного моделирования систем АМОС [15]. Первые две работают под управлением ОС ЕС ЭВМ, а последняя реализована на ЭВМ БЭСМ-6.

Пакет программ АМОС. Формализованная модель, которая была положена в основу разработки такого про-

граммного средства, состоит из двух компонент: одномерного кусочно-линейного агрегата и двухуровневой схемы сопряжения.

Одномерный кусочно-линейный агрегат (в дальнейшем просто агрегат) является частным случаем кусочно-линейного агрегата общего вида и обладает следующими особенностями.

1. Внутреннее состояние агрегата имеет вид $z(t) = v(t)$ либо $z(t) = (v(t), \xi(t))$, где $v(t)$ принимает конечное множество целочисленных положительных значений, $\xi(t)$ — любое положительное число.

2. Выходной сигнал не зависит от непрерывной компоненты $\xi(t)$ и имеет вид

$$x(t) = \mu(t),$$

где $\mu(t)$ является признаком сигнала и принимает конечное число возможных значений.

3. Спонтанное изменение состояний агрегата происходит по следующей схеме.

За время dt состояние агрегата $z = k$ или $z = (k, \xi)$ может измениться скачком, а именно: если $v(t-0) = k$, то с вероятностью $\lambda_{kl} dt$ $v(t+0) = l$. Если такое событие произошло, то в случае $z = (k, \xi)$ реализуют случайную величину η с функцией распределения $H_l^{(0)}(x)$, и новое значение ξ будет $a_{kl}^{(0)}\xi + \eta$; если же $z = k$, то ξ не определяют.

Описанный агрегат совместно с двухуровневой схемой сопряжения позволяет описывать довольно сложные модели исследуемых систем. В частности, таким образом можно описать системы массового обслуживания.

Двухуровневая схема сопряжения. Схема сопряжения 1-го уровня устанавливает пути передачи сигналов между отдельными агрегатами и тем самым позволяет создать базовые модели типичных подсистем (очередь, обслуживающий прибор, рабочий элемент и т. п.).

Схема сопряжения 2-го уровня необходима для установления соответствия между входами и выходами различных подсистем. Следовательно, имеется возможность использовать для исследования конкретных схем гото-

вые программные модули, реализующие функционирование базовых моделей.

На обоих уровнях иерархии схема сопряжения обладает следующими свойствами:

1. Выходной сигнал x , выдаваемый k -м агрегатом при смене состояния его дискретной компоненты ($v(t-0) = i$, $v(t+0) = j$), имеет вид

$$x = x(k, L, i, j)$$

на 1-м уровне и

$$x = x(N, k, i, j, M, L_m)$$

на 2-м уровне. Здесь L — множество номеров агрегатов, на которые подается сигнал x ; N — номер подсистемы, содержащей источник сигнала k -й агрегат; M — множество номеров подсистем, на которые одновременно подается сигнал x . Подсистема с номером $m \in M$ принимает сигнал через приемники сигналов — агрегаты с номерами из множества L_m .

Введение таких сигналов позволяет передавать информацию от одного агрегата (подсистемы) одновременно к нескольким агрегатам (подсистемам).

2. Имеется возможность одновременной посылки различных сигналов от нескольких агрегатов на один и тот же агрегат. Очередность воздействия таких сигналов устанавливается в порядке возрастания номеров агрегатов — источников сигналов.

3. Реакция агрегата с номером l (в дальнейшем употребляется просто агрегат l или l -й агрегат), $l \in L$, на входной сигнал x в момент t задается распределением вероятностей

$$lP_{v(t-0), x, v(t+0)}.$$

Пример 1. Имеется система с параллельным соединением n элементов, структура которой изображена на рис. 2. Время безотказной работы i -го ($i = 1, 2, \dots, n$) элемента случайная величина ξ_{i0} с функцией распределения $H_{i0}(x)$, а время его восстановления — случайная величина ξ_{i1} с функцией распределения $H_{i1}(x)$. Восстановление элемента начинается в момент его отказа. Отказ системы наступает в момент одновременного выхода из строя k элементов ($k \leq n$). Требуется определить вероятность $P(T)$ безотказной работы системы за время T .

Решение. Функционирование системы описывается i -м агрегатом. Состояние i -го агрегата, $i = 1, 2, \dots, n$, в момент t имеет вид

$$z_i(t) = (v_i(t), \xi_i(t)),$$

где

$$v_i(t) = \begin{cases} 0, & \text{если элемент исправен} \\ & \text{в момент } t, \\ 1, & \text{если он находится} \\ & \text{на восстановлении,} \end{cases}$$

а $\xi_i(t)$ соответствует функция распределения $H_{i v_i(t)}(x)$, $\alpha_{i v_i(t)} = 1$. В начальный момент $t_0 = 0$ с вероятностью 1 $v_i(t_0) = 0$.

Смена состояния $v_i(t)$ при обращении непрерывной компоненты в нуль задается законом распределения

$$i p_{00} = i p_{11} = 0, \quad i p_{01} = i p_{10} = 1.$$

$n + 1$ -й агрегат является вероятностным автоматом с состоянием $z_{n+1}(t) = (v_{n+1}(t), \xi_{n+1}(t))$, $v_{n+1}(t) = 0, 1, \dots, n$.

Он моделирует счетчик числа отказавших элементов в момент t и изменяет свое состояние только под воздействием входных сигналов: $v_{n+1}(t_0 + 0) = 0$ с вероятностью 1. Графически схема сопряжения 1-го уровня изображена на рис. 3.

i -й агрегат, $i = 1, 2, \dots, n$, выдает два вида сигналов x_1 и x_2 , которым соответствуют два множества $Z_1 = \{(0, 1)\}$ и $Z_2 = \{(1, 0)\}$. $n + 1$ -й агрегат может только принимать входные сигналы. Его реакция на x_1 и x_2 задается законом распределения

$$n+1 p_{i, x_1, i+1} = 1, \quad i = 0, 1, \dots, k-1, \\ n+1 p_{i, x_2, i-1} = 1, \quad i = 1, 2, \dots, k-1.$$

Остальные вероятности равны нулю.

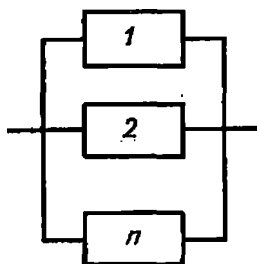


Рис. 2. Система с параллельным соединением элементов

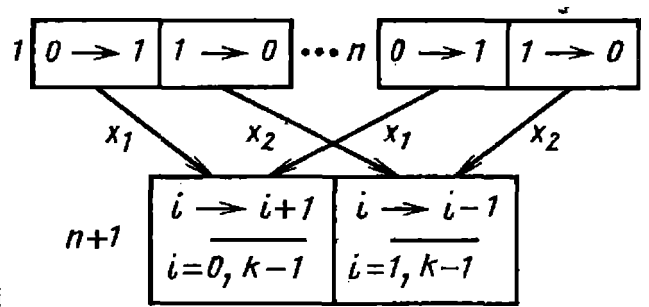


Рис. 3. Графическая схема сопряжения первого уровня

Реализация процесса функционирования системы заканчивается в момент $t^* = \min \{T, \inf (t: v_{n+1}(t+0) = k)\}$.

Если обозначить через

$$\mu_i = \begin{cases} 1, & \text{если } t^* = T, \\ 0 & \text{в противном случае,} \end{cases}$$

случайную величину, определяемую в i -й реализации процесса, то в качестве оценки вероятности безотказной

работы системы на $[0, T]$ можно взять оценку $\bar{P}(T)$ для $M\mu_i$:

$$\bar{P}(T) = \frac{1}{N} \sum_{i=1}^N \mu_i,$$

где число реализаций N выбирается согласно требуемой точности и достоверности оценки.

Пример 2. Если элементы системы, описанной в примере 1, соединить последовательно, то получается система со структурой, изображенной на рис. 4. Система работоспособна в момент t , если все ее элементы исправны в этот момент, и неработоспособна, если хотя бы один элемент отказал. Элементы отказывают и восстанавливаются независимо друг от друга. Требуется определить среднее время T_c безотказной работы системы на $[0, T]$.

Решение. Агрегатная модель настоящей системы отличается от модели предыдущей системы тем, что $k = n$, и $n + 1$ -м агрегатом, который в данном случае является счетчиком времени, а именно: его состояние описывают случайным процессом

$$z_{n+1}(t) = (v_{n+1}(t), \xi_{n+1}(t)),$$

где $v_{n+1}(t) = 0, 1, \dots, n$ изменяется только под воздействием входных сигналов.

Если обозначить через γ_i суммарное время пребывания системы в исправном состоянии на $[0, T]$ в i -й реализации, а через $\zeta_i(t)$ — суммарное время пребывания системы в исправном состоянии в i -й реализации на $[0, t]$, $t \leq T$, то

$$\xi_{n+1}(t) = \sum_{i=1}^m \gamma_i + \zeta_{m+1}(t).$$

$\alpha_{n+1,0} = -1$, $\alpha_{n+1,1} = 0$, m — число реализаций. В начальный момент $t_0 = 0$ $v_{n+1}(t_0 + 0) = 0$ с вероятностью 1, а $\xi_{n+1}(t_0 + 0) = 0$ только в 1-й реализации. В качестве оценки для искомой характеристики может служить оценка \bar{T}_c для $M\gamma_i$ при достаточно больших N , а именно

$$\bar{T}_c = \frac{1}{N} \sum_{i=1}^N \gamma_i,$$

где N выбирается согласно требуемой точности и достоверности оценки.

Пример 3. Имеется система со смешанным соединением элементов, структура которой изображена на рис. 5. Последовательно соединенные элементы образуют m подсистем по n_j элементов в каждой из них, $j = 1, 2, \dots, m$, j -я подсистема функционирует аналогично системе, описанной в примере 2. Система отказывает при одновременном отказе k подсистем, $k \leq m$. Требуется определить вероятность безотказной работы системы на $[0, T]$.

Решение. Агрегатная модель j -й подсистемы аналогична модели системы примера 2. В качестве выхода j -й подсистемы берут $n_j + 1$ -й агрегат, который полностью определяется состоянием подсистемы. Следовательно, вместо j -й подсистемы на втором уровне схемы сопряжения рассматривают $n_j + 1$ -й агрегат. Таким обра-

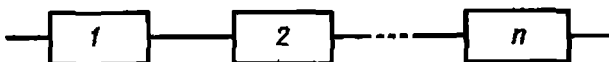


Рис. 4. Система с последовательным соединением элементов

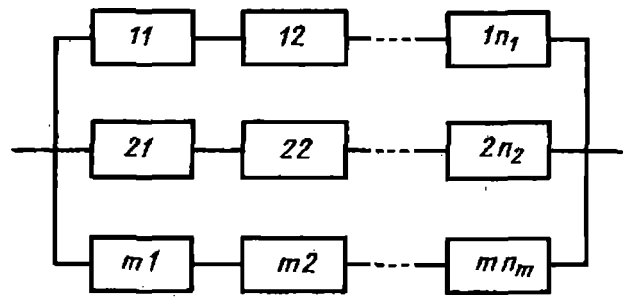


Рис. 5. Система со смешанным соединением элементов

зом, $n_j + 1$ -й агрегат соответствует j -му агрегату, описанному в примере 1, а схема сопряжения второго уровня — схеме сопряжения первого уровня примера 1. Значит, пример 3 свелся к примерам 1 и 2.

Двухуровневая схема сопряжения элементов системы приведена на рис. 6.

Пример 4. Имеется система со следующей структурой (рис. 7). Параллельно соединенные элементы образуют m подсистем по n_j элементов в каждой, $j = 1, 2, \dots, m$, j -я подсистема функционирует аналогично системе, описанной в примере 1. Система отказывает при отказе хотя бы одной подсистемы. Требуется определить ту же характеристику, что и в примере 2.

Решение. Если в качестве j -го элемента системы примера 2 взять выход j -й подсистемы ($n_j + 1$ -й агрегат), описанной в примере 1, то моделирование системы, описанной в настоящем примере, сведется к моделированию системы примера 2.

Двухуровневая схема сопряжения элементов данной системы совпадает с аналогичной схемой системы, описанной в примере 3 (рис. 6).

Структура данных. При агрегатном подходе программная реализация исследуемого объекта не зависит от его структуры, а функционирование характеризуется данными, связанными с программой лишь формой представления информации, тогда как содержание последней определяется характером поведения объекта.

Наиболее эффективным инструментом хранения такой информации в памяти ЭВМ является аппарат списочных структур [25].

Все данные, характеризующие поведение сложной системы, сосредоточены

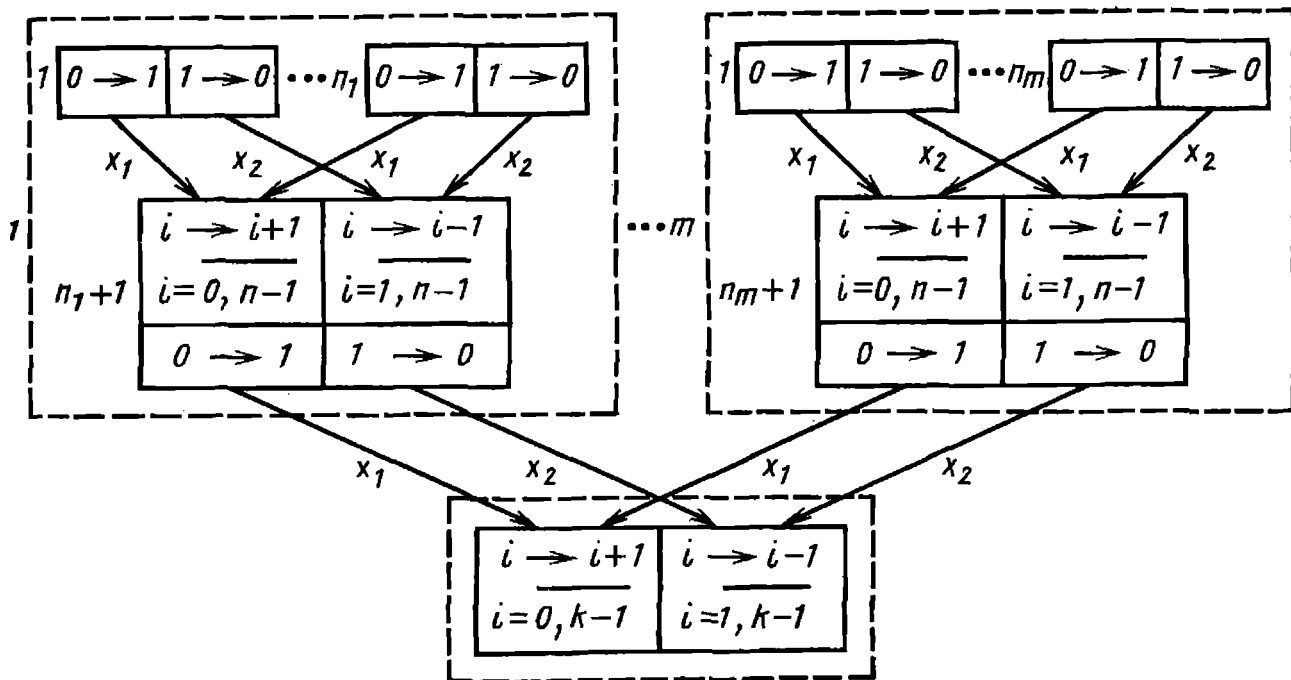
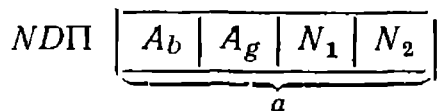


Рис. 6. Двухуровневая схема сопряжения элементов

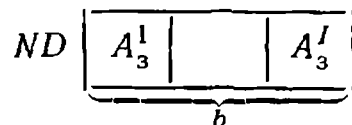
в нескольких массивах, структура которых приведена ниже.

В дальнейшем A_j будет обозначать адрес записи j , а D_j — ее длину.

Каждая запись $a = (A_b, A_g, N_1, N_2)$ соответствует своей подсистеме, так что число записей равно числу подсистем; N_1, N_2 — число агрегатов в подсистеме и число полюсов (агрегатов, через которые подсистема



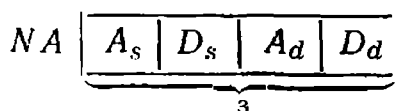
выдает сигналы на другие подсистемы). Такие записи хранятся в ЭВМ в двумерном массиве $NDП(N, 4)$, где N — число подсистем, из которых состоит исследуемая система.



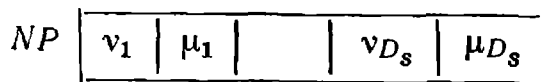
Запись b состоит из I адресов записей z . Длина записи b равна числу агрегатов соответствующей подсистемы. В ЭВМ совокупность таких записей хранится в одномерном массиве ND , причем адрес b совпадает с номером первого ее элемента.

Записи вида z имеют одинаковую структуру и длину 4 и расположены

в двумерном массиве $NA(I, 4)$, где I — число записей. Адрес A_z записи z совпадает с ее номером в массиве NA



Элементом записи s является пара (v_i, μ_i) значений дискретной компоненты агрегата, соответствующая смене состояния $v(t)$ при обращении второй компоненты в нуль. Длина D_s равна числу пар записи s , а



A_s совпадает с номером первой входящей в запись s пары. Совокупность таких записей хранится в двумерном

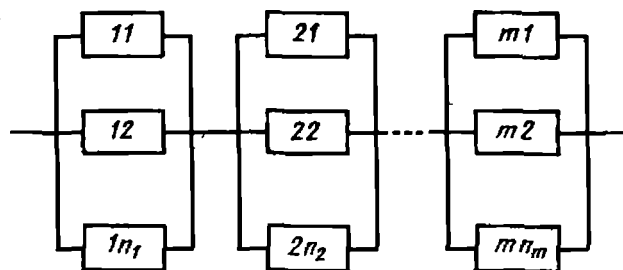


Рис. 7. Система из последовательно соединенных подсистем

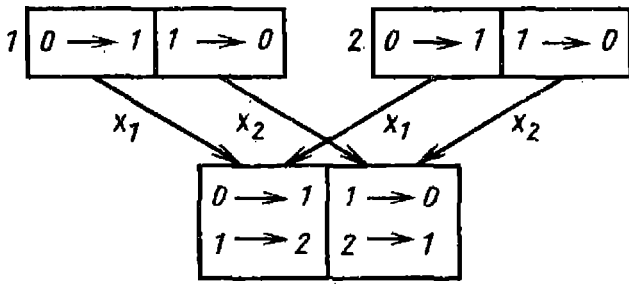


Рис. 8. Схема сопряжения дублированной системы

массиве $NP(I, 2)$, где I — число пар по всем записям массива.

Запись d состоит из D_d четверок одинаковой структуры, которые расположены в двумерном массиве $L1(I, 4)$; I — число четверок по всем записям вида d . Адресом A_d является номер первой входящей в запись d четверки массива $L1$:

$$L1 \left[\begin{array}{|c|c|c|c|} \hline i_1 & j_1 & n_1 & A_{c_1} \\ \hline \end{array} \left| \begin{array}{|c|c|c|c|} \hline i_{D_d} & j_{D_d} & n_{D_d} & A_{c_{D_d}} \\ \hline \end{array} \right. \right]_d$$

а D_d — число четверок записи; n_k — число агрегатов некоторой подсистемы, $k = 1, 2, \dots, D_d$, на которые одновременно поступают сигналы x от соответствующего агрегата (не обязательно той же подсистемы) при смене состояний (i_k, j_k) его дискретной компоненты.

Совокупность записей $\{c\}$ расположена в одномерном массиве $L2$:

$$L2 \left[\begin{array}{|c|c|c|c|} \hline m_1 & l_1 & p_1 & q_1 \\ \hline \end{array} \left| \dots \right. \left| \begin{array}{|c|c|c|c|} \hline p_{l_1} & q_{l_1} & & \\ \hline \end{array} \right. \left| \begin{array}{|c|c|c|c|} \hline m_{n_1} & l_{n_1} & p_1 & q_1 \\ \hline \end{array} \right. \left| \dots \right. \left| \begin{array}{|c|c|} \hline p_{l_{n_1}} & q_{l_{n_1}} \\ \hline \end{array} \right. \right]_c$$

Отыскание элементов внутри записи c осуществляется с помощью косвенной адресации; $m_i, i = 1, \dots, n_1$, равно номеру агрегата, на который поступает сигнал x ; l_i — число возможных реакций m_i -го агрегата на воздействие указанного сигнала, а $\{(p_j, q_j), j = 1, 2, \dots, l_i\}$ — множество таких реакций. A_c совпадает с номером первого элемента записи c в массиве $L2$.

Для учета информации о схеме сопряжения 2-го уровня вводятся записи $\{g\}$, расположенные в одномерном массиве KU , где

$$KU \left[\begin{array}{|c|c|c|c|c|c|c|} \hline L & M & E_1 & A_d^1 & & E_M & A_d^M \\ \hline \end{array} \right]_g$$

L — номер агрегата — источник сигнала x , а $E_i, i = 1, 2, \dots, M$, — номера подсистем, на которые одновременно поступает x . A_g совпадает с номером первого входящего в запись g элемента массива KU .

Задание для каждого состояния дискретной компоненты агрегата функции распределения $F(t)$ с параметрами $\lambda_i, i = 1, 2$, а также скорости α изменения непрерывной компоненты сводится к заданию записи r вида

$$FP \left[\begin{array}{|c|c|c|c|} \hline f & \lambda_1 & \lambda_2 & \alpha \\ \hline \end{array} \right]_r$$

где f — код функции распределения $F(t)$. Такие записи хранятся в двумерном массиве $FP(I, 4)$, где I — число записей. Адрес A_r совпадает с номером записи.

Предполагают, что в начальный мо-

мент времени $t_0 = 0$ с вероятностью $1v(t_0) = 0$ для всех агрегатов всех подсистем.

Пример 5. Если положить в примере 1 (рис. 2) $n = 2$, то получится дублированная система. Ее схема сопряжения изображена на рис. 8.

Исходные данные для такой системы располагаются по описанным выше массивам следующим образом:

<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">$ND\Pi$</td> <td style="padding: 2px;">A_{b_1}</td> <td style="padding: 2px;">A_{g_1}</td> <td style="padding: 2px;">N_1</td> <td style="padding: 2px;">N_2</td> </tr> <tr> <td></td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">0</td> </tr> <tr> <td></td> <td colspan="4" style="padding: 2px;">a_1</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">NA</td> <td style="padding: 2px;">A_{s_1}</td> <td style="padding: 2px;">D_{s_1}</td> <td style="padding: 2px;">A_{d_1}</td> <td style="padding: 2px;">D_{d_1}</td> </tr> <tr> <td></td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">2</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">2</td> </tr> <tr> <td></td> <td colspan="4" style="padding: 2px;">$s_{1,2}$</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">$L1$</td> <td style="padding: 2px;">i_1</td> <td style="padding: 2px;">j_1</td> <td style="padding: 2px;">n_1</td> <td style="padding: 2px;">A_{c_1}</td> <td style="padding: 2px;">i_2</td> <td style="padding: 2px;">j_2</td> <td style="padding: 2px;">n_2</td> <td style="padding: 2px;">A_{c_2}</td> </tr> <tr> <td></td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">7</td> </tr> <tr> <td></td> <td colspan="8" style="padding: 2px;">d_1</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">$L2$</td> <td style="padding: 2px;">m_1</td> <td style="padding: 2px;">l_1</td> <td style="padding: 2px;">p_1</td> <td style="padding: 2px;">q_1</td> <td style="padding: 2px;">p_2</td> <td style="padding: 2px;">q_2</td> <td style="padding: 2px;">m_2</td> <td style="padding: 2px;">l_2</td> <td style="padding: 2px;">p_1</td> <td style="padding: 2px;">q_1</td> <td style="padding: 2px;">p_2</td> <td style="padding: 2px;">q_2</td> </tr> <tr> <td></td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">2</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">2</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">2</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">2</td> <td style="padding: 2px;">1</td> </tr> <tr> <td></td> <td colspan="6" style="padding: 2px;">c_1</td> <td colspan="6" style="padding: 2px;">c_2</td> </tr> </table>	$ND\Pi$	A_{b_1}	A_{g_1}	N_1	N_2		1	1	3	0		a_1				NA	A_{s_1}	D_{s_1}	A_{d_1}	D_{d_1}		1	2	1	2		$s_{1,2}$				$L1$	i_1	j_1	n_1	A_{c_1}	i_2	j_2	n_2	A_{c_2}		0	1	1	1	1	0	1	7		d_1								$L2$	m_1	l_1	p_1	q_1	p_2	q_2	m_2	l_2	p_1	q_1	p_2	q_2		3	2	0	1	1	2	3	2	1	0	2	1		c_1						c_2						<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">ND</td> <td style="padding: 2px;">A_3^1</td> <td style="padding: 2px;">A_3^2</td> <td style="padding: 2px;">A_3^3</td> </tr> <tr> <td></td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td></td> <td colspan="3" style="padding: 2px;">b_1</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">NP</td> <td style="padding: 2px;">v_1</td> <td style="padding: 2px;">μ_1</td> <td style="padding: 2px;">v_2</td> <td style="padding: 2px;">μ_2</td> </tr> <tr> <td></td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td></td> <td colspan="4" style="padding: 2px;">s_1</td> </tr> </table>	ND	A_3^1	A_3^2	A_3^3		1	1	0		b_1			NP	v_1	μ_1	v_2	μ_2		0	1	1	0		s_1			
$ND\Pi$	A_{b_1}	A_{g_1}	N_1	N_2																																																																																																																								
	1	1	3	0																																																																																																																								
	a_1																																																																																																																											
NA	A_{s_1}	D_{s_1}	A_{d_1}	D_{d_1}																																																																																																																								
	1	2	1	2																																																																																																																								
	$s_{1,2}$																																																																																																																											
$L1$	i_1	j_1	n_1	A_{c_1}	i_2	j_2	n_2	A_{c_2}																																																																																																																				
	0	1	1	1	1	0	1	7																																																																																																																				
	d_1																																																																																																																											
$L2$	m_1	l_1	p_1	q_1	p_2	q_2	m_2	l_2	p_1	q_1	p_2	q_2																																																																																																																
	3	2	0	1	1	2	3	2	1	0	2	1																																																																																																																
	c_1						c_2																																																																																																																					
ND	A_3^1	A_3^2	A_3^3																																																																																																																									
	1	1	0																																																																																																																									
	b_1																																																																																																																											
NP	v_1	μ_1	v_2	μ_2																																																																																																																								
	0	1	1	0																																																																																																																								
	s_1																																																																																																																											

Массив KU не задается, поскольку система состоит из одной подсистемы, а задание массива FP очевидно.

Пример 6. Если в примере 3 (рис. 5) положить $n_1 = n_2 = m = 2$, то получится система со следующей структурой (рис. 9).

Решение. Схема сопряжения для такой системы выглядит так (рис. 10).

Система состоит из трех подсистем. Для нее исходные данные располагаются по массивам в следующем порядке:

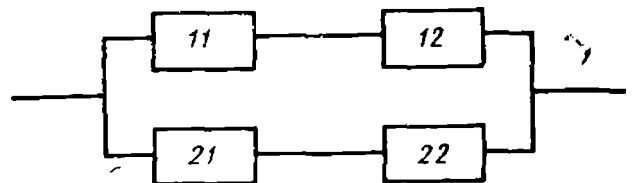


Рис. 9. Система из четырех элементов (первое соединение)

Решение. Схема сопряжения для этого случая представлена на рис. 12. Первая подсистема описывает функ-

<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">$ND\Pi$</td> <td style="padding: 2px;">A_{b_1}</td> <td style="padding: 2px;">A_{g_1}</td> <td style="padding: 2px;">N_1</td> <td style="padding: 2px;">N_2</td> <td style="padding: 2px;">A_{b_2}</td> <td style="padding: 2px;">A_{g_2}</td> <td style="padding: 2px;">N_1</td> <td style="padding: 2px;">N_2</td> <td style="padding: 2px;">A_{b_3}</td> <td style="padding: 2px;">A_{g_3}</td> <td style="padding: 2px;">N_1</td> <td style="padding: 2px;">N_2</td> </tr> <tr> <td></td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td></td> <td colspan="4" style="padding: 2px;">a_1</td> <td colspan="4" style="padding: 2px;">a_2</td> <td colspan="4" style="padding: 2px;">a_3</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">ND</td> <td style="padding: 2px;">A_3^1</td> <td style="padding: 2px;">A_3^2</td> <td style="padding: 2px;">A_3^3</td> </tr> <tr> <td></td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td></td> <td colspan="3" style="padding: 2px;">b_3</td> </tr> <tr> <td></td> <td colspan="3" style="padding: 2px;">$b_{1,2}$</td> </tr> </table>	$ND\Pi$	A_{b_1}	A_{g_1}	N_1	N_2	A_{b_2}	A_{g_2}	N_1	N_2	A_{b_3}	A_{g_3}	N_1	N_2		1	1	3	1	1	3	1	3	0	0	1	0		a_1				a_2				a_3				ND	A_3^1	A_3^2	A_3^3		1	1	0		b_3				$b_{1,2}$			<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">KU</td> <td style="padding: 2px;">L_1</td> <td style="padding: 2px;">M_1</td> <td style="padding: 2px;">E_1</td> <td style="padding: 2px;">A_{d_1}</td> </tr> <tr> <td></td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">3</td> <td style="padding: 2px;">1</td> </tr> <tr> <td></td> <td colspan="4" style="padding: 2px;">$g_{1,2}$</td> </tr> </table>	KU	L_1	M_1	E_1	A_{d_1}		3	1	3	1		$g_{1,2}$			
$ND\Pi$	A_{b_1}	A_{g_1}	N_1	N_2	A_{b_2}	A_{g_2}	N_1	N_2	A_{b_3}	A_{g_3}	N_1	N_2																																																											
	1	1	3	1	1	3	1	3	0	0	1	0																																																											
	a_1				a_2				a_3																																																														
ND	A_3^1	A_3^2	A_3^3																																																																				
	1	1	0																																																																				
	b_3																																																																						
	$b_{1,2}$																																																																						
KU	L_1	M_1	E_1	A_{d_1}																																																																			
	3	1	3	1																																																																			
	$g_{1,2}$																																																																						

Все остальные массивы остаются теми же, что и в примере 5.

Пример 7. Элементы системы, описанной в примере 6, объединены в следующую структуру (рис. 11).

ционирование последовательно соединенных элементов \mathcal{E}_{11} , \mathcal{E}_{12} , а вторая подсистема — параллельно соединенных элементов \mathcal{E}_{13} , \mathcal{E}_{21} .

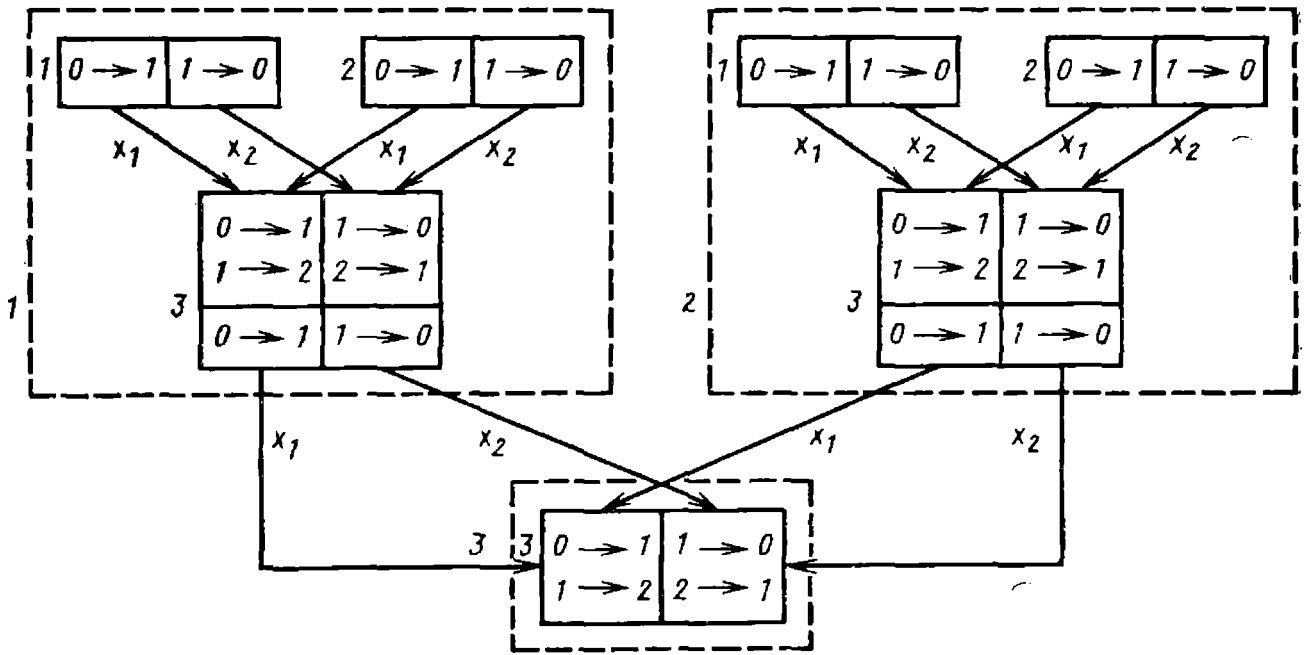


Рис. 10. Схема сопряжения системы, изображенной на рис. 9

Исходные данные для такой системы незначительно отличаются от данных предыдущей системы, а именно: в массиве $L1$ появляется еще одна запись d_2 (в предыдущем примере $d_1 = d_2$), так что значения элементов массива $L1$ таковы:

	i_1	j_1	n_1	A_{c_1}	i_2	j_2	n_2	A_{c_2}	i_1	j_1	n_1	A_{c_1}	i_2	j_2	n_2	A_{c_2}
$L1$	0	1	1	1	1	0	1	7	1	2	1	1	2	1	1	7
	d_1								d_2							

Элементы (i_1, j_1) и (i_2, j_2) записи d_2 соответствуют посылке 3-м агрегатом второй подсистемы выходных сигналов x_1 и x_2 соответственно.

Изменения в массиве $L1$ привели к появлению новой записи g_2 в массиве KU (в предыдущем примере $g_1 = g_2$), что

	L_1	M_1	E_1	A_{d_1}	L_2	M_2	E_2	A_{d_2}
KU	3	1	3	1	3	1	3	3
	g_1				g_2			

повлекло изменение адреса g_2 в массиве $NDП$ ($A_{g_2} = 4$).

Архитектура ППП АМОС. Архитектура пакета определяется тремя компонентами:

1. Функциональным наполнением отражающим специфику предметной области. Оно состоит из предметных модулей (ПМ), предназначенных для формирования структур данных, определяющих функционирование агрегатных моделей подсистем, и модулей свя-

зи (МС), которые задают сопряжение отдельных базовых моделей на втором уровне иерархии.

2. Языком задания, который является средством общения пользователя с пакетом. Входной язык (ВЯ) АМОСа — проблемно-ориентированный, позволяющий формировать задачи на содержа-

тельном уровне. По функциональному назначению все операторы (ВЯ) можно условно разделить на три группы:

- формирующие исследуемую схему;
- осуществляющие процесс счета;

предоставляющие пользователю дополнительные возможности для отладки программ, не имеющих в библиотеке пакета.

3. Системным наполнением, которое, используя элементы функционального наполнения и понимая конструкции языка, реализует дисциплину работы с пакетом. Оно выполняет следующие функции:

- управляет данными и памятью;
- организует вычислительный процесс;
- обслуживает пользователей.

Режимы работы. АМОС может функционировать как в пакетном, так и в диалоговом режимах. Эффективность решения проблемных задач во многих случаях может быть существенно повышена при использовании диалога в двухэтапном режиме. Первый этап позволяет формировать исходные данные задачи, а второй — корректировать развитие вычислительного процесса.

Второй этап целесообразен в случае аналитико-статистического моделирования систем, которое наряду с аналитическими преобразованиями позволяет исследовать систему с помощью статистического моделирования.

Используемые при этом аналитико-статистические методы для расчета вероятностных характеристик высоконадежных систем позволяют сократить

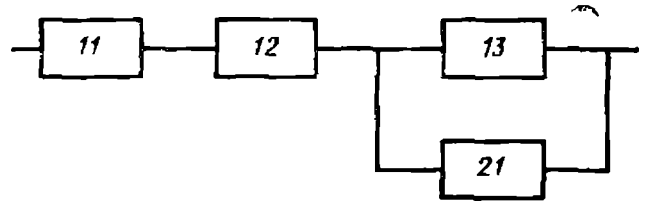


Рис. 11. Система из четырех элементов (второе соединение)

время проведения вычислительного эксперимента от нескольких часов до нескольких минут.

Реализация аналитико-статистических методов непосредственно на агрегатной модели системы позволяет эффективное их использование в рамках ППП АМОС. Это обстоятельство облегчает пользователю применять их к решению задач, не программируя вычислительный алгоритм метода. Для этого достаточно построить агрегатную модель исследуемой системы и указать метод с помощью оператора ВЯ пакета.

Пример 8. Система состоит из четырех элементов, используемых в предыдущих примерах. Их можно условно разделить на два типа по два элемента, в зависимости от функции распределения $F_k(t) = 1 - e^{-\lambda_k t}$ времени безотказной работы k -го элемента и функ-

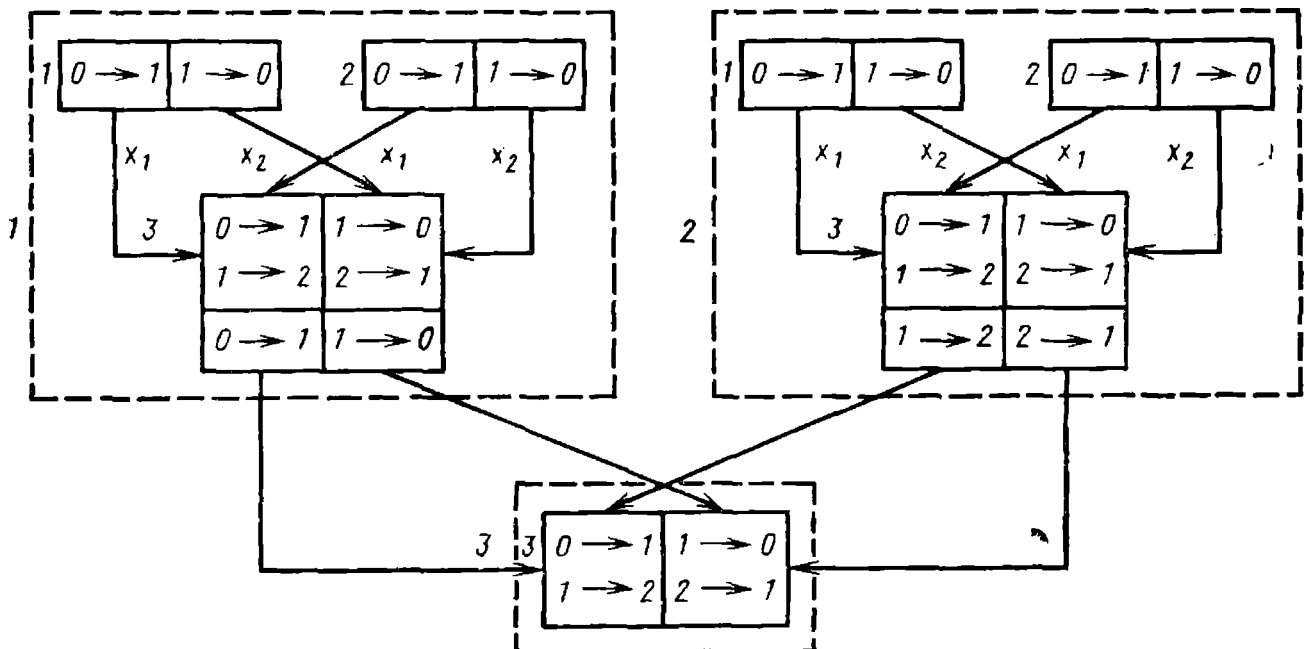


Рис. 12. Схема сопряжения системы, изображенной на рис. 11

ции распределения $G_k(t) = 1 - e^{-\mu_k t^2}$ времени восстановления, $k = 1, 2, 3, 4$. Первые два элемента относятся к первому типу ($\lambda_1 = \lambda_2 = 1,4 \cdot 10^{-5}$; $\mu_1 = \mu_2 = 1,1 \cdot 10^{-2}$), а остальные — ко второму ($\lambda_3 = \lambda_4 = 1,87 \cdot 10^{-5}$; $\mu_3 = \mu_4 = 2,3 \cdot 10^{-2}$). Восстановление k -го элемента начинается сразу же после его отказа.

Отказ системы наступает в момент одновременного отказа не менее одного элемента каждого типа.

Требуется определить вероятность $P(T)$ безотказной работы системы за время T .

Решение. Надежностная схема такой системы представлена на рис. 9, а ее агрегатная модель и схема сопряжения рассмотрены в примере 6.

Задание пакету на его входном языке для такой задачи состоит из 16 операторов:

1. ПАКЕТ
2. ВРЕМЯ: 5
3. РАЗМЕРЫ: P1=3,3,1; P2=3; P3=4; P4=4; P5=8
4. ВЗЯТЬ: СХЕМА 1
5. ФУНКЦИЯ: (EX (0.000014), VEI (0.011)) * 2
6. ВЗЯТЬ: СХЕМА 1
7. ФУНКЦИЯ: (EX 0.0000187), VEI (0.023)) * 2
8. ВЗЯТЬ: СОПРЯЖЕНИЕ 1
9. МЕТОД: 1
10. ЧИСЛО ТИПОВ = 2
11. ИНДЕКС = 1
12. ВЕКТОР = 1,1
13. ВЫЧ: ВБР (1000.)
14. E = 0.000001
15. B = 1.96
16. МОДЕЛ

В конце каждого оператора необходимо ставить знак « \diamond » — признак

конца строки. Оператор 1 указывает режим использования пакета. В 3-м операторе параметры P2, ..., P6 задают размеры массивов ND, NA, NP, L1, L2 соответственно, а вектор P1 указывает на число агрегатов по подсистемам.

СХЕМА 1 соответствует подпрограмме формирования и засылки в соответствующие массивы данных о функционировании первой подсистемы (вторая подсистема совпадает с первой).

EX (λ), VEI (μ) задают функции распределения $F(t)$ и $G(t)$, а коэффициент 2 при них соответствует идентичности двух агрегатов.

СОПРЯЖЕНИЕ 1 вызывает MC, который сопрягает три подсистемы в единую модель описанной системы.

МЕТОД 1 задает метод решения задачи (в данном случае он соответствует аналитико-статистическому методу [10]). По умолчанию, т. е. если не указывается оператор МЕТОД: I, используется метод непосредственного моделирования.

Операторы 10—12 задают параметры метода. В зависимости от метода их может быть разное количество.

ВБР (T); E указывают на вычисляемую характеристику (вероятность безотказной работы системы на $[0, T]$ и точность вычисления), а B соответствует достоверности оценки.

МОДЕЛ передает задачу за счет.

Полученные результаты расчета характеристики $P(T)$ при непосредственном моделировании и с помощью аналитико-статистического метода (МЕТОД: 1) приведены в таблице; они наглядно свидетельствуют об эффективности аналитико-статистического метода, позволяющего существенно сократить время проведения вычислительного эксперимента:

Метод	Время счета	Число реализаций	$1 - P(1000)$	Дисперсия оценки
Непосредственного моделирования	4—5 ч	$29 \cdot 10^6$	$7,9808 \cdot 10^{-6}$	$7,9501 \cdot 10^{-6}$
Аналитико-статистический	1,2 с	132	$7,4802 \cdot 10^{-6}$	$3,3040 \cdot 10^{-11}$

СПИСОК ЛИТЕРАТУРЫ

1. Бусленко В. Н. Автоматизация имитационного моделирования сложных систем. М.: Наука, 1977. 239 с.
2. Бусленко Н. П. Моделирование сложных систем. М.: Наука, 1978. 400 с.
3. Бусленко Н. П., Калашников В. В., Коваленко И. Н. Лекции по теории сложных систем. М.: Сов. радио, 1973. 430 с.
4. Гройсберг Л. Б., Полляк Ю. Г. Методы планирования машинного эксперимента при вероятностном моделировании надежности сложных систем. — Известия АН СССР. Техническая кибернетика, 1976, № 4, с. 85—90.
5. Дал У., Мюрхауг Б., Ньюгорд К. СИМУЛА-67. Универсальный язык программирования. М.: Мир, 1969, 99 с.
6. Ермаков С. М. Метод Монте-Карло и смежные вопросы. М.: Наука, 1975. 472 с.
7. Завадская Л. А. Об одном подходе к ускорению моделирования систем с резервированием. — Электронное моделирование, 1984, 6, № 3, с. 57—60.
8. Калашников В. В. Организация исследования сложных систем на базе агрегатного подхода к моделированию. — Известия АН СССР. Техническая кибернетика, 1982, № 2, с. 92—108.
9. Кендалл М., Стюарт А. Статистические выводы и связи. М.: Наука, 1973. 899 с.
10. Коваленко И. Н. Исследования по анализу надежности сложных систем. Киев: Наукова думка, 1975. 210 с.
11. Коваленко И. Н. Анализ редких событий при оценке эффективности и надежности систем. М.: Сов. радио, 1982. 209 с.
12. Коваленко И. Н. К расчету характеристик высоконадежных систем аналитико-статистическим методом. — Электронное моделирование, 1980, 2, № 4, с. 5—8.
13. Королюк В. С., Турбин А. Ф. Процессы марковского восстановления в задачах надежности систем. Киев: Наукова думка, 1982. 236 с.
14. Крейн М., Лемуан О. Введение в регенеративный метод анализа моделей. М.: Наука, 1982. 104 с.
15. Кривуца В. Г. О диалоговом агрегатном моделировании сложных систем. — В кн.: К анализу надежности сложных систем аналитико-статистическим методом. Изд. Ин-та кибернетики АН УССР, Препринт-82-2, 1982, с. 9—15.
16. Кузнецов Н. Ю. Общий подход к нахождению вероятности безотказной работы структурно-сложных систем аналитико-статистическим методом. — Кибернетика, 1985, № 3, с. 86—94.
17. Кузнецов Н. Ю. Вычисление коэффициента оперативной готовности восстанавливаемой системы аналитико-статистическим методом. — Кибернетика, 1985, № 5, с. 95—101.
18. Линник И. Ю. Улучшение сходимости метода Монте-Карло в некоторых задачах массового обслуживания. — Кибернетика, 1973, № 5, с. 129—132.
19. Марковиц Г., Хауснер Б., Карр Г. СИМСКРИПТ. Алгоритмический язык для моделирования. М.: Сов. радио, 1966. 152 с.
20. Плакс Б. И. Ускоренное вычисление малых вероятностей при моделировании на ЭВМ. — В кн.: Системы обработки и передачи информации, 1982, № 153, с. 81—87.
21. Полляк Ю. Г. Вероятностное моделирование на электронных вычислительных машинах. М.: Сов. радио, 1971. 400 с.

22. **Поляк Ю. Г.** Оценка малых вероятностей при статистическом моделировании систем. — Известия АН СССР. Техническая кибернетика, 1973, № 2, с. 197—203.

23. **Программные средства моделирования непрерывно-дискретных систем/В. М. Глушков, В. В. Гусев, Т. П. Марьянович, М. А. Сахнюк.** Киев: Наукова думка, 1975. 152 с.

24. **Соловьев А. Д.** Асимптотическое поведение момента первого наступления редкого события в регенерирующем процессе. — Известия АН СССР. Техническая кибернетика, 1971, № 6, с. 79—89.

25. **Флорес И.** Структуры и управление данными. М.: Финансы и статистика, 1982. 317 с.

26. **Шпак В. Д.** Об оценке вероятности обрыва регенерирующего процесса в течение фиксированного времени методом статистического моделирования. — Кибернетика, 1983, № 1, с. 75—79.

27. **Шрайбер Т. Дж.** Моделирование на GPSS. М.: Машиностроение, 1980. 592 с.

28. **Buxton J. N., Laski I. G.** Control and simulation language. — Comput. J., 1962, 5, N 3, p. 194—199.

29. **Knuth D. E., McNeley I. L.** SOL — a symbolic language for general purpose systems simulation. — IEEE Trans. EL. C., 1966, N 8, p. 401—408.

1. НЕПРЕРЫВНЫЕ РАСПРЕДЕЛЕНИЯ

Пусть ξ — случайная величина со значениями из R^1 и распределением P . Ее функция распределения пусть $F(x) = P\{\xi \leq x\}$.

Если P абсолютно непрерывна относительно меры Лебега, то существует функция плотности $f(x)$ и

$$F(x) = \int_{-\infty}^x f(y) dy.$$

Если ξ является случайным вектором, $\xi = (\xi_1, \dots, \xi_n)'$, то n -мерной функцией распределения ξ называют функцию

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= \\ &= F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \\ &= P\{\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n\}, \end{aligned}$$

а ее функцией плотности в случае, если P абсолютно непрерывна относительно меры Лебега в R^n , называют функцию $f(x_1, x_2, \dots, x_n) = f_{\xi_1, \xi_2, \dots, \xi_n}(x_1, \dots, x_n)$, определяемую равенством

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= \\ &= \int_{y_1 \leq x_1, \dots, y_n \leq x_n} f(y_1, y_2, \dots, y_n) dy_1 dy_2 \dots dy_n. \end{aligned}$$

Функция распределения случайных величин $\xi_1, \xi_2, \dots, \xi_n$ при условии, что $\xi_{m+1} = x_{m+1}, \dots, \xi_n = x_n$, выражается формулой

$$\begin{aligned} F_{\xi_1, \dots, \xi_m | \xi_{m+1}, \dots, \xi_n}(x_1, \dots, x_m | x_{m+1}, \dots, x_n) = \end{aligned}$$

$$\begin{aligned} &= \frac{1}{f_{\xi_{m+1}, \dots, \xi_n}(x_{m+1}, \dots, x_n)} \times \\ &\times \frac{\partial^{n-m}}{\partial x_{m+1} \dots \partial x_n} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n), \end{aligned}$$

а условная функция плотности есть

$$\begin{aligned} f_{\xi_1, \dots, \xi_m | \xi_{m+1}, \dots, \xi_n}(x_1, \dots, x_m | x_{m+1}, \dots, x_n) &= \\ &= f_{\xi_1, \dots, \xi_n}(x_1, x_2, \dots, x_n) / f_{\xi_{m+1}, \dots, \xi_n}(x_{m+1}, \dots, x_n). \end{aligned}$$

Используя преобразование Розенблатта

$$\begin{aligned} \eta_1 &= F_{\xi_1}(\xi_1), \\ \eta_2 &= F_{\xi_2 | \xi_1}(\xi_2 | \xi_1), \\ \eta_3 &= F_{\xi_3 | \xi_1, \xi_2}(\xi_3 | \xi_1, \xi_2), \\ &\dots \\ \eta_n &= F_{\xi_n | \xi_1, \dots, \xi_{n-1}}(\xi_n | \xi_1, \xi_2, \dots, \xi_{n-1}), \end{aligned}$$

можно преобразовать случайный вектор $\xi = (\xi_1, \dots, \xi_n)$ в случайный вектор $\eta = (\eta_1, \dots, \eta_n)$, имеющий равномерное распределение в n -мерном единичном кубе.

Еще один полезный класс распределений — это распределения на единичной сфере $\Omega_n = \{x : x \in R^n, x'x = 1\}$, $n \geq 1$. В качестве примера от-

метим распределение Мизеса—Фишера [26] с функцией плотности

$$f_n(x) = \frac{1}{a_n(\kappa)} \exp\{\kappa \mu'x\},$$

$$x = (x_1, \dots, x_n) \in \Omega_n,$$

$$a_n(\kappa) = (2\pi)^{n/2} I_{(n/2)-1}(\kappa) \kappa^{-(n/2)+1},$$

$$\mu'\mu = 1, \kappa \geq 0,$$

где $I_s(x)$ — модифицированная функция Бесселя первого рода и порядка s . В частности, при $n = 2$ получаем распределение Мизеса—Фишера на окружности. Оно может быть представлено (см. [3]) в полярных координатах как распределение случайного угла θ , $0 \leq \theta \leq 2\pi$, с плотностью

$$g(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\},$$

$$|\mu| < \infty, \kappa > 0.$$

Рассмотрим некоторые понятия, относящиеся к одномерным случайным величинам. В теории надежности $F(x)$ часто рассматривается как функция распределения моментов отказов, причем $\xi \geq 0$, $F(0^-) = 0$. Функцию $r(x) = f(x)/(1 - F(x))$ называют функцией интенсивности отказов. Распределения, имеющие возрастающие функции интенсивности отказов, называют ВФИ-распределениями (старееющими), а убывающие — УФИ-распределениями (молодеющими).

Моментом порядка r относительно a называют величину, если она существует,

$$\mu_r(a) = \int_{-\infty}^{\infty} (x - a)^r dF(x).$$

Если в качестве a выбрано математическое ожидание случайной величины ξ , т. е.

$$a = m = \mu_1(0) = \int_{-\infty}^{\infty} x dF(x),$$

то такие моменты $\mu_r = \mu_r(m)$ называют центральными моментами, а при $a = 0$ — начальными и обозначают μ_r^0 .

Момент μ_2 называют дисперсией и

обозначают $\sigma^2 = D\xi$. Корень из дисперсии называют средним квадратическим отклонением.

Для симметричных относительно математического ожидания распределений $\mu_{2r+1} = 0$, $r = 0, 1, 2$. В качестве меры несимметричности распределения используют коэффициент асимметрии $\gamma_1 = \mu_3/\mu_2^{3/2}$. Другой часто используемой характеристикой распределения является коэффициент эксцесса $\gamma_2 = (\mu_4/\mu_2^2 - 3)$. Для нормального распределения $\gamma_2 = 0$.

Имеют место соотношения

$$\mu_r^0 = \sum_{j=0}^r C_r^j \mu_{r-j} (\mu_1^0)^j;$$

$$\mu_r = \sum_{j=0}^r C_r^j \mu_{r-j}^0 (-\mu_1^0)^j,$$

т. е. например,

$$\mu_2^0 = \mu_2 + (\mu_1^0)^2;$$

$$\mu_3^0 = \mu_3 + 3\mu_1^0\mu_2 + (\mu_1^0)^3$$

и

$$\mu_2 = \mu_2^0 - (\mu_1^0)^2;$$

$$\mu_3 = \mu_3^0 - 3\mu_1^0\mu_2^0 + 2(\mu_1^0)^3.$$

Другими характеристиками распределений являются семинварианты κ_k , $k = 1, 2$, определяемые при каж-

дом k как коэффициенты при $\frac{(it)^k}{k!}$

в разложении логарифма характеристической функции $\psi(t)$ (в то время как начальные моменты μ_k^0 являются

коэффициентами при $\frac{(it)^k}{k!}$ в разложении $\psi(t)$).

Начальные моменты μ_k^0 являются полиномами от семинвариантов

$$\mu_k^0 = \sum \left(\frac{\kappa_1}{1!}\right)^{\pi_1} \left(\frac{\kappa_2}{2!}\right)^{\pi_2} \dots \left(\frac{\kappa_k}{k!}\right)^{\pi_k} \frac{k!}{\pi_1! \pi_2! \dots \pi_k!},$$

где суммирование ведется по всем наборам неотрицательных целых чисел π_1, \dots, π_k , удовлетворяющих равенству

$$\pi_1 + 2\pi_2 + \dots + k\pi_k = k.$$

Например,

$$\mu_1^0 = \kappa_1;$$

$$\mu_2^0 = \kappa_2 + \kappa_1^2;$$

$$\mu_3^0 = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3.$$

Семиинварианты выражаются через начальные моменты по формуле

$$\begin{aligned} \kappa_k &= \sum (-1)^{\pi_1 + \dots + \pi_k - 1} \times \\ &\times \left(\frac{\mu_1^0}{1!} \right)^{\pi_1} \left(\frac{\mu_2^0}{2!} \right)^{\pi_2} \dots \left(\frac{\mu_k^0}{k!} \right)^{\pi_k} \times \\ &\times \frac{k! (\pi_1 + \dots + \pi_k - 1)!}{\pi_1! \pi_2! \dots \pi_k!}, \end{aligned}$$

где суммирование производится по всем наборам целых неотрицательных чисел $\{\pi_1, \pi_2, \dots, \pi_k\}$, удовлетворяющих равенству $\pi_1 + 2\pi_2 + \dots + k\pi_k = k$.

Например,

$$\kappa_1 = \mu_1^0;$$

$$\kappa_2 = \mu_2^0 - (\mu_1^0)^2;$$

$$\kappa_3 = \mu_3^0 - 3\mu_2^0\mu_1^0 + 2(\mu_1^0)^3.$$

Разложение Грама—Шарлье для непрерывной функции плотности $f(x)$ одномерной случайной величины может быть формально введено следующим образом:

$$f(x) = \sum_{i=0}^{\infty} a_i H_i(x) \varphi(x),$$

где $\varphi(x)$ — функция плотности нормированной нормально распределенной случайной величины (с нулевым математическим ожиданием и единичной дисперсией); $H_i(x)$ — полиномы Эрмита,

$$a_i = \frac{1}{i!} \int_{-\infty}^{\infty} f(x) H_i(x) dx, \quad i = 1, 2, \dots$$

Выражая a_i через начальные моменты μ_3^0 и предполагая, что $f(x)$ — функция плотности нормированной случайной величины, получаем, что

$$\begin{aligned} f(x) &= \varphi(x) \left\{ 1 + \frac{1}{6} \mu_3^0 H_3(x) + \right. \\ &\left. + \frac{1}{24} (\mu_4 - 3) H_4(x) + \dots \right\}. \end{aligned}$$

Переходя к семиинвариантам, получаем, что

$$\begin{aligned} f(x) &= \varphi(x) \left(1 + \frac{\kappa_3}{6} H_3(x) + \right. \\ &+ \frac{\kappa_4}{24} H_4(x) + \frac{\kappa_5}{120} H_5(x) + \\ &\left. + \frac{\kappa_6 + 10\kappa_3^2}{720} H_6(x) + \dots \right). \end{aligned}$$

Разложение Эджворта для функции плотности $f(x)$ нормированной случайной величины ($\kappa_1 = 0, \kappa_2 = 1$) вводится формально так:

$$\begin{aligned} f(x) &= \exp \left\{ -\kappa_3 \frac{D^3}{3!} + \right. \\ &\left. + \kappa_4 \frac{D^4}{4!} - \dots \right\} \varphi(x), \end{aligned}$$

где $D = \frac{d}{dx}$ — оператор дифференцирования.

Квантиль u_p уровня p (p -квантиль, 100 p -процентиль) определяется как корень уравнения $p = F(u_p)$, т. е. $u_p = F^{-1}(p)$. В частности, $u_{1/2}$ — медиана, $u_{1/4}$ и $u_{3/4}$ — нижняя и верхняя квантили, $u_{j/10}$ — j -дециль, $j = 1, 2, \dots, 9$.

Для ряда распределений существуют приближенные формулы для нахождения значений квантили u_p . Если приближенное значение u_p^1 , полученное по такого рода формуле, достаточно мало отличается от истинного значения u_p , то может быть использован следующий способ уточнения значения u_p^1 .

Пусть $p_1 = F(u_p^1)$. Тогда $u_p^1 = u_{p_1}$. Разложение u_p как функции p в окрест-

ности точки p_1 (разложение Корниша—Фишера) имеет вид

$$u_p = u_{p_1} + \sum_{k=1}^{\infty} \frac{c_k(u_{p_1})}{[f(u_{p_1})]^k} (p - p_1)^k.$$

Здесь в предположении, что существуют все произвольные функции распределения $F(x)$, функция $c_k(u_p)$ определяется рекуррентной формулой

$$c_{k+1}(u) = kc_k(u) \kappa(u) + c'_k(u), \\ c_1 = 1, k = 1, 2, \quad \kappa(u) = -f'(u)/f(u).$$

Вопрос об интервале сходимости этого ряда решается для каждого распределения отдельно. Взяв несколько первых членов ряда и вычислив предварительно точное значение $p_1 = F(u_p^1)$, можно получить очередное приближение u_p^2 . На втором шаге вместо p_1 следует взять $p_2 = F(u_p^2)$, а затем процедура повторяется. Значение u_p^i , получаемое на некотором шаге i , рассматривается как приближение для u_p .

Приведем несколько методов численного интегрирования, используемых при вычислении значений различных функций распределения. Пусть требуется вычислить значение интеграла

$$F = \int_a^b f(t) dt.$$

(Если интегрирование производится на бесконечном интервале, то тем или иным способом оцениваются ошибки, возникающие на хвостах, и интеграл сводится к интегралу по конечному интервалу.)

Пусть отрезок $[a, b]$ разбивается на m равных отрезков длины $h = (b - a)/m$. Тогда в соответствии с формулой трапеций

$$F = h \left(\frac{1}{2} f_0 + f_1 + \dots + f_{m-1} + \frac{1}{2} f_m \right) + \varepsilon,$$

$$f_i = f(a + h_i), \quad i = 0, 1, \dots, m,$$

где ошибка ε может быть оценена по формуле

$$\varepsilon = -\frac{(b-a)^3}{12m^2} f''(\tau), \quad \tau \in [a, b].$$

Чаше используют формулу Симпсона

$$F = \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 4f_{2m-3} + 2f_{2m-2} + 4f_{2m-1} + f_{2m}) + \varepsilon,$$

$$\varepsilon = -\frac{(b-a)^5}{180m^4} f^{(4)}(\tau), \quad \tau \in [a, b].$$

Учитывая, что оценка ошибки ε по приведенным формулам бывает затруднительна, на практике используют метод удвоения числа разбиений, когда значение F , полученное при некотором $m = m_0$, сравнивают со значением F при $m = 2m_0$. Обычно есть основания полагать, что первые совпадающие значащие цифры верны.

Рассмотрим теперь метод численного интегрирования Ромберга. Пусть $m = 2^k, k = 0, 1, 2, \dots$ и F_{0k} вычисляют по формуле трапеций

$$F_{0k} = h \left(\frac{1}{2} f_0 + f_1 + \dots + \frac{1}{2} f_m \right).$$

Определяют рекуррентное соотношение

$$F_{sk} = \frac{1}{4^s - 1} (4^s F_{s-1, k+1} - F_{s-1, k}).$$

Затем вычисляют следующую треугольную таблицу значений:

$$\begin{array}{c} F_{00} \\ F_{01} F_{10} \\ F_{02} F_{11} F_{20} \\ \dots \\ F_{0s} F_{1, s-1} \quad F_{s0} \end{array}$$

Значение F_{s0} и служит приближенным значением интеграла F . Здесь удвоение m наступает при увеличении s на 1.

2. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Наиболее важным распределением, встречающимся при применении вероятностных и статистических методов, является нормальное (гауссовское) распределение. Нормальным распределением часто можно с удовлетворительной точностью аппроксимировать характеристики объектов, определяемые действием многих случайных факторов. Ряд основных распределений, таких, как, например, χ^2 -, t - и F -распределение, определяются как распределения некоторых случайных величин, функционально зависящих от нормально распределенных случайных величин. Изучению нормального распределения посвящено большое число работ, однако тем не менее задача вычисления функции многомерного нормального распределения не имеет удовлетворительного решения в случае размерности, превышающей четвертый или пятый порядок.

Рассмотрим одномерное нормальное распределение.

Случайная величина ξ имеет нормальное распределение с математическим ожиданием m и дисперсией σ^2 ($\xi \sim N(\mu, \sigma^2)$), если ее функция распределения есть

$$F(x) = P\{\xi \leq x\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy,$$

а функция плотности соответственно

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Характеристическая функция случайной величины ξ имеет вид

$$\psi(t) = Ee^{i\xi t} = \exp\left\{i\mu t - \frac{\sigma^2 t^2}{2}\right\}.$$

Случайная величина ξ имеет стандартное нормальное распределение, если $m = 0$, $\sigma^2 = 1$, т. е. $\xi \sim N(0, 1)$. Ее функция распределения

$$\Phi(x) = F(x\sigma + \mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy,$$

а функция плотности $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

С функцией нормального распределения связан интеграл ошибок

$$\operatorname{erf}(x) = 2\Phi(\sqrt{2}x) - 1 = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy.$$

Часто используют отношение Миллса

$$R(x) = \frac{1 - \Phi(x)}{\varphi(x)}.$$

Центральные моменты нормального распределения $N(\mu, \sigma^2)$ есть

$$\mu_{2r} = \frac{(2r)!}{2^r r!} \sigma^{2r}.$$

Существуют подробные таблицы нормального распределения [8].

Приведем следующие разложения для функции $\Phi(x)$. Во-первых, это разложение

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left(x - \frac{x^3}{1! 2^1 3} + \frac{x^5}{2! 2^2 5} - \frac{x^7}{3! 2^3 7} + \dots \right),$$

получаемое разложением функции плотности $\varphi(x)$ в ряд по степеням x и дальнейшим почленным интегрированием. Это разложение пригодно для вычислений при $x < 3$.

Во-вторых, это разложение для отношения Миллса

$$R(x) = \frac{1}{x} - \frac{1}{x^3} + \frac{1 \cdot 3}{x^5} - \dots + (-1)^j \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2j-1)}{x^{2j+1}} +$$

$$+ R_j(x).$$

Этот ряд является асимптотическим, причем остаточный член $R_j(x)$ по абсолютной величине не превосходит последний член разложения перед

$R_j(x)$. Этот ряд имеет в основном только теоретическую ценность. Известны и другие разложения [6].

Наиболее полезным с практической точки зрения являются формулы для вычисления функции нормального распределения с использованием непрерывных дробей. Для отношения Миллса она имеет вид

$$R(x) = \frac{1}{x + \frac{1}{x + \frac{2}{x + \frac{3}{x + \dots}}}}$$

или в условной записи

$$R(x) = \frac{1}{x + \frac{1}{x + \frac{2}{x + \frac{3}{x + \dots}}}}$$

Это разложение справедливо при любых значениях x и дает хорошие результаты при $x \geq 3$.

Другое разложение в непрерывную дробь предложено Шентоном [22] для

$$\bar{R}(x) = \frac{\Phi(x) - 1/2}{\varphi(x)} = \frac{x}{1 - \frac{x^2}{3 + \frac{2x^2}{5 - \frac{3x^2}{7 + \dots}}}} \times \frac{4x^2}{9 - \frac{5x^2}{11 + \frac{6x^2}{13 - \frac{7x^2}{15 + \dots}}}}$$

Здесь звездочками обозначены сверху позиции, оборвав на которых непрерывную дробь получаем значение $\bar{R}(x)$ с избытком. В противном случае — с недостатком. Это позволяет оценить ошибку приближения. Разложение Шентона справедливо при $x > 0$, но особенно удобно при $0 \leq x \leq 3$.

Эти две непрерывные дроби уже при небольшом числе первых членов дроби позволяют вычислять значения $\Phi(x)$ с достаточно высокой точностью. Подпрограмма на языке ФОРТРАН, приводимая ниже и предназначенная для вычисления значений функции $\Phi(x)$, основана на описанных непрерывных дробях. Знаменатель, на котором должна обрываться дробь для достижения заданной точности, определяется по некоторой эмпирической формуле автоматически.

FUNCTION NORM (X)

B=ABS (X)

X2=X * X

F=0

P=EXP (-X2/2) * 0.3989422804

IF (B-3) 2,3,3

2 I1=B

I1=I1 * 4+6

M=0

DO 4 I3=2,I1,2

I=I1-M

F=(I-1) * X2/(2 * I-1+
+I * X2/(2 * I+1-F))

4 M=M+2

TNORM=0.5+X * P/(1-F)

RETURN

3 I2=90/B

M1=2 * I+1

F=I2+1

DO 5 I=1,I2

M1=M1-1

5 F=M1/(B+F)

F=P/(B+F)

IF(X)6,7,7

6 TNORM=F

7 TNORM=1-F

RETURN

END

Эта программа позволяет достичь точность 10^{-8} — 10^{-9} (на ЭВМ серии ЕС — 10^{-6} — 10^{-7}). С использованием рассмотренных непрерывных дробей может быть легко получена и значительно более высокая точность. В приведенной программе оказалось удобным непрерывную дробь вычислять в обратном порядке, оборвав ее на некотором знаменателе.

Часть непрерывной дроби вида

$$\frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \dots + \frac{a_n}{b_n}}}}$$

может быть представлена как A_n/B_n , причем A_n и B_n находят с помощью рекуррентных соотношений Уоллиса—Эйлера:

$$A_k = B_k A_{k-1} + a_k A_{k-2};$$

$$B_k = b_k B_{k-1} + a_k B_{k-2}, \quad k = 1, 2,$$

где $A_{-1} = 1$, $A_0 = 0$, $B_{-1} = 0$, $B_0 = 1$.

Для того чтобы получить быстродействующие подпрограммы для вычисления $\Phi(x)$ при ограниченной точности, могут быть использованы аппроксимации для $R(x)$, основанные на

дробно-рациональных функций вида

$$R(x) \approx \frac{a_0 + a_1x + \dots + a_nx^n}{b_0 + b_1x + \dots + b_mx^m}.$$

Коэффициенты $a_i, i = 0, 1, \dots, n, b_j, j = 0, 1, \dots, m$, определяются численными методами, исходя из условия наилучшего приближения $R(x)$ при заданных n и m на некотором интервале. Один пример такой аппроксимации содержится в работе [18].

Если требуется получить подробные таблицы нормального распределения, то наиболее быстрым методом их вычисления может оказаться метод численного интегрирования Симпсона, так как на каждом шаге интегрирования получают очередное значение в таблице.

Функция

$$v = \frac{1}{c_1^2} \varphi(u(c_1t - c_2))$$

при $0 \leq c_1t - c_2 \leq 1$ удовлетворяет дифференциальному уравнению $vv'' = -1$. Константы c_1 и c_2 определяются по начальным условиям.

Рассмотрим задачу вычисления квантилей нормального распределения. Обозначим через $u(p)$ функцию, обратную к функции нормального распределения ($u(p)$ — квантиль уровня p), $0 < p < 1$.

Часто значения функции $u(p)$ находят путем численного решения уравнения $p = \Phi(u(p))$ по методу Ньютона. Однако существуют и более эффективные процедуры. Исходя из разложения Корниша—Файшера (см. с. 254) в рассматриваемом случае

$$u(p) = u(p_1) + \sum_{i=1}^n \frac{V(u(p_1))(p - p_1)^i}{i! \varphi^i(u(p_1))} + r_n(p),$$

где $V_1(x) = 1, V_k(x) = (k-1) \times \times x V_{k-1}(x) + V'_{k-1}(x), k = 2, 3,$

Для остатка $r_n(p)$ справедливо неравенство

$$|r_n(p)| <$$

$$\begin{cases} \frac{1}{2(n+1)} \left(\frac{2|p - p_1|}{p_2} \right)^{n+1} \\ \text{если } -u(p_2) \geq 1, \\ \frac{1}{2(n+1)} (\sqrt{8\pi e} |p - p_1|)^{n+1}, \\ \text{если } -u(p_2) < 1, \end{cases}$$

$$p, p_1 < 1/2, p_2 = \min(p, p_1).$$

Значение $u(p)$ можно вычислять по итерационной формуле

$$X_{N+1} = X_N + \sum_{i=1}^n \frac{V_i(X_N) [p - \Phi(X_N)]^i}{i! \varphi^i(X_N)},$$

$$N = 1, 2,$$

X_1 — какое-либо приближенное значение $u(p)$.

При малых значениях p для получения приближенных значений u_p могут быть полезными следующие формулы:

$$u(p) = -\sqrt{-2 \ln(\sqrt{2\pi p})} + 0 \left(\frac{\ln |\ln p|}{|\ln p|^{1/2}} \right),$$

$$u(p) = -\sqrt{-2 \ln(\sqrt{2\pi p} \sqrt{-2 \ln(\sqrt{2\pi p})})} + 0 \left(\frac{\ln |\ln p|}{|\ln p|^{3/2}} \right).$$

При достаточно малых p и хорошем начальном приближении X_1 сходится следующий итерационный процесс:

$$X_{k+1} = -\sqrt{-2 \ln(p/R(-X_k))},$$

$$k = 0, 1, 2, \dots,$$

причем $\lim_{k \rightarrow \infty} X_k = u(p)$.

Для функции u_p известны также аппроксимации дробно-рациональными функциями

$$u_p = t - \frac{c_0 + c_1t + c_2t^2}{1 + d_1t + d_2t^2 + d_3t^3} + \varepsilon(p),$$

$$t = [-2 \ln(p)]^{1/2},$$

$$|\varepsilon(p)| < 0,00045;$$

$$c_0 = 2,515517; \quad c_1 = 0,802853;$$

$$c_2 = 0,010328; \quad d_1 = 1,432788;$$

$$d_2 = 0,189269; \quad d_3 = 0,001308.$$

Более точные аппроксимации содержатся в работе [18].

Имеет место также разложение [2]

$$u^2(p) = \sum_{i=1}^{\infty} a_i [-\ln(1 - 4(t - 1/2)^2)]^i,$$

причем коэффициенты a_i находят из системы уравнений

$$\sum_{k=0}^{i-1} a_{i-k} b_{k+1, i-k} = \frac{(\pi/2)^i}{(2i)!} c_{2i, i},$$

$$i = 1, 2,$$

где $b_{i,1} = 1$ при $k = 1$, а при $k \geq 2$

$$b_{i,k} =$$

$$= \frac{1}{k-1} \sum_{j=1}^{k-1} \frac{(j(i+1) - (k-1))}{j+1} \times$$

$$\times b_{i, k-j}.$$

Для величин $c_{n,k}$ справедливы рекуррентные соотношения

$$c_{n, n-k} = (n-1) c_{n-1, n-k-1} +$$

$$+ (n - (2k-1)) c_{n-1, n-k}$$

$$0 \leq k \leq n-1, \quad c_{1,1} = 2, \quad c_{n-1, n} = 0$$

при $n \geq 2$.

Коэффициенты a_i :

$$a_1 = \pi/2,$$

$$a_2 = 0,37068870 \cdot 10^{-1}; \quad a_3 =$$

$$= -0,83209445 \cdot 10^{-3};$$

$$a_4 = -0,23232430 \cdot 10^{-3}; \quad a_5 =$$

$$= 0,12935888 \cdot 10^{-4};$$

$$a_6 = 0,29034835 \cdot 10^{-5}; \quad a_7 =$$

$$= -0,22096814 \cdot 10^{-6};$$

$$a_8 = -0,45415913 \cdot 10^{-7}; \quad a_9 =$$

$$= 0,40361657 \cdot 10^{-8};$$

$$a_{10} = 0,79875294 \cdot 10^{-9}.$$

При $0,03 \leq p \leq 0,97$ достигается точность 10^{-8} .

Функция распределения случайного вектора $(X, Y)'$ размера 2, имеющего нормальное распределение, в общем случае есть

$$\Phi_2(x, y; \mu_x, \mu_y, \sigma_x, \sigma_y, r) =$$

$$= P\{X \leq x, Y \leq y\} =$$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \times$$

$$\times \int_{-\infty}^x \int_{-\infty}^y \exp\left\{-\frac{1}{2(1-r^2)} \times$$

$$\times \left[\left(\frac{u-\mu_x}{\sigma_x}\right)^2 - 2r\left(\frac{u-\mu_x}{\sigma_x}\right) \times$$

$$\times \left(\frac{v-\mu_y}{\sigma_y}\right) + \left(\frac{v-\mu_y}{\sigma_y}\right)^2\right]\} du dv.$$

Здесь вектор математических ожиданий есть $E(X, Y) = (\mu_x, \mu_y)$, а ковариационная матрица —

$$= \begin{pmatrix} \sigma_x^2 & \sigma_x\sigma_y\rho \\ \sigma_x\sigma_y\rho & \sigma_y^2 \end{pmatrix}, \quad r = \sigma_x\sigma_y\rho -$$

коэффициент ковариации, ρ — коэффициент корреляции между X и Y .

После замены переменных x и y на $\frac{x-\mu_x}{\sigma_x}$ и $\frac{y-\mu_y}{\sigma_y}$ приходим к стандартной функции двумерного нормального распределения

$$\Phi_2(x, y; \rho) =$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \times$$

$$\times \int_{-\infty}^x \int_{-\infty}^y \exp\left\{-\frac{1}{2(1-\rho^2)} \times$$

$$\times (u^2 - 2\rho uv + v^2)\right\} du dv.$$

Соответствующую функцию плотности будем обозначать $\phi_2(x, y; \rho)$. Зная функцию $\Phi_2(x, y, \rho)$ в первом квадранте $\{x \geq 0, y \geq 0\}$, во всех остальных квадрантах можно вычислить ее по формулам

$$\Phi_2(x, y; \rho) = \Phi(x) - \Phi_2(x, -y; \rho),$$

$$\Phi_2(x, y; \rho) = \Phi(y) - \Phi_2(-x, y; \rho),$$

$$\Phi_2(x, y; \rho) = \Phi(x) + \Phi(y) - 1 - \Phi_2(-x, -y; \rho).$$

Двойной интеграл $\Phi_2(x, y; \rho)$ можно выразить через однократные интегралы, зависящие от двух параметров, простым образом связанных с параметрами x, y и ρ . Это упрощает задачу вычисления двумерного нормального распределения и его табулирования. Один из возможных методов выражения функции двумерного нормального распределения через однократные интегралы [7] приведен ниже.

Вводится функция, предложенная Оуэном,

$$T(h, a) = \frac{1}{2\pi} \int_0^a \frac{\exp\left\{-\frac{h}{2}(1+x^2)\right\}}{1+x^2} dx.$$

Для этой функции

$$T(h, a) + T(ah, 1/a) = 1/4 - [\Phi(h) - 1/2][\Phi(ah) - 1/2],$$

$$a > 0, h \geq 0.$$

Значения двумерного нормального распределения вычисляются по формуле

$$\Phi_2(x, y; \rho) = \frac{1}{2} \Phi(x) + \frac{1}{2} \Phi(y) - T(x, a_1) - T(y, a_2),$$

$$a_1 = \frac{y - \rho x}{x \sqrt{1 - \rho^2}},$$

$$a_2 = \frac{x - \rho y}{y \sqrt{1 - \rho^2}}, \quad x \geq 0, y \geq 0.$$

Алгоритм для вычисления вероятности попадания (X, Y) в произвольную многоугольную область на плоскости содержится в работе [7].

Разложение $\Phi_2(x, y; \rho)$ в ряд по тетрагорическим функциям имеет вид

$$\tau_k(x) = \frac{H_{k-1}(x) \Phi(x)}{(k!)^{1/2}},$$

$$k = 1, 2,$$

где $H_k(x)$ — k -й полином Эрмита.

Этот ряд есть

$$\Phi_2(x, y; \rho) = \Phi(x) \Phi(y) + \sum_{k=1}^{\infty} \rho^k \tau_k(x) \tau_k(y).$$

Ряд сходится равномерно по x и y при всех значениях $|\rho| \leq 1$. Для тетрагорических функций имеет место соотношение

$$\tau_k(x) = -\frac{1}{\sqrt{k}} \frac{d}{dx} \tau_{k-1}(x), \quad n \geq 1.$$

Для функции двумерного нормального распределения имеет место следующая декомпозиция:

$$\Phi_2(x, y; \rho) = \Phi_2\left(x, 0; \frac{(\rho x - y) \operatorname{sign} x}{(x^2 - 2\rho xy + y^2)^{1/2}}\right) + \Phi_2\left(y, 0; \frac{(\rho y - x) \operatorname{sign} y}{(x^2 - 2\rho xy + y^2)^{1/2}}\right) - d,$$

где $d = 0$, если истинно логическое выражение

$$(xy > 0) \vee [(xy = 0) \wedge (x + y \leq 0)],$$

и $d = 1/2$ в противном случае.

Для вычисления $\Phi_2(x, y; \rho)$ может быть также использована следующая формула:

$$\Phi_2(x, y; \rho) = \Phi(x) \Phi(y) + \int_0^{\rho} \Phi_2(x, y; \lambda) d\lambda.$$

Рассмотрим многомерное нормальное распределение. Случайный вектор $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ имеет нормальное распределение $N(\mu, \Sigma)$ с вектором математических ожиданий $\mu = (\mu_1, \dots, \mu_n) = E\mathbf{X}$ и ковариационной матрицей

$$\Sigma = E(\mathbf{X} - \mu)(\mathbf{X} - \mu)' = (r_{ij} : r_{ij} = \operatorname{cov}(X_i, X_j) = E(x_i - \mu_i)(x_j - \mu_j),$$

$$i, j = 1, \dots, n).$$

если его характеристическая функция есть

$$\begin{aligned} \psi(t) &= E e^{it'X} = \\ &= e^{it'\mu - \frac{1}{2} t' \Sigma t} \\ t &= (t_1, \dots, t_n). \end{aligned}$$

Получаемый в результате линейного преобразования вектор $Y = AX + B$, где A — матрица с n столбцами и n строками, а B — вектор размера m , имеет n -мерное нормальное распределение $N(A\mu + B, A\Sigma A')$.

В случае, если распределение $N(\mu, \Sigma)$ невырождено ($\det \Sigma \neq 0$), функция распределения X имеет вид

$$\Phi_n(x; \mu, \Sigma) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} =$$

$$= \int_{y < x} \varphi_n(y; \mu, \Sigma) dy,$$

$$x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$$

($y < x$ — означает, что неравенство выполняется по координатам), φ_n — функция плотности,

$$\begin{aligned} \varphi_n(x; \mu, \Sigma) &= \\ &= \frac{(\det A)^{1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2} (x - \mu)' \times \right. \end{aligned}$$

$$\left. \times A (x - \mu)\right\} =$$

$$= \frac{(\det A)^{1/2}}{(2\pi)^{n/2}} \times$$

$$\times \exp\left\{-\frac{1}{2} \sum_{i,j=1}^n a_{ij} \times \right.$$

$$\left. \times (x_i - \mu_i)(x_j - \mu_j)\right\},$$

причем $A = (a_{ij}) = \Sigma^{-1}$.

В случае вырожденного нормального распределения ($\det \Sigma = 0$) функция распределения $\Phi_n(x; \mu, \Sigma)$ может быть найдена следующим образом. Пусть m — ранг матрицы Σ . Распределение оказывается сосредоточенным на подпространстве размерности m .

Преобразуем случайный вектор X к $Y = Q(X - \mu)$, где Q — ортогональная матрица размера $n \times m$ ($QQ' = I$), такая, что матрица $Q\Sigma Q'$ имеет ненулевыми элементами только первые m диагональных элементов, равных соответственно $\lambda_1, \dots, \lambda_m$ — ненулевым характеристическим числам ковариационной матрицы Σ . В этом случае

$$\begin{aligned} \Phi_n(x; \mu, \Sigma) &= \\ &= \int_V \varphi_m(z; 0, Q\Sigma Q') dz, \end{aligned}$$

где область интегрирования V есть $V\{z = (z_1, \dots, z_m) : Q'(z_1, \dots, z_m,$

$0, 0, \dots, 0)' + \mu < x = (x_1, \dots, x_n)\}$. (Знак неравенства опять понимается по координатам.) Таким образом, область V в R^m задается пересечением n полупространств.

Приведем некоторые неравенства для многомерного нормального распределения. Рассмотрим сначала неравенство Слепяна. Пусть для двух ковариационных матриц $\Sigma_1 = (r_{ij}^{(1)})$ и $\Sigma_2 = (r_{ij}^{(2)})$ размера $n \times m$ $r_{ij}^{(1)} \geq r_{ij}^{(2)}$ при всех i и j и $r_{ii}^{(1)} = r_{ii}^{(2)}$. Тогда

$$\Phi_n(x; 0, \Sigma_1) \geq \Phi_n(x; 0, \Sigma_2).$$

Имеют место также следующие неравенства. Пусть случайный вектор $X = (X_1, \dots, X_n)$ имеет нормальное распределение $N(0, \Sigma)$ и разбит на два подвектора Y и Z , $X = (Y', Z')$ размера m и $n - m$ соответственно. Пусть A_Y и A_Z — выпуклые симметричные относительно нуля области в R^m и R^{n-m} соответственно. Тогда

$$\begin{aligned} P\{Y \in A_Y, Z \in A_Z\} &\geq \\ &\geq P\{Y \in A_Y\} P\{Z \in A_Z\}. \end{aligned}$$

Отсюда следует, например, что

$$P\{|X_i| \leq c_i, i = 1, \dots, n\} \geq$$

$$\geq \prod_{i=1}^n P\{|X_i| \leq c_i\}.$$

Рассмотрим последовательности независимых случайных векторов $\{X_i,$

$i = 1, \dots, k$, таких, что $X_i = (Y'_i, Z'_i)'$, где Y_i имеет размерность m , $Y_i \sim N(0, \Sigma_{YY})$, $i = 1, \dots, k$, $Z \sim N(0, \Sigma_{ZZ})$, $\Sigma_{YZ} = 0$.

Пусть $f(x)$ — измеримая функция: $R^n \rightarrow R^r$, A — измеримое множество в R^r и

$$\beta_k = P\{f(X_i) \in A, i = 1, \dots, k\}.$$

Тогда (см. [23]).

$$\beta_k \geq \beta_{k-1}^{k/k-1} \geq \beta_{k-2}^{k/k-2} \geq \dots \geq \beta_2^{k/2} \geq \dots \geq \beta_1^k + (\beta_2 - \beta_1^2)^{k/2} \geq \beta_1^k.$$

Пусть случайный вектор X представлен в виде $X = (Y', Z')'$, где Y и Z — векторы размеров соответственно m и n , $EY = \mu_Y$, $EZ = \mu_Z$, а ковариационная матрица Σ разбита соответствующим образом на подматрицы

$$\Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{YZ} & \Sigma_{ZZ} \end{pmatrix}.$$

Тогда условное распределение Y при заданном векторе $Z = z$ будет снова нормальным

$$N(\mu_Y + \Sigma_{YZ} \Sigma_{ZZ}^{-1}(z - \mu_Z), \Sigma_{YY} - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Sigma'_{YZ}).$$

Пусть D_Z — матрица, диагональные элементы которой совпадают с соответствующими элементами матрицы Σ , а остальные элементы равны нулю. Нормализующее преобразование при $\det \Sigma \neq 0$

$$Y = D_Z^{-1/2}(X - \mu) = \left(\frac{X_1 - \mu_1}{\sigma_1}, \frac{X_2 - \mu_2}{\sigma_2}, \dots, \dots, \frac{X_n - \mu_n}{\sigma_n} \right)$$

сводит задачу вычисления $\Phi_n(x; \mu, \Sigma)$ к задаче вычисления $\Phi_n(y; 0, R)$, где $y = D_Z^{-1/2}(x - \mu)$ и $R = D_Z^{-1/2} \Sigma D_Z^{-1/2}$ — корреляционная матрица, соответствующая ковариационной матрице Σ . В дальнейшем, если X — нормированный вектор, т. е. его ковариационная матрица совпадает с

корреляционной, функция распределения и функция плотности будут помечаться сверху индексом «0», например, $\Phi_n^0(y; \mu, R)$, $\varphi_n^0(y; \mu, R)$. Отметим, что

$$\frac{\partial \varphi_n^0(x; \mu, R)}{\partial \rho_{ij}} = \frac{\partial^2 \varphi_n(x; \mu, R)}{dx_i dx_j}.$$

Для вычисления значений функции трехмерного нормального распределения $\Phi_3^0(x; 0, R)$ в работе [24] получен следующий алгоритм, основанный на функции:

$$S(h, a, b) = \int_{-\infty}^h T(ay, b) \varphi(y) dy = \frac{b}{2\pi} \int_0^1 \frac{\Phi[h(1 + a^2 + a^2 b^2 y^2)^{1/2}]}{(1 + b^2 y^2)(1 + a^2 + a^2 b^2 y^2)^{1/2}} dy.$$

Этот алгоритм заключается в следующем:

а) если $x > 0$ или $x < 0$ по координатно, то

$$\Phi_3^0(x; 0, R) = \frac{1}{2} \sum_{i=1}^3 \left[\lambda \Phi(x_i) - \sum_{\substack{i=1 \\ j \neq i}}^3 (T(x_i, a_{ij}) + 2S(x_i, a_{ij}, b_{ij})) \right];$$

б) при $x_1 \geq 0, x_2 \geq 0, x_3 \leq 0$ или $x_1 \leq 0, x_2 \leq 0, x_3 \geq 0$ следует воспользоваться преобразованием

$$\Phi_3^0(x; 0, R) = \frac{1}{2} [\Phi(x_1) + \Phi(x_2) - \delta(x_1, x_2) - T(x_1, a_{12}) - T(x_3, a_{21}) - \Phi_3((x_1, x_2, -x_3)'; 0, \tilde{R})].$$

Здесь использованы следующие обозначения:

$$\tilde{R} = \begin{pmatrix} 1 & \rho_{12} & -\rho_{13} \\ \rho_{21} & 1 & -\rho_{23} \\ -\rho_{31} & -\rho_{32} & 1 \end{pmatrix},$$

$$\delta(u, v) = \begin{cases} 0, & \text{если } (\text{sign } u)(\text{sign } v) = 1, \\ 1 & \text{в противном случае,} \end{cases}$$

$$\lambda_i = (1 - \delta(a_{ij}, a_{ik})), \quad i = 1, 2, 3,$$

где $i \neq j \neq k$,

$$a_{ij} = \frac{x_j - x_i \rho_{ij}}{x_i (1 - \rho_{ij}^2)^{1/2}}, \quad i \neq j,$$

$$b_{ij} = b_{ijk} = [(1 - \rho_{ij}^2)(x_k - x_i \rho_{ik}) - (\rho_{jk} - \rho_{ij} \rho_{ik})(x_j - x_i \rho_{ij})] / [(x_j - x_i \rho_{ij})(\det R)^{1/2}],$$

где $\{i, j, k\}$ — перестановка $\{1, 2, 3\}$.

Интегрируя в выражении для $S(h, a, b)$ по частям, получаем, что

$$S(h, a, b) = \frac{1}{2\pi} \Phi[h(1 + a^2 + a^2 b^2)] \times \\ \times \operatorname{arctg} \frac{b}{\sqrt{1 + a^2 + a^2 b^2}} - \\ - \frac{ha^2 b^2}{4\pi} \int_0^1 \operatorname{arctg} \frac{by}{\sqrt{1 + a^2 b^2 y^2}} \times \\ \times e^{-h^2(1 + a^2 + a^2 b^2 y^2)/2} \frac{y dy}{\sqrt{1 + a^2 + a^2 b^2 y^2}}.$$

Таким образом, функция трехмерного нормального распределения, как и функция двумерного, выражается через однократные интегралы. Следовательно, функция четырехмерного нормального распределения выражается через двукратные интегралы.

Отметим некоторые частные случаи вычисления $\Phi_n^0(x; \mathbf{0}, \mathbf{R})$. Если элементы корреляционной матрицы \mathbf{R} могут быть представлены в виде

$$\rho_{ij} = \sum_{s=1}^k \beta_{is} \beta_{sj}, \quad i \neq j,$$

то

$$\Phi_n(x; \mathbf{0}, \mathbf{R}) = \\ = \int_{R^k} \prod_{i=1}^n \Phi \left(\frac{x_i - \sum_{j=1}^k \beta_{ij} y_j}{\left(1 - \sum_{s=1}^k \beta_{is}^2\right)^{1/2}} \right) \times \\ \times \varphi(y_1) \times \dots \times \varphi(y_k) dy_1 \dots dy_k.$$

Существуют также формулы для вычисления стандартной вероятности

первого квадранта $K_n(\mathbf{R}) = \Phi_n^0(0; \mathbf{0}, \mathbf{R})$ при $n = 2$ и $n = 3$. Это

$$K_2(\mathbf{R}) = \frac{1}{4} + \frac{1}{2\pi} \arcsin \rho_{12},$$

$$K_3(\mathbf{R}) = \frac{1}{4\pi} (2\pi - \arccos \rho_{12} - \arccos \rho_{13} - \arccos \rho_{23}).$$

При $n \geq 4$ элементарные формулы отсутствуют. В случае, если $\rho_{ij} = \rho$ при всех $i \neq j$, то (обозначая $P_n(\rho) = K_n(\mathbf{R})$)

$$P_n(\rho) = \\ = \int_{-\infty}^{\infty} \Phi^n \left(-\sqrt{\frac{\rho}{1-\rho}} y \right) \varphi(y) dy.$$

Эта формула справедлива при $-1/(n-1) < \rho < 1$. Имеет место также следующая формула:

$$P_4(\rho) = \frac{1}{16} + \frac{3}{4\pi} \arcsin \rho + \\ + \frac{3}{2\pi^2} \int_0^\rho \frac{\arcsin(x/(1+2x))}{\sqrt{1+x^2}} dx \approx \\ \approx \frac{1}{16} + \frac{3}{4\pi} \tau + \frac{1}{4\pi^2} \frac{\tau(3+5\tau)}{(1+\tau)(1+2\tau)},$$

где $\rho = \sin \tau$. Последнее приближение при $3 < 1/\rho < 12$ дает не менее пяти значащих цифр.

Плэкет [21] предложил так называемую формулу редукции, позволяющую свести вычисление функции n -мерного нормального распределения $\Phi_n^0(x; \mathbf{0}, \mathbf{R})$ к выражению, содержащему не более чем $(n-1)$ -кратные интегралы от некоторых функций. Эта формула может применяться рекуррентно до тех пор, пока порядок интегрирования не сократится примерно вдвое. Еще больше этот порядок можно сократить, применяя прием, изложенный в конце этого раздела.

Метод Плэкета заключается в следующем. Пусть имеется некоторая корреляционная матрица $\mathbf{K}_0 = (\kappa_{ij})$. Рассмотрим семейство матриц

$$\mathbf{K}_t = (\lambda_{ij}(t)) = t\mathbf{R} + (1-t)\mathbf{K}_0, \\ t \in [0, 1], \lambda_{ij}(0) = \kappa_{ij}, \lambda_{ij}(1) = \rho_{ij}.$$

Эти матрицы являются также корреляционными. Разобьем матрицу следующим образом:

$$K_t = \begin{pmatrix} K_t^{11} & K_t^{12} \\ K_t^{21} & K_t^{22} \end{pmatrix},$$

где K_t^{11} — ковариационная матрица второго порядка.

Аналогично разобьем матрицу $C_t = K_t^{-1}$:

$$C_t = \begin{pmatrix} C_t^{11} & C_t^{12} \\ C_t^{21} & C_t^{22} \end{pmatrix}.$$

Такого же типа обозначения будем использовать для разбиений матриц $R = K_1$ и $C = C_1$. Далее разобьем вектор $x = (x_1, x_2, \dots, x_n)'$ на две части $x = (x^{(1)'}, x^{(2)'})'$ и $x^{(1)} = (x_1, x_2)'$.

По формуле полного дифференциала имеем

$$\Phi_n^0(x; 0, R) = \Phi_n^0(x; 0, K_0) + \sum_{i>j} \int_{\kappa_{ij}}^{\rho_{ij}} \frac{\partial \Phi_n^0(x; 0, K_t)}{\partial \lambda_{ij}(t)} d\lambda_{ij}(t).$$

Функции $\frac{\partial \Phi_n^0(x; 0, K_t)}{\partial \lambda_{ij}}$ могут быть записаны так же, как и

$$\frac{\partial \Phi_n^0(x; 0, K_t)}{\partial \lambda_{12}} = \varphi_2((x_1, x_2)'; \lambda_{12}) \times \Phi_{n-2}(x^{(2)}; b, (C_t^{22})^{-1}),$$

где $b = K_t^{21} (K_t^{11})^{-1} x^{(1)}$.

Выберем матрицу K_0 так, чтобы она отличалась от матрицы R только одним элементом ρ_{12} (и соответственно ρ_{21}), который обозначим κ_{12} , причем $\det K_0 = 0$. В качестве κ_{12} можно взять одно из решений квадратного уравнения

$$K_0^{21} (R^{11})^{-1} K_0^{12} = 1.$$

Тогда

$$\Phi_n^0(x; 0, R) = \Phi_n^0(x; 0, K_0) +$$

$$+ \int_{\kappa_{12}}^{\rho_{12}} \varphi_2(x^{(1)}; \lambda) \times \Phi_{n-2}(x^{(2)}; b_\lambda, (C_\lambda^{22})^{-1}) d\lambda.$$

Здесь матрица C и вектор b зависят от λ . Таким образом, $\Phi_n^0(x; 0, K_0)$ выражается через невырожденные функции нормального распределения порядка не выше $n - 1$, а порядок интегрирования в одном оставшемся интеграле этой формулы равен $n - 1$.

Учитывая, что функция трехмерного нормального распределения выражается через однократные интегралы, заключаем, что функции распределения четырехмерного и пятимерного нормального распределения могут быть выражены через двукратные интегралы.

Проанализируем далее последнюю формулу. Приведем $\Phi_{n-2}(x^{(2)}; b_\lambda, (C_\lambda^{22})^{-1})$ к стандартному виду. Пусть это $\Phi_{n-2}^0(y_\lambda; 0, L_\lambda)$. Здесь элементы вектора $y_\lambda = (y_1^\lambda, \dots, y_{n-2}^\lambda)$ и корреляционной матрицы $L_\lambda = (l_{ij}^\lambda)$ размера $(n - 2) \times (n - 2)$ являются элементарными, хотя и громоздкими функциями λ . В интеграле

$$T = \int_{\kappa_{12}}^{\rho_{12}} \varphi_2(x^{(1)}; \lambda) \Phi_{n-2}^0(y_\lambda; 0, L_\lambda) d\lambda$$

проведем интегрирование по частям. Тогда

$$T = \Phi_{n-2}^0(y_{\rho_{12}}; 0, L_{\rho_{12}}) \times$$

$$\times \int_{\kappa_{12}}^{\rho_{12}} \varphi_2(x^{(1)}; \tau) d\tau -$$

$$- \int_{\kappa_{12}}^{\rho_{12}} \int_{\kappa_{12}}^{\lambda} \varphi_2(x^{(1)}; \tau) d\tau \times$$

$$\times \frac{d}{d\lambda} \Phi_{n-2}^0(y_\lambda; 0, L_\lambda) d\lambda.$$

Далее

$$\begin{aligned} & \frac{d}{d\lambda} \Phi_{n-2}^0(y_\lambda; 0, L_\lambda) = \\ & = \sum_{i=1}^{n-2} \frac{\partial \Phi_{n-2}^2(y_\lambda; 0, L_\lambda)}{\partial y_i^\lambda} \times \\ & \times \frac{dy_i^\lambda}{d\lambda} + \sum_{i < j \leq n-2} \frac{\partial \Phi_{n-2}^0(y_\lambda; 0, K_\lambda)}{\partial l_{ij}^\lambda} \times \\ & \times \frac{dl_{ij}^\lambda}{d\lambda}. \end{aligned}$$

Слагаемые в первой сумме выражаются через произведение одномерной функции плотности нормального распределения, элементарной функции $\frac{dy_i^\lambda}{d\lambda}$ и функций вида $\Phi_{n-3}(y_\lambda; a, \Lambda)$, где вектор a и ковариационная матрица Λ определяются из формулы для условной функции плотности. Слагаемые во второй сумме полученного выражения приводятся к произведению двумерной функции плотности, элементарных функций $dl_{ij}^\lambda/d\lambda$, $dy_i^\lambda/d\lambda$ и функций не более чем $(n - 4)$ -мерных нормальных распределений. Таким образом, функция шестимерного нормального распределения может быть выражена через некоторые двукратные интегралы (интегралы от произведений однократных интегралов) и однократные интегралы.

Применяя эту процедуру неоднократно (так, что при интегрировании по частям под знак дифференциала будут попадать кратные интегралы) совместно с описанной ранее процедурой Плэкета понижения порядка интегрирования в многомерном нормальном интеграле на единицу, можно показать, что максимальный порядок интегрирования в сумме интегралов, к которой может быть сведен многомерный нормальный интеграл, зависит от его размерности следующим образом:

Размерность многомерного нормального интеграла	Максимальный порядок интегралов
1, 2, 3	1
4, 5, 6	2
7, 8, 9, 10	3
11, 12, 13, 14, 15	4

Максимальный порядок интегралов, входящих в сумму, к которой может быть сведена функция n -мерного нормального распределения, выражается ближайшим сверху целым числом к величине

$$\frac{1}{2} (-3 + \sqrt{8n + 1}), \quad n > 1.$$

При возрастании n , к сожалению, очень быстро растет число интегралов в сумме, к которой приводится функция нормального распределения. Сравнительно легко реализовать этим методом вычисление значений функции нормального распределения размерности до 6.

Рассмотрим теперь задачу вычисления нормальной вероятности выпуклой многогранной области в R^n . Алгоритм вычисления нормальной вероятности многоугольной области в R^2 описан в работе [7]. Приводимый ниже алгоритм может быть реализован при умеренных размерности n и числе ограничивающих область плоскостей. При малых размерностях n многогранная область может служить достаточно хорошей аппроксимацией для других областей.

Пусть в n -мерном пространстве R^n задана многогранная область, образованная пересечением m полупространств A_1, A_2, \dots, A_m , причем каждое полупространство задано, например, следующим образом:

$$A_i = \{x = (x_1, \dots, x_n) \mid a_{i1}x_1 + \dots + a_{in}x_n + a_i \leq 0\}.$$

Итак, требуется найти нормальную вероятность

$$P\{A_1 A_2 \dots A_m\}.$$

Предполагаем, что распределение невырождено, а условия, задающие A_i , линейно независимы. Понадобится следующая формула:

$$\begin{aligned}
 P \{A_1 A_2 \dots A_m\} &= \\
 &= P \{A_1 \dots A_{i-1} A_{i+1} \dots A_m\} - \\
 &- P \{A_1 \dots A_{i-1} A_i^c A_{i+1} A \dots A_m\}.
 \end{aligned}
 \tag{1}$$

Рассмотрим сначала задачу вычисления вероятности многомерной области, образованной пересечением $n+1$ полупространств

$$A = A \cdot A_{n+1}.$$

Обозначим

$$\begin{aligned}
 L_i(x) &= a_{i1}x_1 + \dots + a_{in}x_n, \\
 L_i^*(x) &= L_i(x) + a_i, \quad i = 1, \dots, n+1.
 \end{aligned}$$

Тогда

$$\begin{aligned}
 A_i &= \{x = (x_1, \dots, x_n) : L_i^*(x) \leq 0\}, \\
 A_i^c &= \{x : L_i^*(x) > 0\}, \quad i = 1, \dots, n+1.
 \end{aligned}$$

Найдем такие константы β_1, \dots, β_n , что

$$\begin{aligned}
 L_{n+1}(x) &= \beta_1 L_1(x) + \dots + \beta_n L_n(x) \\
 &\text{(по крайней мере, одна из величин } \beta_1, \dots, \beta_n \text{ не равна нулю).}
 \end{aligned}$$

Тогда

$$L_{n+1}^* = \beta_1 L_1^* + \dots + \beta_n L_n^* + \beta_{n+1}, \tag{2}$$

где

$$\beta_{n+1} = -(\beta_1 a_1 + \dots + \beta_n a_n) + a_{n+1}.$$

Обозначим $A_i^0 = A_i$, $A_i^1 = A_i^c$.

Найдем величины r_1, r_2, \dots, r_{n+1} , принимающие значения 0 или 1, такие, что вероятность множества $A_1^{r_1} \dots \times A_{n+1}^{r_{n+1}}$ равна нулю. Положим

$$r_{n+1} = \begin{cases} 0, & \text{если } \beta_{n+1} \geq 0, \\ 1, & \text{если } \beta_{n+1} < 0. \end{cases}$$

Затем выберем

$$r_i = \begin{cases} 0, & \text{если } \beta_i \text{ sign } \beta_{n+1} \leq 0, \\ 1, & \text{если } \beta_i \text{ sign } \beta_{n+1} > 0, \\ i = 1, & \dots, n. \end{cases}$$

Здесь и ниже под $\text{sign } a$ подразумевается функция

$$\text{sign } a = \begin{cases} 1, & a \geq 0, \\ -1, & a < 0. \end{cases}$$

Представим $A_{n+1}^{r_{n+1}}$ теперь в виде

$$\begin{aligned}
 A_{n+1}^{r_{n+1}} &= \{x : \beta_1 \text{ sign } \beta_{n+1} L_1^*(x) + \\
 &+ \beta_n \text{ sign } \beta_{n+1} L_n^*(x) + \\
 &+ |\beta_{n+1}| \leq (<) 0\},
 \end{aligned}
 \tag{3}$$

причем при $r_{n+1} = 0$ в левой части используется знак « \leq », а при $r_{n+1} = 1$ — знак « $<$ ».

Выберем теперь произвольную точку x из R^n . Возможны два случая:

а) $x \in A_1^{r_1} \dots A_n^{r_n}$,

тогда $x \in A_1^{r_1} \dots A_{n+1}^{r_{n+1}}$,

б) $x \in A_1^{r_1} \dots A_n^{r_n}$.

Тогда в силу выбора r_i одновременно выполняются неравенства

$$\beta_i \text{ sign } \beta_{n+1} L_i^*(x) \geq 0, \quad i = 1, \dots, n.$$

Следовательно, x может принадлежать только границе множества $A_{n+1}^{r_{n+1}}$. Поэтому вследствие невырожденности распределения вероятность множества $A_1^{r_1} \dots A_{n+1}^{r_{n+1}}$ равна нулю.

Пусть множество $A_1^{r_1} \dots A_{n+1}^{r_{n+1}}$ отличается от $A_1 \dots A_{n+1}$ только тем, что $r_{v_1} = r_{v_2} = \dots = r_{v_s} = 1$, а остальные значения r_i равны нулю. В этом случае

$$A_{v_1}^{r_{v_1}} = A_{v_1}^c, \quad A_{v_s}^{r_{v_s}} = A_{v_s}^c$$

Используя формулу (1), сначала выражаем вероятность множества $A_1 \dots A_{n+1}$ через вероятности множества $A_1 \dots A_{v_1-1} A_{v_1+1} \dots A_{n+1}$, ограниченного n плоскостями, и множества $A_1 \dots A_{v_1-1} A_{v_1}^c A_{v_1+1} \dots A_{n+1}$.

Вероятность первого множества после соответствующих преобразований, описанных ниже, выражается

через функцию n -мерного нормального распределения.

Вероятность второго множества преобразуем снова таким же образом. Теперь она выражается через вероятность множества

$$A_1 \cdot A_{v_1-1} A_{v_1}^c A_{v_1+1} \times \\ \times A_{v_2-1} A_{v_2+1} \cdot A_{n+1},$$

ограниченного n плоскостями, и множества

$$A_1 \cdot A_{v_1-1} A_{v_1}^c A_{v_1+1} \times \\ \times A_{v_2-1} A_{v_2}^c A_{v_2+1} \cdot A_{n+1}.$$

Поступая таким же образом и далее, выразим вероятность множества $A_1 \times \dots \times A_{n+1}$ через вероятности s множеств, ограниченных n плоскостями, и вероятность множества

$$A_1^{r_1} \cdot A_{n+1}^{r_{n+1}}$$

Вероятность же этого множества равна нулю вследствие выбора r_1, \dots, r_{n+1} .

Рассмотрим теперь задачу вычисления вероятности многогранной области с произвольным числом m ограничивающих плоскостей. Пусть это множество

$$A_1 \cdot A_m, \quad m > n + 1.$$

Выделим первые A_1, \dots, A_{n+1} полуплоскостей и расставим в соответствии с только что описанным алгоритмом знаки дополнений так, что множество $A_1^{r_1} \cdot \dots \cdot A_{n+1}^{r_{n+1}}$ оказывается пустым (или содержит только некоторые границы). Тогда множество

$$A_1^{r_1} \cdot A_{n+1}^{r_{n+1}} A_{n+2} \cdot A_m$$

будет таким же.

Используя формулу (1), выразим вероятность $A_1 \cdot \dots \cdot A_m$ через вероятности множеств, образованных $m - 1$ полуплоскостями. Эту процедуру проводим рекуррентно до тех пор, пока искомая вероятность не будет выражена через вероятности множеств, образованных n плоскостями.

Если некоторые границы оказываются параллельными, то это может быть

использовано для упрощения алгоритма следующим образом.

Пусть полупространства A_1, \dots, A_m таковы, что среди них имеется k полупространств $A_{\tau_1}, \dots, A_{\tau_k}$, таких, что их границы параллельны. Разобьем их на два класса A_1^*, \dots, A_s^* и A_{s+1}^*, \dots, A_k^* , таких, что $A_1^* \subset A_2^* \subset \dots \subset A_s^*$ и $A_{s+1}^* \subset A_{s+2}^* \subset \dots \subset A_k^*$, причем A_1^* не содержится в A_{s+1}^* и наоборот. Очевидно, что

$$A_1^* \cdot A_s^* = A_1^* \\ \text{и} \\ A_{s+1}^* \cdot A_k^* = A_{s+1}^*.$$

Таким образом, количество полуплоскостей в $A_1 \cdot \dots \cdot A_m$ сокращается на $k - 1$ или $k - 2$.

В этом случае, если один из классов отсутствует, переходим к поиску других параллельных между собой границ, а если присутствуют оба класса, то анализ продолжается следующим образом.

Перегруппируем $A_1 \cdot \dots \cdot A_m$ так, что A_1^* перейдет в A_1 , а A_{s+1}^* — в A_2 , и выбросим лишние полупространства. Тогда задача сводится к нахождению вероятности множества

$$A_1 A_2 \cdot A_{m-k+2}.$$

Вновь преобразованные полупространства A_1 и A_2 не содержатся друг в друге, и пусть их пересечение не пусто иначе задача решена. Тогда

$$P(A_1 A_2 A_3 \cdot \dots \cdot A_{m-k+2}) = \\ = P(A_1 A_3 \cdot \dots \cdot A_{m-k+2}) - \\ - P(A_2^c A_3 \cdot \dots \cdot A_{m-k+2}).$$

Нормальная вероятность области (конуса) $A_1 A_2 \cdot \dots \cdot A_n$ в R^n может быть записана в виде

$$\int_{Ax+a \leq 0} \varphi_n(x; m, \Sigma) dx,$$

где $A = (a_{ij})$ (как и ранее, векторное неравенство понимается покомпонентно). При замене x на $A^{-1}y$ задача сводится к вычислению значения

$$\Phi_n(-a; A^{-1}m, A^{-1}\Sigma(A^{-1})').$$

Это преобразование возможно, так как ранее предполагалось, что ранг матрицы A равен n .

3. РАСПРЕДЕЛЕНИЕ χ^2 И ЕГО ОБОБЩЕНИЯ

Случайная величина X имеет (центральное) распределение χ^2 с n степенями свободы ($X \sim \chi_n^2$), если ее функция плотности

$$f_n(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Такое распределение имеет случайная величина, которая может быть представлена в виде суммы квадратов n ($n = 1, 2, \dots$) независимых случайных величин X_1, X_2, \dots, X_n , имеющих стандартное нормальное распределение $N(0, 1)$,

$$X = X_1^2 + X_2^2 + \dots + X_n^2.$$

Начальные моменты X

$$EX^k = n(n+2) \cdot \dots \cdot (n+2(k-1)),$$

а характеристическая функция

$$\Psi_n(t) = (1 - 2it)^{-n/2}.$$

Нецентральное распределение χ^2 с n степенями свободы и параметром нецентральности δ^2 совпадает с распределением случайной величины $X = X_1^2 + X_2^2 + \dots + X_n^2$, где $X_i, i = 1, 2, \dots, n$, независимы между собой и $X_i \sim N(\delta_i, 1)$, причем $\delta^2 = \delta_1^2 + \delta_2^2 + \dots + \delta_n^2$. Распределение X при этом зависит только от δ^2 , но не от конкретных значений $\delta_1, \delta_2, \dots, \delta_n$. Его функция плотности $f_n(x; \delta^2)$ может быть выражена через функции плотности центрального распределения χ^2 так:

$$f_n(x; \delta^2) = \begin{cases} e^{-\delta^2/2} \sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{\delta^2}{2}\right)^i f_{n+2i}(x), & \text{если } x > 0, \\ 0, & \text{если } x \leq 0, \end{cases}$$

а характеристическая функция

$$\Psi_n(t; \delta^2) = (1 - 2it)^{-n/2} \times \exp\left\{\frac{i\delta^2 t}{1 - 2it}\right\}.$$

Преобразование $\xi = 2\lambda X$ переводит случайную величину X , имеющую распределение χ^2 (центральное или нецентральное), в случайную величину ξ , имеющую гамма-распределение (соответственно центральное или нецентральное) с параметрами λ и $\mu = n/2$ и параметром нецентральности $\gamma = \delta^2/2$. Здесь параметр λ — это параметр масштаба, а параметр μ — это параметр формы. Параметр μ в гамма-распределении может быть любым действительным числом большим нуля. Функция плотности (центрального) гамма-распределения

$$f(x; \mu, \lambda) = \begin{cases} \frac{\lambda^\mu x^{\mu-1} e^{-\lambda x}}{\Gamma(\mu)}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Характеристическая функция нецентрального гамма-распределения

$$\Psi(t; \mu, \lambda, \gamma) = \left(1 - \frac{it}{\lambda}\right)^\mu \times \exp\left(\frac{it\gamma}{\lambda - it}\right).$$

Для функции распределения $F_n(x)$ центрального распределения χ^2 имеет место следующая рекуррентная формула:

$$F_n(x) = F_{n-2}(x) - 2f_n(x), \quad x > 0.$$

Если n четное, то вычисления начинаются с

$$F_2(x) = 1 - e^{-x/2},$$

а при нечетном n с

$$F_1(x) = 2\Phi(\sqrt{x}) - 1.$$

При четных n , таким образом,

$$F_n(x) = 1 - \sum_{i=0}^{n/2} \frac{(x/2)^i}{i!} e^{-x/2}.$$

Несмотря на простоту этой формулы, вычисления по ней при больших значениях n , а следовательно, и x затруднительны.

Рассмотрим методы вычисления функции центрального гамма-распределения $G(x; \mu, \lambda)$. Для стандартного гамма-распределения $G(x; \mu, 1)$ (отметим, что $\Gamma(\mu) = G(x, \mu, 1)$ — неполная гамма-функция) имеют место следующие разложения в ряд:

$$G(x; \mu, 1) = \frac{e^{-x} x^\mu}{\Gamma(\mu + 1)} \times \left[1 + \sum_{i=1}^{\infty} \frac{x^i}{(\mu + 1)(\mu + 2) \dots (\mu + i)} \right]$$

и в непрерывную дробь:

$$G(x; \mu, 1) = 1 - \frac{e^{-x} x^\mu}{\Gamma(\mu)} \times \left(\frac{1}{x+1} \frac{1-a}{1+a} \frac{1}{x+1} \frac{2-a}{1+a} \frac{2}{x+1} \dots \right).$$

При $a \leq x \leq 1$ или $x < a$ вычисления в разложении в ряд прекращаются, когда очередное слагаемое становится меньше заданной ошибки вычислений. При использовании непрерывной дроби вычисления прекращаются, когда разность между двумя последовательными частными дробями становится меньше заданной ошибки.

Другие непрерывные дроби для гамма-распределения можно найти в работе [15].

Функция нецентрального χ^2 -распределения при четном числе степеней свободы n выражается формулой

$$F_n(x; \delta^2) = \sum_{i=0}^{\infty} \sum_{j=i+n/2}^{\infty} \frac{\left(\frac{\delta^2}{2}\right)^i \left(\frac{x}{2}\right)^j}{i! j!} \times e^{-\frac{x+\delta^2}{2}} \text{ при } x > 0.$$

При $n = 2k + 2$ (см. [12])

$$F_n(x; \delta^2) = F_2(x; \delta^2) + \sum_{i=1}^k C_k^i 2^i F_2^{(i)}(x; \delta^2),$$

причем

$$F_2(x; \delta^2) = \frac{1}{2\pi} \int_{-\sqrt{x}}^{\sqrt{x}} \int_{-\sqrt{x-x_1^2-\delta^2}}^{\sqrt{x-x_1^2-\delta^2}} e^{-\frac{x_1^2+x_2^2}{2}} \times dx_1 dx_2,$$

а j -я производная от $F_2(x; \delta^2)$ есть

$$F_2^{(j)}(x; \delta^2) = - \sum_{i=1}^{\infty} (2^{i+j-1} (i+j-1)!)^{-2} \times x^{i+j-1} i(i+1) \dots (i+j-1) \times (\delta^2)^{i-1} \exp\left[-\frac{1}{2}(\delta^2+x)\right] - \sum_{m=1}^{j-1} C_{k-1}^{j-1} 2^{m-j} F_2^{(m)}(x; \delta^2),$$

$j = 2, 3, \dots, k$.

В общем случае для функции распределения нецентрального χ^2 -распределения имеет место формула [20]

$$F_n(x; \delta^2) = \int_0^x \frac{e^{-y/2} e^{-\delta^2/2}}{2^{n/2}} \times \sum_{i=0}^{\infty} \frac{y^{(n/2)+i-1} (\delta^2)^i}{\Gamma[(n/2)+i] 2^{2i} i!} dy.$$

$F_n(x; \delta^2)$ может быть сравнительно легко вычислена путем численного интегрирования по формуле (см. [5])

$$F_n(x; \delta^2) = \frac{1}{2} - \frac{1}{2\pi} \int_0^\pi \frac{1}{\sin \frac{\varphi}{2}} \times \exp\left\{-\frac{x+\delta^2}{2} (1-\cos \varphi)\right\} \times \sin\left(\frac{n-1}{2} \varphi - \frac{x-\delta^2}{2} \sin \varphi\right) \times d\varphi + R_n(x; \delta^2),$$

где

$$R_n(x; \delta^2) = \frac{(-1)^{[n/2]}}{\pi} \times \int_2^\infty \frac{1}{\lambda(\lambda-1)^{n/2}} \times \exp\left\{-\frac{\lambda}{2}\left(x + \frac{\delta^2}{\lambda-1}\right)\right\} d\lambda$$

при нечетных n и $R_n(x, \delta^2) = 0$ при четных n . При $n \rightarrow \infty$ $R_n(x, \delta^2)$ быстро стремится к нулю в области значений x , примаыкающих к математическому ожиданию, т. е. к $n + \delta^2$.

Рассмотрим теперь задачу вычисления функции распределения линейной комбинации

$$Q = \sum_{i=1}^r \alpha_i \chi_{s_i}^2, \delta_i^2, i$$

независимых случайных величин $\chi_{s_i}^2, \delta_i^2, i$, имеющих распределение χ^2

с числами степеней свободы s_i и параметрами нецентральности $\delta_i^2, i = 1, 2, \dots, k$. (Коэффициенты $\{\alpha_i\}$ могут быть как положительными, так и отрицательными.)

Математическое ожидание Q есть

$$EQ = \sum_{i=1}^r s_i \alpha_i (1 + \delta_i^2),$$

а дисперсия

$$DQ = \sum_{i=1}^r s_i \alpha_i^2 (2 + 4\delta_i^2).$$

Имеет место следующая формула для функции распределения Q (см. [4, 5]):

$$P(Q < x) = 1 - \frac{1}{\pi} \arctg \frac{1}{a} - \frac{1}{\pi} \int_0^\infty \frac{e^{-atx/2} \sin(\theta(t, x))}{\rho(t)} dt,$$

где

$$\theta(t, x) = \sum_{k=1}^r \left\{ \frac{s_k}{2} \omega_k(t) + \frac{\lambda_k \delta_k^2 t}{2[(\lambda_k - at)^2 + t^2]} \right\} - \frac{tx}{2},$$

$$\rho(t) = t \prod_{k=1}^r \left\{ \left[\left(1 - \frac{at}{\lambda_k}\right)^2 + \frac{t^2}{\lambda_k^2} \right]^{s_k/4} \times \exp\left\{-\frac{\delta_k^2 t [a\lambda_k - (a^2 + 1)t]}{2[(\lambda_k - at)^2 + t^2]}\right\}\right\},$$

$$\omega_k(t) = \arctg \frac{t}{\lambda_k - at}, \quad \lambda_k = \frac{1}{\alpha_k},$$

$k = 1, 2, \dots, r$.

Здесь параметр a — произвольное неотрицательное число. При $a = 0$ эта формула переходит в формулу Имгофа [16]. При значениях a , отличных от нуля, появляющаяся под знаком интеграла экспонента сокращает количество значимых колебаний подынтегральной функции. На практике оказывается удобный выбор $a = 2,4$. Интегрирование может быть проведено по формуле Симпсона. Указанная формула годится в частном случае и для вычисления как распределения χ^2 -распределенных случайных величин, так и после соответствующего преобразования параметров в Q для вычисления распределения гамма-распределенных случайных величин и их линейных комбинаций.

Простейшие аппроксимации для квантилей $\chi^2(p)$ центрального χ^2 -распределения есть аппроксимация Фишера.

$$\chi^2(p) \approx \frac{1}{2} \{u(p) + \sqrt{2n-1}\}^2$$

и Вильсона — Хильферти

$$\chi^2(p) = n \left\{ u(p) \sqrt{\frac{2}{9n} + 1} - \frac{2}{9n} \right\},$$

где $u(p)$ — квантиль уровня p стандартного нормального распределения. Более точные и более строгие аппроксимации содержатся, например, в работах [11, 27].

4. T-РАСПРЕДЕЛЕНИЯ СТЬЮДЕНТА

Случайная величина X имеет (центральное) t -распределение (распреде-

ление Стьюдента) с n степенями свободы, если ее функция плотности

$$f_n(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2) \sqrt{\pi n}} \times \\ \times \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \\ -\infty < x < \infty.$$

Таким распределением обладает случайная величина

$$X = \frac{\xi}{\sqrt{\eta/n}},$$

где случайная величина ξ имеет стандартное нормальное распределение $N(0, 1)$, η — независимая от ξ случайная величина с распределением χ^2 с n степенями свободы. Моменты распределения Стьюдента существуют только при $\mu_{2s+1} = 0$, а

$$\mu_{2s} = \frac{n^s \Gamma(s+1/2) \Gamma(n/2 - s)}{\Gamma(1/2) \Gamma(n/2)},$$

$$2s < n,$$

$$DX = \frac{n}{n-2} \quad \text{при } n > 2.$$

Характеристическая функция t -распределения при нечетных n есть

$$\psi(t) = \frac{\sqrt{\pi} \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \times \\ \times \frac{e^{-\sqrt{n}|t|}}{2^{2(n-1)} (n-1)!} \times \\ \times \sum_{k=0}^{n-1} (2k!) C_{n-1+k}^{2k} (2\sqrt{n}|t|)^{n-1-k},$$

В случае, если $\xi \sim N(m, 1)$, $m \neq 0$, то говорят, что X имеет нецентральное t -распределение с параметром нецентральности m , а если еще и η имеет нецентральное χ^2 -распределение с n степенями свободы и параметром нецентральности δ^2 , то такое t -распределение называют дважды нецентральным.

Функция плотности нецентрального t -распределения (с одним параметром нецентральности m)

$$f_n(x; m) =$$

$$\frac{\exp[-(m^2/(2(n+x^2)))] \times \\ \times [n/(n+x^2)]^{(n+1)/2}}{2^{(n-1)/2} \Gamma(n+1) \Gamma(1/2) n^{1/2}} \times \\ \times \int_0^{\infty} z^n \exp[-(z - \\ - mx(m+x^2)^{-1/2})^2/2] dz.$$

Для функции центрального t -распределения $F_n(x)$ справедливы формулы: при $n = 1$

$$F_1(x) = \frac{1}{\pi} \arctg(x/\sqrt{n}) + 1/2;$$

при n нечетном (обозначая $\theta = \arctg(x/\sqrt{n})$)

$$F_n(x) = \frac{1}{\pi} \left[\theta + \sin \theta \left(\cos \theta + \right. \right. \\ \left. \left. + \frac{2}{3} \cos^3 \theta + \right. \right. \\ \left. \left. + \frac{2 \cdot 4 \cdot \dots \cdot (n-3)}{1 \cdot 3 \cdot \dots \cdot (n-2)} \cos^{n-2} \theta \right) \right] + 1/2,$$

а при n четном

$$F_n(x) = \frac{1}{2} \sin \theta \left[1 + \frac{1}{2} \cos^2 \theta + \right. \\ \left. + \frac{1 \cdot 3}{2 \cdot 4} \cos^4 \theta + \right. \\ \left. + \frac{1 \cdot 3 \cdot \dots \cdot (n-3)}{2 \cdot 4 \cdot \dots \cdot (n-2)} \cos^{n-2} \theta \right] + 1/2.$$

Имеет место также следующее разложение:

$$Q_n(x) = 2[1 - F_n(x)] = \\ = 2f_n(x) \frac{\sqrt{n}}{t} \left[\frac{1}{n} + \frac{t^2}{2(n+2)} + \right. \\ \left. + \frac{3t^4}{2 \cdot 4(n+4)} + \frac{5t^6}{16(n+6)} + \dots \right],$$

$f_n(x)$ — функция плотности t -распределения, $t = \left(1 + \frac{x^2}{n}\right)^{-1}$

Функция t -распределения выражается также через бета-распределение $I_x(a, b)$

$$F_n(x) = 1 - \frac{1}{2} I_t\left(\frac{n}{2}, \frac{1}{2}\right),$$

$$t = \left(1 + \frac{x^2}{n}\right)^{-1}$$

Квадрат t -распределенной случайной величины имеет F -распределение, так что центральное t -распределение (в отличие от нецентрального) может быть вычислено с использованием алгоритмов для F -распределения.

Для вычисления функции нецентрального t -распределения $F_n(x; m)$ может использоваться следующая формула:

$$T_n(x; m) = \frac{\sqrt{2n}}{Q(n/2)} \times \int_0^\infty \frac{\Phi(xz - m)}{z} \lambda^n(z) dz,$$

где

$$Q(x) = \Gamma(x+1)/(x^x e^{-x} \sqrt{x}) = \sqrt{2\pi} \left(1 + \frac{1}{12x} + \frac{1}{288x^2} + \dots\right),$$

$$\lambda(z) = z \exp\{-(z^2 - 1)/2\},$$

$\Phi(x)$ — функция стандартного нормального распределения. Численное интегрирование в этой формуле проводится при больших и малых значениях n в окрестности точки $z = 1$.

Функция дважды нецентрального t -распределения с параметрами m и δ^2 может быть вычислена по формуле [10]

$$F_n(x; m, \delta^2) = \int_0^\infty \Phi(x\sqrt{y/n} - m) g(y) dy,$$

$$g(y) = \frac{1}{2} \left(\frac{y}{\delta^2}\right)^{\frac{n-2}{4}} \times e^{-\frac{(y+\delta^2)}{2}} I_{\frac{n-2}{2}}(\sqrt{\delta^2 y}),$$

$$I_\gamma(z) = \sum_{m=0}^\infty \frac{(z/2)^{\gamma+2m}}{m! \Gamma(\gamma+m+1)}.$$

Рассмотрим методы вычисления квантилей t -распределения. Можно получить (см. [18]), что первые члены разложения Корниша—Фишера для квантилей t -распределения есть

$$u(p) = u(p_1) + \frac{p - p_1}{f_n(u(p_1))} + \frac{\kappa(u(p_1))(p - p_1)^2}{2f_n^2(u(p_1))} + \frac{[2\kappa^2(u(p_1)) + \kappa'(u(p_1))](p - p_1)^3}{6f_n^3(u(p_1))},$$

где $\kappa(t) = \frac{t(n+1)}{(n+t^2)}$.

Приближенные значения для квантилей t -распределения, необходимые для последующих итераций, могут быть найдены в работах [17, 18]. Например, для $t_{n, \alpha}$ -квантилей t -распределения уровня α получена формула

$$t_{n, \alpha} \approx 1/(-0,0953 - 0,631/(n+1) + 0,81/\sqrt{-\ln(4\alpha(1-\alpha))} + 0,76(4\alpha\sqrt{n})^{1/n}).$$

5. F-РАСПРЕДЕЛЕНИЕ

Случайная величина X имеет F -распределение с n_1 и n_2 степенями свободы, если ее функция плотности есть

$$f_{n_1, n_2}(x) = \frac{\Gamma((n_1+n_2)/2)}{\Gamma(n_1/2)\Gamma(n_2/2)} \times n_1^{n_1/2} n_2^{n_2/2} x^{n_1/2-1} \times (n_1+n_2x)^{-(n_1+n_2)/2}, \quad x > 0;$$

Таким распределением обладает случайная величина

$$X = \frac{\chi_{n_1, 1}^2/n_1}{\chi_{n_2, 2}^2/n_2},$$

где $\chi_{n_i, i}$, $i = 1, 2$, — независимые случайные величины, имеющие распределение χ^2 соответственно с n_1 и n_2 степенями свободы.

Начальные моменты X находят по формуле

$$EX^k = \frac{\Gamma\left(\frac{n_1}{2} + k\right) \Gamma\left(\frac{n_2}{2} - k\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} \times \left(\frac{n_2}{n_1}\right)^k \quad \text{при } 2k < n_2,$$

а математическое ожидание и дисперсию по формулам

$$MX = \frac{n_1}{n_2 - 2}, \quad n_2 > 2, \quad \text{и } DX = \frac{2n_1^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}, \quad n_2 > 4.$$

Если $\chi_{n_1, \delta_1, 1}^2$ и $\chi_{n_2, \delta_2, 2}^2$ независимые случайные величины, имеющие нецентральное χ^2 -распределение с параметрами нецентральности δ_1^2 и δ_2^2 , то случайная величина

$$X = \frac{\chi_{n_1, \delta_1, 1}^2/n_1}{\chi_{n_2, \delta_2, 2}^2/n_2}$$

имеет дважды нецентральное F -распределение, а если $\delta^2 = 0$, то распределение называют нецентральным F -распределением. Обозначим его функцию распределения через $F_{n_1, n_2}(x; \delta_1^2, \delta_2^2)$.

Имеет место формула

$$F_{n_1, n_2}(x; \delta_1^2, \delta_2^2) = 1 - F_{n_2, n_1}\left(\frac{1}{x}, \delta_2^2, \delta_1^2\right), \quad x > 0.$$

В случае, если n_2 — четное число, то функция центрального распределения

$$F_{n_1, n_2}(x) = (1 - y)^{n_1/2} \times \left[1 + \frac{n_1}{2} y + \frac{n_1(n_1 + 2)}{2 \cdot 4} y^2 + \dots + \frac{n_1(n_1 + 2) \dots (n_1 + n_2 - 4)}{2 \cdot 4 \dots (n_2 - 2)} \times y^{n_2/2 - 1} \right],$$

где $y = n_2/(n_2 + n_1 x)$.

Если n_1 и n_2 — любые целые числа, то (см. [13, 18])

$$F_{n_1, n_2}(x) = \frac{2\Gamma\left(\frac{n_1 + n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} \times \int_0^{\arcsin \sqrt{y}} \sin^{n_1 - 1} \theta \cos^{n_2 - 1} \theta d\theta.$$

Функция F -распределения следующим образом связана с функцией бета-распределения $I_z(a, b)$:

$$F_{n_1, n_2}(x) = 1 - I_y\left(\frac{n_2}{2}, \frac{n_1}{2}\right),$$

где $y = \frac{n_2}{n_2 + n_1 x}$.

Аппроксимации для функции бета-распределения и ее квантилей могут быть перенесены и на F -распределение и наоборот [1, 18].

Функция дважды нецентрального F -распределения может быть вычислена по формуле [25]

$$F_{n_1, n_2}(x; \delta_1^2, \delta_2^2) = 1 - \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{(-\delta_1^2/2)^r (-\delta_2^2/2)^s}{r! s!} \times \left\{ \sum_{i=0}^r \sum_{j=0}^s (-1)^{i+j} C_r^i C_s^j \times I_y\left(\frac{n_2}{2} + i, \frac{n_1}{2} + j\right) \right\},$$

где $y = 1/(1 + n_1 x/n_2)$; $I_y(a, b)$ — бета-распределение.

Другая формула имеет следующий вид [5]:

$$F_{n_1, n_2}(x; \delta_1^2, \delta_2^2) = \frac{1}{2} - \frac{1}{\pi} \times \int_0^{\infty} \frac{\exp\left\{-\frac{\delta_1^2 t^2}{2(1+t^2)} - \frac{\delta_2^2 y^2 t^2}{2(1+y^2 t^2)}\right\}}{t(1+t^2)^{n_1/4} (1+y^2 t^2)^{n_2/4}} \times \sin\left[\frac{n_1}{4} \arctg t - \frac{n_2}{4} \arctg(yt) + \frac{\delta_1^2 t}{2(1+t^2)} - \frac{\delta_2^2 yt}{2(1+y^2 t^2)}\right] dt,$$

$y = n_1 x/n_2$.

Преобразование $s = t^5$ под знаком интеграла улучшает процедуру вычисления $F(x)$ по методу Симпсона.

Другие методы вычисления функции нецентрального F -распределения содержатся в работах [9, 14 и 19].

Приведем программу, реализующую эту формулу:

```

FUNCTION FS(X,N1,N2,D1,D2,N)
Y=N1 * X/N2
S=Y * Y
R=(N1+N2) * 2.5
H=EXP(-(ALOG(R * 1.E-07)
 * +N2 * ALOG(Y)/2)/R)/N
F=0
DO I=1,N
T=I * H
P=T * * 5
Q=P * P
R=1+Q
Z=1+S * Q
1 F=F+EXP(-(D1/R+S * D2/Z) *
 * Q/2-(N2 * ALOG(Z)+N1 *
 * ALOG(R))/4) * SIN((N1
 * * ATAN(P)-N2 * ATAN(Y * P)+
 * (D1/R-Y * D2/Z) * P)/2)/T *
 * (4+2 * (I-I/2 * 2-1))
FS=.5-F * H * .5305165
RETURN
END
    
```

Формальным параметрам этой процедуры-функции $X, N1, N2, D1$ и $D2$ соответствуют x, n_1, n_2, δ_1^2 и δ_2^2 , а параметр N — число разбиений в формуле Симпсона, деленное пополам. Можно выбирать $N = 50-100$. Точность контролируется увеличением N в 2 раза и повторным просчетом.

6. БЕТА-РАСПРЕДЕЛЕНИЕ

Функция бета-распределения имеет вид

$$B(u; a, b) = \frac{1}{B(a, b)} \int_0^u t^{a-1} (1-t)^{b-1} dt,$$

$$0 \leq u \leq 1,$$

$$B(a, b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)},$$

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt -$$

гамма-функция Эйлера. Имеет место тождество

$$B(u; a, b) = 1 - B(1-u; b, a), \quad (4)$$

поэтому можно табулировать лишь случай $0 < a \leq b$. Справедливы рекуррентные соотношения

$$B(u; a, b) = \frac{u^a (1-u)^{b-1}}{aB(a, b)} + B(u; a+1, b-1) \quad (5)$$

$$B(u; a, b) = u B(u; a-1, b) + (1-u) B(u; a, b-1), \quad (6)$$

с помощью которых можно рассматривать лишь те значения a и b , для которых $B(u; a, b)$ легко находится интегрированием. Если $a = 1$, то

$$B(u; 1, b) = 1 - (1-u)^b,$$

$$0 \leq u \leq 1;$$

если $b = 1$, то

$$B(u; a, 1) = 1 - u^a, \quad 0 \leq u \leq 1.$$

Если $a \neq 1, b \neq 1$, то для вычисления $B(u; a, b)$ можно использовать ее разложение в ряд Тейлора в окрестности нуля

$$B(u; a, b) = \frac{u^a}{aB(a, b)} (1+u) \times \sum_{n=1}^{\infty} \frac{u^n (1-b)(2-b) \dots (n-b)}{(n-a)n!}. \quad (7)$$

В формуле (7) можно считать, что $u \in \left[0, \frac{1}{2}\right]$, а параметр $b < 1$. Этого всегда можно добиться, используя тождества (4)–(6). При $u \leq \frac{1}{2}$ и $b < 1$ ряд (7) сходится быстро. Первые 10 членов разложения гарантируют относительную погрешность $\delta < 10^{-4}$, а 14 членов — относительную погрешность $\delta < 10^{-5}$.

Функция $B(u; a, b)$ связана с $F(u; \nu_1, \nu_2)$ — ф. р. Фишера с (ν_1, ν_2) — степенями свободы

$$F(u; \nu_1, \nu_2) = 1 - B\left(\frac{\nu_2}{\nu_2 + \nu_1 u}; \frac{\nu_1}{2}, \frac{\nu_2}{2}\right).$$

Для вычисления квантилей $V^{-1}(\alpha; a, b)$ функции бета-распределения можно использовать соответствующие таблицы и программы для F -распределения. Полезна следующая аппроксимация:

$$V^{-1}(\alpha; a, b) \approx \frac{a}{a + b \exp\{2\omega\}}, \quad (8)$$

где

$$\left\{ \begin{aligned} \omega &= \frac{u_\alpha (h + \lambda)^{1/2}}{h} - \left(\frac{1}{2b-1} - \frac{1}{2a-1} \right) \left(\lambda + \frac{5}{6} - \frac{3}{h} \right); \\ h &= 2 \left(\frac{1}{2a-1} + \frac{1}{2b-1} \right)^{-1}, \\ \lambda &= \frac{u_\alpha^2 - 3}{6}, \end{aligned} \right.$$

где u_α — квантиль уровня α .

Если a или b близко к $1/2$, то (8) дает значительную погрешность.

7. ДИСКРЕТНЫЕ РАСПРЕДЕЛЕНИЯ

Биномиальное распределение. Функция биномиального распределения

$$V_n(r, p) = \sum_{k=0}^r b_n(k, p), \quad (9)$$

где

$$b_n(k, p) = C^k p^k (1-p)^{n-k}.$$

Справедливы равенства

$$V_n(r, p) = 1 - V_n(r-1, 1-p),$$

$$b_n(r, p) = b_n(n-r, 1-p),$$

из которых вытекает, что необходимы таблицы и программы лишь для $p \leq \frac{1}{2}$. Для ξ , имеющей распределение

$$(9), \quad M\xi = np, \quad D\xi = npq.$$

Коэффициент асимметрии

$$\beta_1 = (q-p)/\sqrt{npq}.$$

Коэффициент эксцесса

$$\beta_2 = 3 - \frac{6}{n} + \frac{1}{npq}.$$

При расчете таблиц и составлении программ вычисления $V_n(r, p)$ и $b_n(r, p)$ используют рекуррентные соотношения типа «вперед»:

$$V_{n+1}(r, p) = pV_n(r-1, p) + qV_n(r, p);$$

$$nb_{n+1}(r, p) = pb_n(r-1, p) + (1-p)b_n(r, p)$$

или типа «назад»:

$$nV_{n-1}(r, p) = (n-r-1)V_n(r, p) + (r+1)V_n(r+1, p);$$

$$nb_{n-1}(r, p) = (n-r)b_n(r, p) + (r-1)b_n(r, p).$$

Функцию $V_n(r, p)$ выражают также через функцию $V(p; a, b)$:

$$V_n(r, p) = 1 - V(p; r+1, n-r).$$

При больших n используют следующие приближения. Пусть

$$\sigma = \sqrt{np(1-p)};$$

$$k_1 = \frac{\lambda - np + 1/2}{\sigma};$$

$$k_2 = \frac{\mu - np + 1/2}{\sigma}.$$

Тогда

$$\sum_{r=\lambda}^{\mu} b_n(r, p) = \Phi(k_2) - \Phi(k_1) +$$

$$+ \frac{1-2p}{6\sigma\sqrt{2M}} \times$$

$$\times [(1-k^2)e^{-k^2/2}] \Big|_{k=k_1}^{k_2} + \omega,$$

где

$$|\omega| < \frac{0,13 + 0,18|1-2p|}{\sigma^2} +$$

$$+ e^{-3\sigma/2}.$$

Для аппроксимации на хвостах (т. е. для вероятностей, принадлежащих отрезкам $[0,005; 0,05]$ и $[0,93; 0,995]$) при прикидочных расчетах используют формулу

$$B_n(r, p) = \Phi(2\sqrt{(r+1)(1-p)} - 2\sqrt{(n-r)p}).$$

Для аппроксимации между хвостами (т. е. для вероятности, принадлежащих отрезку [0,05; 0,93]) при прикидочных расчетах полагают

$$B_n(r, p) = \Phi(\sqrt{(4r+3)(1-p)} - \sqrt{(4n-4r-1)p}).$$

Высокую точность дает формула

$$B_n(r, p) = \Phi\left\{\sqrt{A(r, v)} - \sqrt{A(n-r-1, -v)}\right\},$$

где $v = 1 - rp$,

$$A(r, v) = 2(r+1)(1+v) - \frac{15 + 12v + 5v^2}{18} + \frac{(3 + 6v + 7v^2)(2r - n + 1 + nv)^2}{36n(1-v^2)}.$$

При приближенном нахождении квантили для $B_n(r, p)$ следует полагать

$$B_n(r, p) = \Phi\left\{\left[r - up + q + \frac{1}{3}(u_\alpha^2 - 1)\right] \times \left(1 - 2\frac{r+1}{n+1}\right) \times \sqrt{\frac{n+1}{(r+1)(n-r)}}\right\}.$$

Известны аппроксимации биномиального распределения пуассоновским распределением $\Pi(r, \lambda)$.

Для прикидочных расчетов ($p < 0,2, n \leq 300$)

$$B_n(r, p) = \Pi(r, \lambda_1);$$

$$\lambda_1 = \frac{(2n-r)p}{2-p}.$$

С высокой точностью ($p \leq 0,12, n \leq 300$)

$$B_n(r, p) = \Pi(r, \lambda),$$

где

$$\lambda = \frac{\lambda_1}{1 - \frac{(2\lambda_1^2 - r\lambda_1 - r^2 - 2r)}{6(2n-r)^2}}.$$

Пуассоновское распределение. Функция пуассоновского распределения

$$\Pi(r, \lambda) = \sum_{j=0}^r \pi(j, \lambda), \tag{10}$$

где

$$\pi(r, \lambda) = \frac{\lambda^r e^{-\lambda}}{r!}.$$

Для ξ , имеющей распределение (10),

$$M\xi = \lambda; D\xi = \lambda;$$

$$\beta_1 = 1/\sqrt{\lambda}; \beta_2 = 1/\lambda.$$

Для прикидочных расчетов используют формулы]

$$\Pi(r, \lambda) = \begin{cases} \Phi(2\sqrt{r+1} - 2\sqrt{\lambda}), & \text{на хвостах} \\ \Phi(2\sqrt{r+0,75} - 2\sqrt{\lambda}), & \text{между хвостами.} \end{cases}$$

Аппроксимация высокой точности имеет вид

$$\Pi(r, \lambda) \approx \Phi \times \left(2\sqrt{r + \frac{t+4}{9}} - 2\sqrt{\lambda + \frac{t-8}{36}}\right),$$

где $t = (r - \lambda + 1/6)^2/\lambda$.

Гипергеометрическое распределение. Функция гипергеометрического распределения

$$H(n, d; N, D) = \sum_{i=0}^d h(n, i; N, D), \tag{11}$$

где

$$h(n, d; N, D) = \frac{C_D^d C_{N-D}^{n-d}}{C_N^n},$$

$d = 0, 1, \dots, n;$

N, D, n — целые неотрицательные числа.

Для случайной величины, имеющей распределение (11),

$$E\xi = n \frac{D}{N},$$

$$D\xi = n \frac{D}{N-1} \left(1 - \frac{D}{N}\right) \left(1 - \frac{n}{N}\right),$$

$$\beta_1 = \frac{\left(1 - 2 \frac{D}{N}\right) (N - 2n) \sqrt{N-1}}{\sqrt{n \frac{D}{N} \left(1 - \frac{D}{N}\right) \times (N-2) \sqrt{N-n}}},$$

$$\beta_2 = \frac{c_1(N) + c_2(N) 6 \frac{D}{N} \left(1 - \frac{D}{N}\right)}{n \frac{D}{N} \left(1 - \frac{D}{N}\right)} +$$

$$+ c_3(N) + c_4(N),$$

где

$$c_1(N) = \frac{(N-1)N(N+1)}{(N-2)(N-3)(N-n)};$$

$$c_2(N) = \frac{(N-1)N^2}{(N-2)(N-3)(N-n)};$$

$$c_3(N) = 3 \times$$

$$\times \left[\frac{(N-1)N^2}{(N-2)(N-3)(N-n)} - 1 \right];$$

$$c_4(N) = \frac{18(N-1)}{(N-2)(N-3)} -$$

$$\frac{6(N-1)}{(N-2)(N-3)n \frac{D}{N} \left(1 - \frac{D}{N}\right)} -$$

$$\frac{3(N-1)Nn}{(N-2)(N-3)(N-n)}.$$

СПИСОК ЛИТЕРАТУРЫ

1. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. М.: Наука, 1983. 416 с.
2. **Кутин В. А.** Одна формула обращения интеграла вероятности. — Уч. зап. Перм. ун-та, 1974, № 309, с. 99—101.
3. **Мардна К.** Статистический анализ угловых наблюдений. М.: Наука, 1978. 240 с.

4. **Мартынов Г. В.** Вычисление функций распределения квадратичных форм от нормальных случайных величин. — Теория вероятностей и ее применение, 1975, т. XX, № 4, 797—809.

5. **Мартынов Г. В.** Критерии омега-квадрат. М.: Наука, 1978. 80 с.

6. **Мартынов Г. В.** Вычисление функции нормального распределения. — В кн.: Теория вероятностей. Математическая статистика. Теоретическая кибернетика. Т. 17. М.: ВИНТИ, 1979, с. 57—84.

7. **Смирнов Н. В., Большев Л. Н.** Таблицы для вычисления функции двумерного нормального распределения. М.: Изд-во АН СССР, 1962. 204 с.

8. **Таблицы вероятностных функций.** М.: ВЦ АН СССР, т. 1, 1970. 290 с., т. 2, 1970. 344 с.

9. **Amos D. E.** Computation of the central and noncentral F-distributions. — Commun. Statist. — Theor. Met., A5, 1976, N. 3, p. 263—281.

10. **Bulgren W. G.** Probability integral of double noncentral t-distribution with degrees of freedom n and noncentrality parameters δ and λ . — Selected Tables in Mathematical Statistics, vol. 11., Providens, Rhod Island, 1974, p. 1—138.

11. **Chernick M. L., Murthy V. K.,** Chi-squared percentiles: old and new approximations, with applications to sample size determination. — American Journal of Mathematical and Management Sciences, 1983, 3, N. 2, p. 145—161.

12. **Chien Pai Han.** On the computation of noncentral chi-squared distributions with even degrees of freedom. — J. Statist. Comput. Simul., 1979, 9, p. 25—29.

13. **Davis W. E., Khalil H. M.** Algorithm and series expansion for the F-distribution. — Signum Newsletter, 7, 1972, p. 21—23.

14. **Han C. P., Wang J. L. J.** Computation of noncentral F-distributions. — Commun. Statist. — Simula. Comput, 1983, 12, N. 1, p. 1—9.

15. **Leone F. C.** Tables of cumulative non-central chi-square distribution. — Selected Tables in Mathematical Statistics, vol. I, Providence, Rhod Island, 1974, p. 1—78.

16. Imhof J. P. Computing the distribution of quadratic forms on normal variables. — *Biometrika*, 1961, 48, 1961, p. 419—426.

17. Koechler K. J. A simple approximation for the percentiles of t distribution. — *Appl. Statist.*, 1978, 27, N. 3, p. 280—290.

18. Kennedy W. J., Gentle J. E. *Statistical computing*. Marcel Dekker Inc., New York and Basel, 1980. 591 p.

19. Mudholcar G. S., Chaubery Y. P. Some approximations for noncentral-F distribution — *Technometrics*, 1976, 18, N. 3, p. 351—358.

20. Patnaik P. B. The noncentral χ^2 and F distributions and their applications. — *Biometrika*, 1949, 36, p. 202—232.

21. Plackett R. L. A reduction formula for normal multivariate integrals. — *Biometrika*, 1954, 41, N. 3—4, p. 351—360.

22. Shenton L. R. Inequalities for

the normal integral including a new continued fraction. — *Biometrika*, 1954, 41, N. 1—2, p. 177—178.

23. Sidák Z. A chain of inequalities for some types of multivariate distributions, with nine special cases. — *Appl. Math.*, 1979, A18, N. 2, p. 110—118.

24. Steck G. P. A table for computing trivariate normal probabilities. — *Ann. Math. Statist.*, 1958, 29, N. 3, p. 780—800.

25. Tiku M. L. Double noncentral F-distribution — tables and applications. — *Selected Tables in Mathematical Statistics*, vol. II, Providence, Rhode Island, 1974, p. 139—176.

26. Watson G. S. Large sample theory of the Langevin distribution. — *J. Stat. Plan. and Inf.*, 1983, 8, N. 3, p. 245—256.

27. Zar J. H. Approximations for the percentage points of the chi-squared distribution. — *Appl. Statist.*, 1978, 27, N. 3, p. 280—290.

Предметный указатель

А

Агрегат кусочно-линейный 237, 238 —
 Моделирование 238, 293
 Аксиомы исчисления высказываний 27
 Алгебра борелевская 17
 — булева 23

Б

Бета-распределение 273, 274

В

Вектор случайный дискретный 75
 Величина случайная 70
 — дискретная 71, 72 — Реализация 209, 210
 — многомерная 74 — Моделирование 212, 213
 — непрерывная 72—74
 — одномерная с заданным распределением 210—212
 — положительная 74
 Величины случайные — Дисперсия 79 —
 Ковариация 81 — Математическое ожидание 78 — Независимость двух случайных величин 76 — Сходимость последовательностей 82 — Функция распре-

деления 76, 77 — Числовые характеристики 78—81

Вероятности апостериорные 70

— априорные 70
 — гипергеометрические 198
 — переходные 87

Вероятность события 66
 — условная 69

Время дискретное 168
 — остаточное 93

Выбор случайный без возвращения 196
 Выборка 196

Г

Гамма-распределение 130

Гипотезы статистические — Проверка 112—114

Граница контрольная 207

— Крамера—Рао 110
 — функции γ -доверительная 111, 112

Граф 32 — Маршруты 33, 34 — Степени и полустепени вершин 32, 33 — Части 34, 35

— неориентированный 32, 33
 — ориентированный 32, 33 — Конденсация и база 36, 37

Графы — Деревья 43 — Изоморфизм 37 — Кодеревья 43, 44 — Остовы 43, 44 —

- Специальные типы 41—44 — Типы связности 35, 36
 — гамильтоновы 42, 43
 — двойственные 41, 42
 — двухдольные 41
 — дополнительные 42
 — неразделимые 44
 — однородные 41
 — планарные 41
 — полные 41
 — регулярные 41
 — эйлеровы 42

Д

- Датчик случайных чисел 208
 Движение броуновское 95
 Дизъюнкция 23
 Дубликат графа неориентированный 32, 33

З

- Задача оптимизации 164, 170—172, 175, 177, 179, 180
 Закон арксинуса 98
 — Вейбулла—Гнеденко 130
 — нормальный 130
 — 0 или 1 100
 — повторного логарифма 96
 — экспоненциальный (показательный) 130
 Законы больших чисел 82—84

И

- Импликант функции простой 24
 Инвариантность 114
 Интеграл стохастический 95
 — уравнения 50—52
 Интервал γ -доверительный 111
 Интерпретация формул 25, 26
 Исчисление высказываний 26—28
 — предикатов 31

К

- Карты контрольные 207
 Кванторы 28
 Классы стратегий 167, 168
 Контроль выборочный 196
 — непрерывного потока штучной продукции 204—207
 — разрушающий 196
 — по плану Доджа 204, 205
 — последовательный — Планы 203
 — сплошной 196
 — статистический — Планы 195—198 — Последующие оценки 202—204 — Системы планов 201—202
 Контроль по признаку альтернативному 195, 196
 — качественному 195, 196
 — количественному (векторному) 195, 196
 — функциональному 196
 Конъюнкция элементарная 23
 Коэффициент доверия 111
 — асимптотический 112
 Кратность точки 93
 Критерий качества 127
 — факторизации 104
 — фундаментальности системы решений уравнения 54

Л

- Логика высказываний 25, 26

М

- Марковость строгая 87
 Марковский вверх процесс 152—154
 Мартингалы 93, 94
 Матрица достижимости 46
 — инцидентий 45
 — ковариационная 81
 — переходных вероятностей 87
 — полустепеней захода 46
 — расстояний 46
 — смежности 45, 46
 — стохастическая 87
 — Фишера информационная 109
 Матрицы графа 45—47
 Метод выделения главной части 219
 — двойной выборки 216
 — зависимых испытаний 223
 — искусственного базиса 123, 124
 — квантилей 107
 — многошаговый 217, 218
 — моделирования обобщенный 212
 — моментов 107
 — последовательный 216, 217
 — путей и сечений 138, 140
 — расслоенной выборки 221—223
 — рекуррентный 140
 — существенной выборки 219—221
 — уменьшения дисперсии 219—223
 Методы аналитико-статистические 223—225
 М-метод 124
 Множества — Алгебра 16 — Кольцо 16 — Минимальная σ -алгебра 17 — Объединение (сумма) множеств 13, 14 — Пересечение 14 — Пределы 15 — Разность 14 — σ -алгебра 15, 17 — σ -кольцо 16 — Эквивалентность 12
 — бесконечные 12
 — борелевские 17
 — выпуклые 116, 117
 — измеримые 17
 — конечные 12
 — непересекающиеся 14
 — несчетные 12
 — счетные 12
 — упорядоченные 18
 — частично упорядоченные 18
 Множество — Дополнение 15 — Мощность 13 — Образ 18 — Символ включения 11 — Способы задания 11
 — элементарных событий 65
 Моделирование — Основные системы 239
 — полумарковских процессов 214, 215
 — пуассоновского потока с параметром $\lambda(t)$ 213
 — случайных величин 208—213
 — цепи Маркова с непрерывным временем 213, 214
 Модель резервирования с восстановлением марковская 157
 — общая — Двусторонние неравенства для характеристик надежности 159—160 — Определение 157—159
 Модель сложной восстанавливаемой системы 160—162
 — статистическая оценка функционала 209
 Модель технического обслуживания системы без учета структуры 175—177
 — дублированной 177—180
 Моменты восстановлений 92
 Мультиграф 32, 33

Н

- Надежность элемента восстанавливаемого 131—135

- невозстанавливаемого 129, 130
- системы с независимо отказывающимися элементами 136—138
- Некоррелированность** 81
- Неравенство Колмогорова** 82
- Ферника 101
- Чебышева 81, 82

О

- Образ множества** 18
- Обслуживание техническое** — Физическое описание задачи 163, 164
- Объект управления** 163
- управляемый 164—166
- Объем выборки** 196
- партии 197
- Операции над графами бинарные** 40, 41
- унарные 37—40
- Отказы элементов в системе** — Модели зависимости 141, 142
- Отношение бинарное** 47
- порядка 17
- слабого упорядочения 18
- эквивалентности 17
- Отображение взаимно однозначное** 18
- обратное 19
- Оценивание доверительное** 111, 112
- Оценка максимального правдоподобия** 106
- нестационарных характеристик надежности систем 227—237
- распределения времени безотказной работы систем 225—227
- Оценки Горвица—Томпсона** 106

П

- Параметр сильной выпуклости** 118
- План Вальда** 197, 198
- План контроля двуступенчатый** 197
- критичный 205, 206
- одноступенчатый 196, 197
- по количественному признаку 206, 207
- последовательного 197, 198
- статистического 198, 199
- Плотность распределения** 72
- совместного 75
- Подграфы** 34, 35
- Подмножество** 12
- Показатели базовые** 207
- Поле случайное вещественное** 86
- гауссовское 100, 101
- Поток событий** 87
- Предел среднего выходного качества** 199
- Предикаты** 28—31 — Аксиоматическая теория исчисления 31
- Принцип двойственности** 15
- оптимальности Беллмана 127, 128
- Программирование выпуклое** 118—120
- динамическое — Оптимизация дискретных процессов 126—128 — Сущность 124, 125
- линейное (ЛП) — Двойственные задачи ЛП со смешанными ограничениями 121, 122 — Задача ЛП и ее свойства 120, 121 — Методы решения задач 122—124
- Прообраз множества** 19
- Пространство вероятностное** 65
- дискретное вероятностное 67—69
- естественное параметрическое 104
- значений 18
- измеримое 17
- состояний 86
- управлений 164, 166
- управлений и стратегий 169

- Процесс авторегрессии** 97
- гибели и рождения 189—191
- винеровский 95—96
- восстановления 91, 92 — Анализ систем, описываемых процессом 192—194
- Процесс марковский** 90, 91
- однородный — Вероятности состояний 143 — Вычисление основных характеристик 143—145 — Определение 142, 143 — Процесс чистой гибели и его обобщение 145—148
- однородный с конечным числом состояний — Управление 172, 173
- с дискретным временем — Анализ систем, описываемых процессом 187—189
- с непрерывным временем — Анализ систем, описываемых процессом 189—191
- Процесс Орнштейна—Уленбека** 100
- полумарковский — Анализ систем, описываемых процессом 191, 192
- полумарковский управляемый 168—172
- рождения и смерти 155—157
- скользящего среднего 97
- Процесс случайный** 86, 164—166
- вещественный 86
- гауссовский 99—102
- регенерирующий 149, 150, 173, 174 — Управление 173—175
- регенерирующий специального типа 152
- согласованный 87
- стационарный 96 — Линейное преобразование 98, 99
- точечный 93
- управляемый 165, 176
- управляемый с непрерывным временем и дискретным управлением 166, 167
- Процессы с ортогональными приращениями** 95, 96
- Псевдограф** 32

Р

- Распределение абсолютно непрерывное** 72
- априорное 200
- бета-распределение 73
- биномиальное 71, 276
- биномиальное отрицательное 72
- Вейбулла—Гнеденко 73
- гамма-распределение 73
- геометрическое 71
- гипергеометрическое 72, 275, 276
- логарифмически нормальное 72, 73
- нормальное 72, 255—266
- нормальное двумерное 76
- нормальное усеченное 72
- Пойа 200
- Пуассона 72
- пуассоновское 275
- равномерное 72, 73
- стареющее 131
- суббиномиальное 200
- супербиномиальное 200
- Стьюдента 74
- Фишера 74
- хи-квадрат 73, 267—271
- эквибиномиальное 200
- экспоненциальное 73
- Эрланга 73
- Распределения дискретные** 274—276
- непрерывные 251—254

С

- Связки логические** 25
- Семейство экспоненциального типа** 104

Симплекс-метод 122—124
 Система аксиом 66
 — γ -доверительных множеств 111, 112
 Системы показательные последовательные с компонентами зависимыми — Анализ надежности 183—187
 — независимыми — Анализ надежности 181—183
 Системы случайных величин 74—76
 Следствия логические 26
 Совокупность генеральная 103
 Статистика математическая — Оценки 106, 107 — Показатели качества оценок 108—111
 Статистики 104, 105
 Стратегия марковская 167
 — нерандомизированная 167
 — однородная 167
 — оптимальная 168
 — рандомизированная 167
 — управления 165, 166
 Структура дохода 176—179

Т

Теорема Блэкуэлла 93
 — интегральная Муавра — Лапласа 84
 — Куна—Таккера 119
 — локальная 84
 — Ляпунова 85
 — Пуассона 84
 — существования 168
 — центральная предельная 84, 85
 — узловая 93
 — умножения 69
 — Фаркаша 117
 Теоремы предельные 84, 85, 150—152
 Теория вероятностей — Основные понятия 65—67

У

Управление процессом $\xi(t)$ дискретное с временем дискретным 164, 165
 — непрерывным 165—167
 Уравнение Беллмана 127, 128
 Уравнения в частных производных первого порядка 62—64
 Уравнения обыкновенные дифференциальные — Задачи Коши для уравнений n -го порядка 50, 51 и систем уравнений 49, 50 — Зависимость решений задачи Коши от начальных данных и параметров 52, 53 — Интегралы уравнения n -го порядка 50, 51 — Краевые задачи 61 — Первые интегралы системы 51, 52 —

Свойства решений дифференциального уравнения второго порядка 57, 58 — Система дифференциальных уравнений, 49 — Системы линейных уравнений 58—60 — Устойчивость решений систем уравнений 60, 61 — Фундаментальная система решений 58 — Фундаментальное решение 56, 57
 Уравнения обыкновенные дифференциальные линейные 54—56
 Условие регулярности 109, 119

Ф

Формы нормальные 23, 24
 Формула Байеса 69
 — исчисления высказываний 25
 — общезначимая 26
 — полной вероятности 69
 — противоречивая 26
 Функции булевы 21, 22
 — вогнутые 117, 118
 — восстановления 92
 — выпуклые 117, 118
 — от случайных величин 76—78
 Функционал 20
 — дохода 169, 170, 174, 175
 Функция 18
 — измеримая 19
 — Лагранжа 119
 — множеств 19, 20
 — неизмеримая 19
 — правдоподобия 103, 104
 — распределения 70, 71, 74, 75
 — экстремальная 175
 F-распределение 271—273

Ц

Цепи марковские 87—91
 Цепь Маркова дискретная — Реализация 210
 — с конечным множеством состояний — Управление 173

Ч

Числа случайные 208
 Число восстановлений 92

Э

Эксперимент 65 — Математическая модель 65—67
 Эллипсоид рассеяния 108