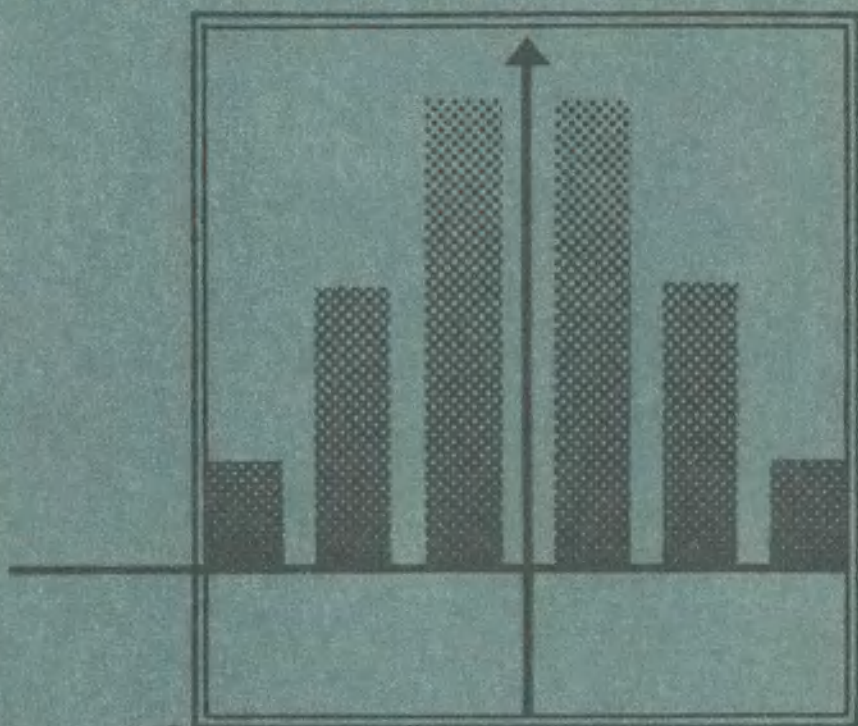


*И. Н. Коваленко
Б. В. Гнеденко*

Теория ВЕРОЯТНОСТЕЙ



**Вища
школа**

И.Н.Коваленко
Б.В.Гнеденко

Теория ВЕРОЯТНОСТЕЙ

Допущено
Министерством высшего и среднего
специального образования СССР
в качестве учебника для
студентов университетов и вузов

КИЕВ
«ВЫЩА ШКОЛА»
1990

ББК 22.171я73

К 56

УДК 519.21(075.8)

Рецензенты:

д-р физ.-мат. наук, проф. *В. В. Булдыгин* (Киевский политехнический институт); чл.-кор. АН УССР, проф. *М. И. Ядренко* и канд. физ.-мат. наук, доц. *Н. Н. Леоненко* (Киевский государственный университет)

Редакция литературы по математике и физике

Редактор *Л. П. Онищенко*

Коваленко И. Н., Гнеденко Б. В.

К 56

Теория вероятностей : Учебник. — К. : Выща шк., 1990. — 328 с. : ил.

ISBN 5-11-001842-1

Излагаются основные разделы теории вероятностей, теории случайных процессов и математической статистики. Фундаментальные понятия (вероятность, случайная величина, математическое ожидание) приведены в терминах аксиоматического подхода А. Н. Колмогорова. Большое внимание уделяется разъяснению этих понятий на примерах. Случайные величины излагаются в векторной концепции. Цепи Маркова даются параллельно в дискретном и непрерывном вариантах. Рассматриваются стационарные, гауссовские, регенерирующие, полумарковские процессы. Одна из глав посвящена теории массового обслуживания.

Для студентов университетов и вузов

К 1602090000—115
М 211(04)—90 64—90

ББК 22.171я73

ISBN 5-11-001842-1

© И. Н. Коваленко,
Б. В. Гнеденко, 1990

ПРЕДИСЛОВИЕ

Теория вероятностей, математическая статистика и теория случайных процессов, будучи самостоятельными математическими науками, служат теоретической основой изложения ряда специальных дисциплин, а следовательно, представляют собой неотъемлемую составную часть фундаментальной подготовки будущих специалистов. Роль этих наук возрастает в связи с привлечением студентов к решению научных и прикладных задач.

В настоящем учебнике изложение основных положений ведется на базе аксиоматического метода А. Н. Колмогорова. Основное понятие — математическое ожидание — вводится как абстрактный интеграл Лебега. Достаточно внимания уделяется наглядной интерпретации этого понятия. При изложении опускаются излишние подробности. Но в рамках постановок, которые часто могут встретиться на практике (интеграл Римана), изложение вполне строго и подробно. Использование интеграла Лебега во многих случаях способствует более четкому и ясному изложению материала.

При рассмотрении последовательности независимых испытаний не только доказаны относящиеся к ней факты, но на этой простой модели проиллюстрировано проявление таких глубоких общих закономерностей теории вероятностей, как оценка вероятностей больших отклонений, закон больших чисел, усиленный закон больших чисел. Рассмотрены ограниченные и неограниченные случайные блуждания. На модели независимых испытаний демонстрируется применение метода решения задач, связанных с вероятностями редких событий с разнообразными приложениями, например, к вычислению вероятности сохранения связи при обрывах линий.

При рассмотрении случайных величин принята векторная концепция изложения. Приведены физические и механические интерпретации, которые позволяют студенту создать единое представление о плотности в пространстве произвольной размерности. Достаточно внимания уделено матричным преобразованиям случайных величин, знание которых необходимо при обработке статистических наблюдений (например, при определении параметров траекторий объектов по радиолокационным наблюдениям).

Цепи Маркова излагаются в двух параллельных вариантах: дискретном и непрерывном. В обоих вариантах особо подчеркивается единство аналитического аппарата. Уделено внимание не только эргодическому распределению, но и оценке отклонения от него.

Среди примеров цепи Маркова с непрерывным временем рассмотрен ветвящийся процесс, описывающий различные физические явления.

Приведены необходимые сведения о типах сходимости случайных величин, полноте пространств. Они необходимы для строгого изложения теории случайных процессов: без них было бы невозможно обосновать каноническое разложение случайного процесса. Рассмотрены теорема Биркгофа — Хинчина и понятие эргодического процесса.

Центральная предельная теорема — один из фундаментальных результатов теории вероятностей — изложена так, что читатель познакомится с методом оценок распределений по характеристическим функциям. Обсуждаются следствия из этой теоремы.

Изложение перечисленного выше материала обеспечивает теоретическую основу для изучения более сложных разделов теории вероятностей. Так, в главе о корреляционном анализе случайных процессов постоянно используется подробно рассмотренная ранее теорема о полноте пространства случайных величин с конечными дисперсиями.

Рассмотрены разнообразные классы случайных процессов, наиболее часто применяемые в прикладных исследованиях. С одной стороны, это процессы, свойственные статистической радиотехнике (стационарные, гауссовские), с другой — процессы, на основе которых строятся модели дискретных событий (дискретные автоматы, различного рода системы управления): марковские, полумарковские, потоки однородных событий.

На единой концепции рекуррентного принципа построена глава «Теория массового обслуживания». В отличие от других книг авторами предложен общий подход к построению модели реальных сложных систем в виде системы обслуживания. Приведены примеры расчета резервированных систем.

Отдельная глава посвящена изложению математической статистики. Изложение всего материала сопровождается рассмотрением полезных и интересных примеров из различных областей приложений.

Главы 1—11 и «Дополнение» написаны И. Н. Коваленко. В них обобщен его опыт преподавания теории вероятностей и ее приложений в Московском институте электронного машиностроения, Киевском университете им. Т. Г. Шевченко, Киевском высшем инженерном радиотехническом училище ПВО им. А. И. Покрышкина, а также опыт работы над учебными пособиями совместно с О. В. Сармановым, А. А. Филипповой, Г. И. Ивченко и В. А. Каштановым.

Исторический обзор написан Б. В. Гнеденко на основании его исследований по истории теории вероятностей, в том числе исследований, выполненных им в последнее время.

§ 1. Случайные события. Статистическая вероятность

Рассмотрим некоторый опыт, в результате которого может наступить либо не наступить событие A . Примерами такого опыта могут быть:

- а) опыт — бросание монеты, событие A — выпадение герба;
- б) опыт — стрельба 10 выстрелами по мишени, событие A — набор ровно 90 очков;
- в) опыт — суточный прогноз средней температуры воздуха, событие A — ошибка не более, чем на 5°C ;
- г) опыт — применение лекарственного препарата, событие A — улучшение состояния больного;
- д) опыт — ввод в ЭВМ программы, событие A — безошибочный ввод.

Общее в этих опытах то, что каждый из них может реализоваться в данных условиях в принципе сколько угодно раз. Такие опыты называются *испытаниями*. Частотой $p_n(A)$ события A в первых n испытаниях называется отношение числа $m_n(A)$ наступлений события в этих испытаниях к числу испытаний n :

$$p_n(A) = m_n(A)/n. \quad (1)$$

Во многих реальных случаях с увеличением n частота события A стабилизируется, практически мало отличаясь от некоторого числа p . Такое число p , являющееся мерой возможности наступления события A в испытании, может быть принято в качестве *вероятности* события A :

$$p = P(A), \quad (2)$$

где символом $P(A)$ (от франц. probabilité — вероятность) обозначается вероятность события A . В индивидуальном испытании невозможно предвидеть, произойдет ли событие A ; можно лишь на основании большого числа предыдущих испытаний определить меру возможности наступления данного события.

Вероятность события представляется как предел частоты при неограниченном увеличении числа испытаний:

$$P(A) = \lim_{n \rightarrow \infty} p_n(A). \quad (3)$$

Это определение является математической идеализацией реальных явлений, поскольку в действительности экспериментатор может реализовать лишь конечное число испытаний, а требование

полного совпадения их условий выполнимо лишь с некоторым приближением.

Вероятность события A , найденная на основании испытаний, называется его *статистической вероятностью*. Событие, не прогнозируемое в отдельных испытаниях и обладающее устойчивостью частоты в большом числе испытаний, т. е. такое событие, которому можно приписать некоторую вероятность, называется *случайным событием*.

§ 2. Пространство элементарных событий. Действия над событиями

Испытание может иметь различные *исходы*. Так, при бросании монеты имеется два возможных исхода — герб или надпись; при испытании изделия на долговечность исходом является момент его отказа. Что принимать в качестве исхода испытания, зависит от условий задачи. Например, в испытании «выстрел по мишени» можно считать исходом число выбитых очков, в другой постановке задачи — координаты точки попадания.

Множество Ω исходов испытания называется *пространством элементарных событий*, сами же исходы — *элементарными событиями*. Имеем $\Omega = \{\omega\}$, где ω — символ элементарного события.

Пусть A — событие, которое может наступить либо не наступить в данном испытании. Факт наступления события A определяется его исходом, т. е. элементарным событием. Если при данном ω событие A наступает, то говорят, что элементарное событие ω *благоприятствует* событию A ; если при данном ω событие A не наступает, то ω *не благоприятствует* A . Так, в случае стрельбы по мишени положим, что элементарное событие ω — расстояние от центра мишени до точки попадания. Тогда событию «попадание в десятку» благоприятствуют элементарные события ω , для которых $0 \leq \omega \leq R$, где R — радиус центрального круга мишени, остальные же ω не благоприятствуют. Следовательно, событие A можно отождествить с множеством благоприятствующих ему элементарных событий. Это множество обозначается тем же символом A , что и событие. Таким образом, событие есть подмножество A пространства элементарных событий Ω .

Достоверным событием Ω называется событие, которому благоприятствуют все элементарные события ω . *Невозможным событием* \emptyset называется событие, которому не благоприятствует никакое элементарное событие.

Событием, *противоположным* событию A , называется событие \bar{A} , которому благоприятствуют ω , не благоприятствующие A , и не благоприятствующие ω , благоприятствующие A . Можно записать $\bar{A} = \Omega \setminus A$. *Объединением (суммой)* $A \cup B$ событий A и B называется событие, состоящее в наступлении хотя бы одного из них. Объединению благоприятствуют те и только те ω , которые благоприятствуют либо A , либо B . *Произведением (пересечением)* $AB =$

$= A \cap B$ событий A и B называется событие, состоящее в наступлении обоих этих событий. Аналогично определяются объединение и пересечение конечного и бесконечного множеств событий:

$$A_1 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i, \quad \bigcup_{i=1}^{\infty} A_i, \quad \bigcup_{\alpha \in I} A_{\alpha};$$

$$A_1 \dots A_n = \bigcap_{i=1}^n A_i, \quad A_1 A_2 \dots = \bigcap_{i=1}^{\infty} A_i, \quad \bigcap_{\alpha \in I} A_{\alpha}.$$

Событие $\bigcap_{\alpha \in I} A_{\alpha}$ означает, что любому элементу α множества I сопоставлено событие A_{α} и рассматривается событие, состоящее в наступлении всех A_{α} одновременно.

События A, B называются *несовместными*, если $AB = \emptyset$, т. е. не существует элементарного события, благоприятствующего как A , так и B .

События A_1, A_2, \dots называются *попарно несовместными*, если $A_i A_j = \emptyset$ при любых $i \neq j$.

Событие B называется *следствием* события A ($A \subset B$), если все ω , благоприятствующие A , благоприятствуют B . Если $A \subset B$ и $B \subset A$, то события A и B называются *равными*: $A = B$. Равным событиям благоприятствуют одни и те же элементарные события.

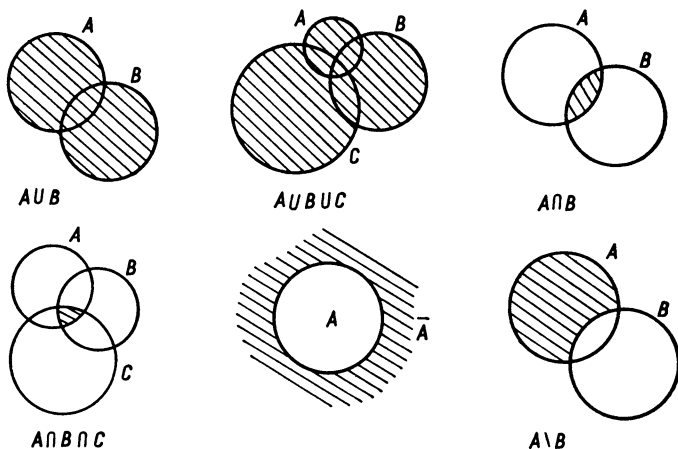


Рис. 1

Разностью $A \setminus B$ событий A и B называется событие, которому благоприятствуют те и только те ω , которые благоприятствуют A и не благоприятствуют B .

Введенные отношения между событиями — обычные отношения между множествами. Эти отношения удобно представлять графически (рис. 1).

Объединение и пересечение имеют ряд очевидных свойств, а именно:

коммутативность

$$A \cup B = B \cup A, \quad AB = BA;$$

ассоциативность

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (AB)C = A(BC);$$

дистрибутивность

$$(A \cup B)C = (AC) \cup (BC), \quad (AB) \cup C = (A \cup C)(B \cup C).$$

Введем функцию $I_A(\omega)$ элементарного события ω , положив $I_A(\omega) = 1$ при $\omega \in A$, $I_A(\omega) = 0$ при $\omega \notin A$. Функция $I_A(\omega)$ называется *индикатором* события A . Выясним, какие операции над индикаторами соответствуют операциям над множествами. Имеем

$$I_{\bar{A}} = 1 - I_A,$$

$$I_{A \cup B} = \max\{I_A, I_B\},$$

$$I_{AB} = I_A I_B, \quad I_{A \setminus B} = I_A(1 - I_B).$$

Соотношение $A \subset B$ равносильно неравенству $I_A \leq I_B$.

Тождество двух выражений относительно событий A_1, \dots, A_n сводится к тождеству двух выражений от индикаторов этих событий. Это последнее проверяется подстановкой вместо $(I_{A_1}, \dots, I_{A_n})$ всевозможных n -мерных векторов из нулей и единиц. Подобным же образом проверка следования двух выражений $\Phi(A_1, \dots, A_n) \subset \Psi(A_1, \dots, A_n)$ сводится к проверке неравенства для соответствующих индикаторов.

Пример 1. Доказать правила де-Моргана

$$\overline{A_1 \dots A_n} = \bigcup_{i=1}^n \bar{A}_i, \quad \bigcap_{i=1}^n A_i = \bar{A}_1 \dots \bar{A}_n. \quad (1)$$

Обозначив для краткости $I_i = I_{A_i}$, получим

$$I_{\overline{A_1 \dots A_n}} = 1 - I_{A_1 \dots A_n} = 1 - I_1 \dots I_n;$$

в то же время

$$I_{\bigcup_i \bar{A}_i} = \max_i I_{\bar{A}_i} = \max_i (1 - I_i).$$

Если этот максимум равен 1, то некоторое $I_i = 0$, а следовательно, $1 - I_1 \dots I_n = 1$. Если максимум равен 0, то все $I_i = 1$ и $1 - I_1 \dots I_n = 0$. Первое правило де-Моргана доказано. Второе правило доказывается аналогично.

Пример 2. Доказать соотношение

$$I_{A_1 \cup \dots \cup A_n} = \sum_{r=1}^n (-1)^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq n} I_{A_{i_1} \dots A_{i_r}}. \quad (2)$$

Применим первое правило де-Моргана, заменив в нем \bar{A}_i на A_i (тогда, очевидно, A_i заменится на \bar{A}_i):

$$I_{A_1 \cup \dots \cup A_n} = I_{\overline{\bar{A}_1 \dots \bar{A}_n}} = 1 - I_{\bar{A}_1 \dots \bar{A}_n} = 1 - I_{\bar{A}_1} \dots I_{\bar{A}_n} = 1 - (1 - I_1) \dots$$

$$\dots (1 - I_n) = 1 - \sum_{r=0}^n (-1)^r \sum_{1 \leq i_1 < \dots < i_r \leq n} I_{i_1} \dots I_{i_r}.$$

Остается заметить, что в последней сумме нулевое слагаемое сокращается с единицей, а $I_{i_1} \dots I_{i_r} = I_{A_{i_1} \dots A_{i_r}}$.

Отметим частные случаи формулы (2). Положив $n = 2$, $A_1 = A$, $A_2 = B$, получим

$$I_{A \cup B} = I_A + I_B - I_{AB}. \quad (3)$$

Положив $n = 3$, $A_1 = A$, $A_2 = B$, $A_3 = C$, получим

$$I_{A \cup B \cup C} = I_A + I_B + I_C - I_{AB} - I_{AC} - I_{BC} + I_{ABC}. \quad (4)$$

Индикаторы можно применять и при действиях с бесконечным множеством событий. Так, для любых событий A_n имеем

$$I_{\bigcup_{n=1}^{\infty} A_n} \leq \sum_{n=1}^{\infty} I_{A_n}. \quad (5)$$

Действительно, если правая часть неравенства (5) равна 0, то $I_{A_n} = 0$ для всех n , т. е. $\omega \notin \bigcup_{n=1}^{\infty} A_n$, а следовательно, левая часть также равна 0. Если же правая часть больше или равна 1, то левая часть, очевидно, не превосходит ее.

Если $I_i = 0$, $1 \leq i \leq n$, то все слагаемые в правой части соотношения (2) равны нулю. Если $\sum_{i=1}^n I_i = k > 0$, то соотношение (2) превращается в тождество

$$1 = \sum_{r=1}^k (-1)^{r-1} C_k^r.$$

Поскольку по формуле Тейлора

$$0 = (1 - 1)^k = \sum_{r=0}^m (-1)^r C_k^r + (-1)^{m+1} C_k^{m+1} \theta^{m+1}, \quad 0 < \theta < 1,$$

то

$$\begin{aligned} \sum_{r=1}^{2s} (-1)^{r-1} \sum_{i_1 < \dots < i_r} I_{A_{i_1} \dots A_{i_r}} &\leq \\ &\leq I_{A_1 \cup \dots \cup A_n} \leq \sum_{r=1}^{2s-1} (-1)^{r-1} \sum_{i_1 < \dots < i_r} I_{A_{i_1} \dots A_{i_r}}. \end{aligned} \quad (6)$$

При $s = 1$ получаем наиболее важный частный случай (6):

$$\sum_{i=1}^n I_{A_i} - \sum_{1 \leq i < j \leq n} I_{A_i A_j} \leq I_{A_1 \cup \dots \cup A_n} \leq \sum_{i=1}^n I_{A_i}. \quad (7)$$

§ 3. Аксиоматическое определение вероятности

Данное в § 1 определение статистической вероятности относится к реально наблюдаемым событиям. В то же время оказалось необходимым аксиоматическое определение вероятности, на основании которого строится теория вероятностей — строгая математическая теория, с помощью которой создаются вероятностные модели реальных явлений. Исходные данные для модели берутся из экспери-

мента, но после того, как модель построена, исследуется она методами математики. Выводы, полученные в результате такого исследования, снова применяются к реальному объекту. Способы получения из реальных наблюдений исходных данных для построения вероятностных моделей дает математическая статистика.

Изложим аксиоматическое определение вероятности, данное А. Н. Колмогоровым в 1933 г. Используемые при этом понятия σ -алгебры, меры и интеграла разъясняются в «Дополнении».

Задается пространство элементарных событий ω — произвольное множество Ω и σ -алгебра \mathcal{F} его подмножеств, называемых *событиями*. Любому событию A ставится в соответствие неотрицательное число $P(A)$ — *вероятность* события A . Как функция множества A вероятность представляет собой нормированную меру, называемую *вероятностной мерой* (*распределением вероятностей*), а следовательно, удовлетворяет следующим аксиомам:

$$1. P(A) \geq 0.$$

$$2. P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \quad (1)$$

для попарно несовместных событий A_n .

$$3. P(\Omega) = 1.$$

Аксиома 2 называется *расширенной аксиомой сложения* (*аксиомой счетной аддитивности*).

Тройка $\langle \Omega, \mathcal{F}, P \rangle$, т. е. измеримое пространство $\langle \Omega, \mathcal{F} \rangle$ с вероятностной мерой, называется *вероятностным пространством*.

Поскольку $\emptyset = \emptyset \cup \emptyset \cup \dots$, причем события $\emptyset, \emptyset, \dots$ попарно несовместны, то по расширенной аксиоме сложения $P(\emptyset) = P(\emptyset) + P(\emptyset) + \dots$, а следовательно,

$$P(\emptyset) = 0. \quad (2)$$

Из (1) следует аксиома сложения

$$P\left(\bigcup_{n=1}^N A_n\right) = \sum_{n=1}^N P(A_n) \quad (3)$$

для попарно несовместных A_n ; для доказательства достаточно в формуле (1) положить $A_n = \emptyset$ при $n > N$ и воспользоваться равенством (2).

Из расширенной аксиомы сложения вытекают следующие утверждения, называемые *аксиомой непрерывности*.

1. Если $A_n \subset A_{n+1}$, $n \geq 1$, то

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n). \quad (4)$$

2. Если $A_{n+1} \subset A_n$, $n \geq 1$, то

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n). \quad (5)$$

Измеримая числовая функция $\xi = f(\omega)$ элементарного события называется *случайной величиной*.

Интеграл $\int f(\omega) \mathbf{P}(d\omega)$, если он существует, называется *математическим ожиданием* случайной величины ξ и обозначается $\mathbf{M}\xi$. Случайная величина ξ называется *дискретной*, если $f(\omega)$ — ступенчатая функция, т. е. Ω разбивается на события A_n , на которых эта функция постоянна. Из определения интеграла (см. «Дополнение»)

$$\mathbf{M}\xi = \sum_n x_n \mathbf{P}(A_n), \quad (6)$$

где $x_n = f(\omega)$, $\omega \in A_n$.

Распределением случайной величины $\xi = f(\omega)$ называется функция множества

$$\mathbf{P}(\xi \in A) = \mathbf{P}(\{\omega : f(\omega) \in A\}), \quad (7)$$

заданная для любых борелевских множеств A . Распределение случайной величины ξ полностью определяется ее функцией распределения

$$F_\xi(x) = \mathbf{P}(\xi < x) = \mathbf{P}(\{\omega : f(\omega) < x\}), \quad (8)$$

заданной для любых $x \in \mathbb{R}$.

Пример. Пусть Ω — интервал $(0; 1)$ и вероятностная мера \mathbf{P} характеризуется тем, что мера любого интервала $(a; b)$, $0 \leq a \leq b \leq 1$, равна длине этого интервала, т. е. равна $b - a$.

Определим случайную величину $\xi = 1/\omega$. Тогда при $x > 1$ $\xi < x$ в том и только том случае, если $\omega > 1/x$, т. е. $\{\omega : f(\omega) < x\} = (1/x, 1)$. Следовательно, $F_\xi(x) = \mathbf{P}((1/x, 1)) = 1 - 1/x$. При $x \leq 1$ множество тех ω , для которых $\xi = 1/\omega < x$, пусто, а следовательно, $F_\xi(x) = 0$.

Распределение дискретной случайной величины ξ задается *рядом распределения* (x_n, p_n) , где x_n — возможные значения ξ , p_n — их вероятности:

$$p_n = \mathbf{P}(\xi = x_n) = \mathbf{P}(\{\omega : f(\omega) = x_n\}). \quad (9)$$

Вероятность можно выразить через математическое ожидание, а именно:

$$\mathbf{P}(A) = \mathbf{M}I_A(\omega) = \int I_A(\omega) \mathbf{P}(d\omega). \quad (10)$$

Поскольку интеграл линеен относительно подынтегральной функции, то из определения $\mathbf{M}\xi$ следует важная формула

$$\mathbf{M}(c_1\xi_1 + \dots + c_n\xi_n) = c_1\mathbf{M}\xi_1 + \dots + c_n\mathbf{M}\xi_n. \quad (11)$$

В качестве следствия введенных понятий получим формулу для вероятности объединения событий в общем случае, т. е. без предположения об их попарной независимости. Проинтегрировав обе части равенства (2.2)¹ по вероятностной мере \mathbf{P} , придем к формуле

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{r=1}^n (-1)^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq n} \mathbf{P}(A_{i_1} \dots A_{i_r}), \quad (12)$$

¹ При ссылках на формулы с двойной нумерацией первая цифра обозначает номер параграфа, а вторая номер формулы, соответственно при ссылках на параграфы — первая цифра номер главы, а вторая номер параграфа.

называемой *формулой включения и исключения*. В частности, при $n = 2$ и $n = 3$ (см. (2.3), (2.4)) имеем

$$P(A \cup B) = P(A) + P(B) - P(AB), \quad (13)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - \\ - P(BC) + P(ABC).$$

Пример. Между двумя пунктами имеется три канала связи. Вероятность исправности каждого из них в отдельности равна a , вероятность исправности любых фиксированных двух каналов равна b , вероятность исправности всех трех каналов равна c . Тогда вероятность исправности хотя бы одного канала равна $3a - 3b + c$.

Из равенства (13) находим

$$P(A \cup B) \leq P(A) + P(B), \quad (14)$$

откуда по индукции

$$P(A_1 \cup \dots \cup A_n) \leq P(A_1) + \dots + P(A_n). \quad (15)$$

Для бесконечной последовательности событий также имеем неравенство

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n), \quad (16)$$

которое получаем при интегрировании неравенства (2.5) и применении теоремы Лебега (см. «Дополнение»).

§ 4. Классическое понятие вероятности

Пусть Ω состоит из n элементарных событий $\omega_1, \dots, \omega_n$, каждое из которых имеет вероятность $P(\{\omega_i\}) = 1/n$. Тогда, если событию A благоприятствуют m элементарных событий, то можно записать

$$A = \bigcup_{k=1}^m \{\omega_{i_k}\};$$

а так как события $\{\omega_{i_k}\}$ попарно несовместны, то по аксиоме сложения

$$P(A) = \sum_{k=1}^m P(\{\omega_{i_k}\}) = \frac{m}{n}. \quad (1)$$

В классическом определении вероятностей, возникшем на базе задач, связанных с азартными играми, еще в XVIII в., в основу определения полагалось n равновероятных исходов; в таком случае вероятность события определялась как отношение числа исходов, благоприятствующих этому событию, к общему числу исходов. Мы вывели эту формулу в качестве частного случая аксиоматического определения вероятности.

Поскольку элементарные события имеют определенные вероятности, т. е. входят в \mathcal{F} , то и любые их множества по свойству

σ -алгебры входят в \mathcal{F} . Итак, в данном случае σ -алгебра событий включает все подмножества Ω . Всего имеется 2^n различных событий: каждое ω_i , независимо от остальных, можно включать либо не включать в множество благоприятствующих событий, так что всего имеется $(2 \text{ возможности для } \omega_1) \times \dots \times (2 \text{ возможности для } \omega_n) = 2^n$ возможностей. В число этих событий входят, в частности, достоверное $\Omega = \{\omega_1, \dots, \omega_n\}$, невозможное $\emptyset = \{ \}$.

Пусть задана функция $\xi = f(\omega)$, ставящая в соответствие исходу ω_i число $x_i = f(\omega_i)$. Эта функция, согласно § 3, является случайной величиной. Имеем

$$M\xi = \sum_{i=1}^n x_i \cdot \frac{1}{n} = \bar{x} \quad (2)$$

(черта над буквой обозначает среднее арифметическое).

Поскольку $I_A(\omega) = 1$ для ω , благоприятствующих событию A , $I_A(\omega) = 0$ для остальных ω , то

$$\sum_{i=1}^n I_A(\omega) = m, \quad (3)$$

где m — число исходов, благоприятствующих A . Отсюда получаем следующее выражение для вероятности:

$$P(A) = \frac{1}{n} \sum_{i=1}^n I_A(\omega_i). \quad (4)$$

Применяя классическое понятие вероятности к реальным событиям, обычно апеллируют к понятию физической симметрии. Так, при бросании игрального кубика ввиду его симметрии выпадение любого числа очков (1, 2, 3, 4, 5, 6) представляется одинаково возможным; следовательно, элементарные исходы можно считать равновероятными.

Пример 1. Среди N радиодеталей, имеющихся на складе, M неисправных. Случайно выбирается n деталей. Требуется найти вероятность $p_{N,M;n,m}$ того, что среди выбранных деталей попадается m неисправных.

Занумеруем детали числами от 1 до N . Вероятность выбора деталей с номерами i_1, \dots, i_m одна и та же для любых $1 \leq i_1 < i_2 < \dots < i_m \leq N$. Назовем (i_1, \dots, i_m) элементарным событием ω . Тогда число возможных элементарных событий равно C_N^m (число способов выбора m элементов из N). Определенному в условии примера событию благоприятствуют те и только те ω , для которых число неисправных деталей равно m , число исправных деталей равно $n - m$. Неисправные детали можно выбрать C_M^m способами, исправные C_{N-M}^{n-m} способами. При этом каждый способ выбора неисправных деталей можно сочетать с каждым способом выбора исправных. Отсюда находим, что число благоприятствующих данному событию исходов равно $C_M^m C_{N-M}^{n-m}$, а следовательно,

$$p_{N,M;n,m} = C_M^m C_{N-M}^{n-m} / C_N^n. \quad (5)$$

Пример 2. Имеется учебное пособие, изображенное на рис. 2 и называемое *доской Шеннарда*. На $n + 1$ уровнях в доску вбиты гвоздики; счет уровней осуществляется сверху вниз, от нулевого до n -го. Просветы между гвоздиками k -го уровня занумеруем слева направо от 0 до k . Шарик, точно проходящий между

соседними гвоздиками, опускается в верхний просвет. Найти вероятность p_{nk} того, что он выпадет в k -й просвет n -го уровня.

На каждом уровне от нулевого до $(n - 1)$ -го шарик «выбирает» одну из двух возможностей скатится с гвоздика влево или вправо. Таким образом, число возможных траекторий шарика равно 2^n . Естественно считать все эти траектории равновероятными.

Для того чтобы шарик выпал в k -й просвет n -го уровня, необходимо и достаточно, чтобы из n «шагов» было k шагов вправо и $n - k$ влево. Номера уровней, на которых делаются шаги вправо, полностью определяют благоприятствующую траекторию. Эти номера можно выбрать C_n^k способами. Итак,

$$p_{nk} = C_n^k \cdot 2^{-n}. \quad (6)$$

Пример 3. Схема испытания та же, что и в примере 2. Пусть v — номер просвета, в который выпал шарик. Доказать, что v — случайная величина, и найти Mv .

Очевидно, v — функция элементарного события, принимающая значения $0, 1, \dots, n$ ($v = f(\omega)$). Любому множеству B значений v соответствует некоторое множество A элементарных событий, но все такие множества, согласно предыдущему, являются событиями. Итак, v — случайная величина. Эта величина равна k на C_n^k значениях ω . Поэтому

$$\sum_{\omega} f(\omega) = 0 \cdot C_n^0 + 1 \cdot C_n^1 + 2 \cdot C_n^2 + \dots + n C_n^n. \quad (7)$$

Рассмотрим сумму k -го слева и k -го справа слагаемых и воспользуемся тем, что $C_n^{n-k} = C_n^k$.

$$k C_n^k + (n - k) C_n^{n-k} = n C_n^k = \frac{n}{2} C_n^k + \frac{n}{2} C_n^{n-k}.$$

Среднее слагаемое существует при четном n и равно $\frac{n}{2} C_n^{n/2} = \frac{n}{2} C_n^k, k = \frac{n}{2}$.

Итак, сумма (7) равна $\frac{n}{2} (C_n^0 + C_n^1 + \dots + C_n^n)$. По формуле (2), в которой нужно заменить n на 2^n (число элементарных событий), имеем

$$Mv = \frac{n}{2} (C_n^0 + C_n^1 + \dots + C_n^n) \cdot 2^{-n} = \frac{n}{2}. \quad (8)$$

Действительно, общее число 2^n элементарных событий можно перечислить так: C_n^0 благоприятствующих событию $\{v = 0\}$, C_n^1 благоприятствующих событию $\{v = 1\}$ и т. д., т. е.

$$2^n = C_n^0 + \dots + C_n^n. \quad (9)$$

Этот же результат получаем разложением бинома Ньютона $2^n = (1 + 1)^n$.

Пример 4. n писем, предназначенных n различным адресатам, вынуты из конвертов, перепутаны и вложены обратно в случайном порядке по одному в каждый конверт. Событие A_i состоит в том, что i -е письмо вложено в свой конверт. Найти $P(A_{i_1} \dots A_{i_r})$, где $1 \leq i_1 < \dots < i_r \leq n$.

Элементарное событие ω можно отождествить с подстановкой $\begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$, обозначив через j_i номер конверта, в который вложено i -е письмо. Число способов выбора j_1 равно n ; если j_1 выбрано, то j_2 можно выбрать лишь $n - 1$ способами, и т. д. Число равновероятных событий равно $n(n - 1) \dots 2 \cdot 1 = n!$.

Благоприятствующие подстановки те, для которых $j_k = i_k$ при $k = 1, \dots, r$. Таким образом, r чисел из n закреплены, остальные же можно переставлять произвольно; следовательно, число благоприятствующих элементарных событий равно $(n - r)!$. Итак,

$$P(A_{i_1} \dots A_{i_r}) = (n - r)!/n!. \quad (10)$$

Пример 5. Пользуясь условием примера 4, найти вероятность события B_n : хотя бы одно письмо вложено в свой конверт. Найти $\lim_{n \rightarrow \infty} P(B_n)$.

По формуле (3.8)

$$\begin{aligned} P(B_n) &= P\left(\bigcup_{i=1}^n A_i\right) = \sum_{r=1}^n (-1)^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq n} (n - r)!/n! = \\ &= \sum_{r=1}^n (-1)^{r-1}/r!, \end{aligned} \quad (11)$$

так как во внутренней сумме $C_n^r = n!/(r!(n - r)!)$ слагаемых и все они равны $(n - r)!/n!$. Наконец, заметив, что

$$\begin{aligned} e^{-1} &= \sum_{r=0}^{\infty} (-1)^r/r! = 1 + \sum_{r=1}^n (-1)^r/r! + \sum_{r=n+1}^{\infty} (-1)^r/r! = \\ &= 1 - P(B_n) + \sum_{r=n+1}^{\infty} (-1)^r/r!, \end{aligned}$$

причем последняя сумма бесконечно мала как остаток сходящегося ряда, получим

$$\lim_{n \rightarrow \infty} P(B_n) = 1 - e^{-1}. \quad (12)$$

Пример 6. Пользуясь условием примера 4, найти математическое ожидание числа ξ писем, адреса которых не перепутаны.

Имеем $\xi = I_{A_1} + \dots + I_{A_n}$. Отсюда по формуле (3.11)

$$M\xi = M I_{A_1} + \dots + M I_{A_n}.$$

Теперь применим формулу (3.6), а также (10) при $r = 1$:

$$M\xi = n(n - 1)!/n! = 1. \quad (13)$$

§ 5. Условные вероятности. Независимость. Формула Байеса

Рассмотрим следующий пример. Футбольной команде до конца чемпионата предстоит сыграть два матча; чтобы стать чемпионом (событие A), ей достаточно набрать два очка.

Возможные исходы испытания — двухзвенные цепочки, изображенные на рис. 3. Припишем каждой из них вероятность $1/9$. События, благоприятствующие A , отмечены затемненными кружочками. Имеем $P(A) = 6/9 = 2/3$. Пусть событие B состоит в том, что в первом матче команда не проиграла. Зная, что наступило событие B (результаты первого матча обведены кружком), естественно оценить вероятность A числом $5/6$. Это означает, что исходы, не благоприятствующие B , отбрасываются, а благоприятствующие переоцениваются.

Заметим, что

$$P(AB)/P(B) = \frac{5}{9} / \frac{2}{3} = \frac{5}{6}.$$

Пусть A, B — события, причем $P(B) > 0$. Условной вероятностью события A при условии B (точнее, при условии, что наступило событие B) называется число

$$P(A|B) = P(AB)/P(B). \quad (1)$$

Вероятность произведения событий A и B

$$P(AB) = P(B) P(A|B). \quad (2)$$

Если $P(A) > 0$, то, очевидно, можно записать и так:

$$P(AB) = P(A) P(B|A). \quad (3)$$

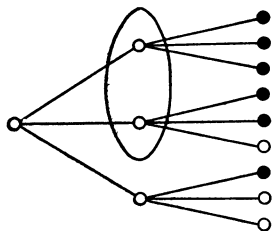


Рис. 3

Пусть A_1, \dots, A_n — события, причем $P(A_1 \dots A_{n-1}) > 0$. Тогда из (2) по индукции получаем

$$P(A_1 \dots A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 A_2) \dots P(A_n|A_1 \dots A_{n-1}). \quad (4)$$

Пусть ξ — случайная величина, B — событие, $P(B) > 0$. Условным математическим ожиданием ξ при условии B называется число

$$M(\xi|B) = M(\xi I_B)/P(B). \quad (5)$$

Используется также понятие *частичного математического ожидания*:

$$M(\xi, B) = M\xi I_B. \quad (6)$$

Тогда формула (5) запишется в виде

$$M(\xi|B) = M(\xi, B)/P(B). \quad (7)$$

Сопоставляя (7) с (1), находим, что условная вероятность — частный случай условного математического ожидания:

$$P(A|B) = M(I_A|B) = M(I_A I_B)/M I_B = M I_{AB}/M I_B. \quad (8)$$

Пример. Найти математическое ожидание числа ξ выпавших очков при бросании игрального кубика при условии, что это число — простое (событие B).

Событию B благоприятствуют три исхода: 2, 3, 5. Следовательно, $I_B(\omega)$ равно 0, 1, 1, 0, 1, 0 соответственно при ω , равном 1, 2, 3, 4, 5, 6. Значения $\xi I_A = \omega I_A$ равны соответственно 0, 2, 3, 0, 5, 0. Отсюда

$$M\xi I_A = \frac{1}{6} (0 + 2 + 3 + 0 + 5 + 0) = 10/6 = 5/3;$$

$$M(\xi|B) = \frac{5}{3} / \frac{1}{2} = \frac{10}{3}.$$

В отличие от $P(A|B)$ вероятность $P(A)$ называется *безусловной вероятностью* события A . Понятие безусловности относительно: согласно § 1, вероятность любого события относится к определен-

ным условиям опыта. Безусловность означает, что в данной задаче эти условия фиксированы.

Можно записать

$$P(A) = P(A | \Omega),$$

поскольку $A\Omega = A$, $P(\Omega) = 1$, а следовательно,

$$P(A) = P(A\Omega)/P(\Omega).$$

События A , B называются *независимыми*, если

$$P(AB) = P(A)P(B). \quad (9)$$

Сопоставив (9) с (2), приходим к следующему выводу. События A и B независимы в том и только в том случае, если вероятность одного из них равна нулю, либо условная вероятность одного при условии, что наступило другое, равна безусловной вероятности.

Полной группой событий называется конечная или бесконечная последовательность событий (A_n) , для которой

$$\bigcup_n A_n = \Omega; \quad A_n A_m = \emptyset, \quad n \neq m. \quad (10)$$

Иначе говоря, при любом ω происходит одно и только одно событие из (A_n) . Условие (10) равносильно условию

$$\sum_n I_{A_n} = 1, \quad \omega \in \Omega. \quad (11)$$

Пусть ξ — случайная величина, имеющая математическое ожидание. Тогда из (11)

$$M\xi = M\xi \sum_n I_{A_n} = \sum_n M\xi I_{A_n}, \quad (12)$$

или

$$M\xi = \sum_n M(\xi, A_n) = \sum_n P(A_n) M(\xi | A_n). \quad (13)$$

При бесконечной последовательности (A_n) возможность поменять местами знаки M и \sum в (12) следует из теоремы Лебега (см. «Дополнение») и того, что $M|\xi| < \infty$:

$$\left| M\xi \sum_{n \geq N} I_{A_n} \right| \leq M|\xi| \sum_{n \geq N} I_{A_n} \xrightarrow{N \rightarrow \infty} 0, \quad (14)$$

так как величина, стоящая после знака M , монотонно убывает к нулю при любом ω .

Формула (13) называется *формулой полного математического ожидания*. Положив $\xi = I_B$, где B — любое событие, из (13) получим формулу полной вероятности

$$P(B) = \sum_n P(A_n B) = \sum_n P(A_n) P(B | A_n). \quad (15)$$

Формулы (13) и (15) справедливы и в том случае, если некоторые $P(A_n) = 0$: для этих n можно определить $P(B | A_n)$ произвольным образом. Заметим также, что для справедливости формулы (15) достаточно, чтобы A_n были несовместны и $B \subset \bigcup_n A_n$.

Пусть $P(B) > 0$. Имеем

$$P(A_i|B) = P(A_i B)/P(B) = P(A_i) P(B|A_i)/P(B)$$

или, после подстановки формулы (15),

$$P(A_i|B) = \frac{P(A_i) P(B|A_i)}{\sum_n P(A_n) P(B|A_n)}. \quad (16)$$

Формула (16) называется *формулой Байеса (формулой гипотез)*. Ее интерпретируют следующим образом.

Пусть событие B может наступить лишь в случае, если наступило некоторое A_n (n -я гипотеза), причем гипотезы несовместны. Гипотезы имеют *априорные* (a priori — до опыта) *вероятности* $P(A_i)$: «до опыта» здесь означает ситуацию, когда не известно, наступило ли событие B . Вероятность $P(A_i|B)$ называется *апостериорной* (a posteriori — после опыта) *вероятностью* события A_i : опыт закончился наступлением события B .

Пример 1. При пристрелке цели на данном делении прицела получено два недолета и два перелета (событие B). Имеется 7 гипотез: A_0 — прицел установлен верно, A_{-6} , A_{-4} , A_{-2} — прицел занижен на 6, 4, 2 деления, A_2 , A_4 , A_6 — прицел завышен на 2, 4, 6 делений. Априорные вероятности этих гипотез равны соответственно 0,3; 0,05; 0,1; 0,2; 0,2; 0,1; 0,05 в порядке их перечисления. Условные вероятности события B при этих гипотезах равны соответственно 0,9; 0,01; 0,1; 0,4; 0,4; 0,1; 0,01. Требуется найти апостериорную вероятность того, что прицел установлен верно. По формуле Байеса

$$P(A_0|B) = \frac{0,3 \cdot 0,9}{0,3 \cdot 0,9 + 2(0,05 \cdot 0,01 + 0,1 \cdot 0,1 + 0,2 \cdot 0,4)} = \frac{0,27}{0,451} \approx 0,6.$$

В результате проведения опыта вероятность гипотезы A_0 повысилась примерно в 2 раза.

Пример 2. Датчик устройства, считывающего почтовые индексы, принимает черточку за пробел или наоборот с вероятностью 0,1. Вероятность ошибки в k данных элементах и отсутствия ошибки в остальных $9 - k$ элементах равна $0,1^k \cdot 0,9^{9-k}$. Априорная вероятность всех 10 цифр одна и та же. На выходе дат-

чика получен сигнал, соответствующий изображению $\overline{7}$ (событие B). Вычислить апостериорную вероятность того, что была записана цифра 7.

Число ошибочно считанных элементов при гипотезах 0, 1, ..., 9 получим, подсчитав число несовпадений элементов цифр с указанным изображением (*расстояние Хемминга*). Это соответственно 5, 4, 7, 3, 5, 4, 2, 2, 4, 6. Вероятности $P(B|A_i)$, вычисленные по приведенной формуле, равны соответственно

$$0,000007; 0,000059; 0,000000; 0,000531; 0,000007; \\ 0,000059; 0,004782; 0,004782; 0,000059; 0,000001.$$

Поскольку в данном случае все $P(A_i)$ равны и в формуле Байеса сокращаются, то $P(A_7|B)$ равна отношению числа 0,004 782 к сумме выписанных 10 чисел, откуда

$$P(A_7|B) \approx 0,5$$

(заметим, что такую же апостериорную вероятность имеет гипотеза A_6).

Пример 3. В цифровых устройствах применяется контроль на четность, устанавливающий факт наличия ошибок, если их число нечетно. Вероятности 0, 1, 2, 3, 4, 5 ошибок в схеме из 5 элементов равны соответственно 0,9; 0,09; 0,009; 0,0009; 0,00009; 0,00001. Найти условную вероятность отсутствия ошибок при условии (событие B), что число ошибок четно.

В данном случае $P(B | A_i)$ равна 1 при $i = 0, 2, 4$ и равна 0 при $i = 1, 3, 5$.
Имеем

$$P(A_0 | B) = 0,9 / (0,9 + 0,009 + 0,00009) \approx 0,99.$$

События A_1, \dots, A_n называются *независимыми* (более подробно: *независимыми в совокупности*), если для любого k и любого набора $1 \leq i_1 < \dots < i_k \leq n$ выполняются равенства

$$P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}). \quad (17)$$

Так, независимость четырех событий A, B, C, D сводится к системе равенств

$$\begin{aligned} P(AB) &= P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(AD) = P(A)P(D); \\ P(BC) &= P(B)P(C), \quad P(BD) = P(B)P(D), \quad P(CD) = P(C)P(D); \\ P(ABC) &= P(A)P(B)P(C), \quad P(ABD) = P(A)P(B)P(D), \quad P(ACD) = \\ &= P(A)P(C)P(D), \quad P(BCD) = P(B)P(C)P(D); \\ P(ABCD) &= P(A)P(B)P(C)P(D). \end{aligned}$$

Если события A_1, \dots, A_n независимы в совокупности, то равенства (17) сохраняются, когда некоторые из A_i заменить на \bar{A}_i , т. е., например, из независимости A, B, C, D следует независимость A, \bar{B}, \bar{C}, D . Принцип доказательства этого свойства виден на следующем примере. Пусть A, B независимы. Поскольку $\bar{A}B \cup AB = B$, причем AB и $\bar{A}B$ несовместны, то

$$\begin{aligned} P(\bar{A}B) &= P(B) - P(AB) = P(B) - P(A)P(B) = \\ &= (1 - P(A))P(B) = P(\bar{A})P(B), \end{aligned}$$

т. е. \bar{A} и B независимы.

События A_1, \dots, A_n называются *попарно независимыми*, если A_i и A_j независимы при любых $1 \leq i < j \leq n$. События $A_n, n \geq 1$, называются *независимыми* (попарно независимыми), если этим свойством обладают A_1, \dots, A_N при любом N .

Понятие независимости, подробнее называемой *статистической независимостью*, — одно из важнейших в теории вероятностей. Оно служит также основанием изучения реальных явлений методами теории вероятностей. Проверка независимости большого числа реальных событий путем статистического определения их вероятностей и проверки равенств (17) практически невозможна. Однако для того чтобы считать события независимыми, имеется веское физическое основание. Именно, события, определяемые пренебрежимо слабо физически связанными процессами, практически считают статистически независимыми. Примеры событий A и B , удовлетворяющих этому условию, приводятся в следующей таблице:

№ п/п	Событие A	Событие B
1	Безошибочность составления программы	Сбой ЭВМ при выполнении программы
2	Скорость ветра от 10 до 15 м/с	Масса снаряда выше номинальной
3	Исправность считывающего автомата	Пропуск символа в читаемом тексте
4	При бросании 10 монет первые 5 выпали гербами	Последние 5 монет выпали гербами

Однако бывает, что и в сильно физически связанных процессах и даже в одном и том же процессе происходят статистически независимые события. Например, при бросании игрального кубика события {число очков четно} и {число очков кратно трем} независимы.

Случайные величины ξ и η называются *независимыми*, если любые события, определяемые значениями этих величин, независимы:

$$P(\xi \in A, \eta \in B) = P(\xi \in A) P(\eta \in B), \quad (18)$$

где A, B — любые борелевские множества из \mathbb{R} .

Случайные величины ξ_1, \dots, ξ_n называются *независимыми*, если

$$P(\xi_i \in A_i, 1 \leq i \leq n) = P(\xi_1 \in A_1) \dots P(\xi_n \in A_n) \quad (19)$$

для любых борелевских множеств A_i из \mathbb{R} .

Если (ξ_i) — бесконечная последовательность случайных величин и для любого n случайные величины ξ_1, \dots, ξ_n независимы, то ξ_i называются *независимыми (независимыми в совокупности)*.

Дискретные случайные величины ξ_1, \dots, ξ_n независимы в том и только том случае, если

$$P(\xi_1 = x^{(1)}, \dots, \xi_n = x^{(n)}) = P(\xi_1 = x^{(1)}) \dots P(\xi_n = x^{(n)}), \quad (20)$$

где $x^{(k)}$ — любые числа, для которых $P(\xi_k = x^{(k)}) > 0$, т. е. возможные значения ξ_k . Из (20) получаем (19) путем суммирования по $x^{(k)} \in A_k$.

Задачи

1. Бросается два игральных кубика. Элементарное событие $\omega = (\omega_1, \omega_2)$, где ω_i — число выпавших очков i -го кубика. Перечислить элементарные события, благоприятствующие событию:

а) $\omega_1 + \omega_2$ — простое число;

б) $\max\{\omega_1, \omega_2\}$ делится на 4;

в) $\omega_1 \geq 2\omega_2$; г) $|\omega_1 - \omega_2| \leq 2$; д) $\omega_1 = \omega_2^2$.

2. Бросается три игральных кубика. Элементарные события зададим формулой $\omega = (\omega_1, \omega_2 + \omega_3)$, где ω_i — число выпавших очков i -го кубика. Определить, какие из следующих ниже выражений будут событиями в данном пространстве элементарных событий:

а) $\omega_1 + \omega_2 + \omega_3 = 4$; б) $\omega_1 + \omega_2 = 4$;

в) $\omega_1 > \omega_2$; г) $\omega_1^2 + \omega_2^2 + \omega_3^2 = 97$.

3. Пространство Ω содержит 5 элементарных событий. Сколько существует различных пар событий (A, B) , где B — следствие A ?

4. Пусть A, B, C, D — события. С помощью символов объединения, пересечения и взятия противоположного события записать следующие события:

а) из событий A, B, C, D наступило ровно одно;

б) два или три;

в) либо ни одного, либо все четыре.

5. Два игрока А и Б играют в следующую игру. Вначале игрок А игроку Б платит a копеек. Затем бросается пять монет. Если выпало четыре или пять гербов, игрок Б игроку А платит 32 копейки. Каким нужно взять a , чтобы математическое ожидание выигрыша каждого из игроков было нулевым?

6. ω_1, ω_2 — те же, что и в задаче 1. Доказать, что $\omega_1^2 + \omega_2^2$ — случайная величина, и найти ее математическое ожидание.

7. Грани двадцатигранника помечены числами 1, 2, 4, ..., 2^{19} . Вероятность выпадения каждой грани одна и та же. Найти математическое ожидание числа на выпавшей грани.

8. В сказочном городе любой местный житель дает на вопросы пришельца правильный ответ с вероятностью 0,6 и ложный с вероятностью 0,4. Чтобы найти правильный путь, можно опросить $2m + 1$ жителей и пойти в сторону, указываемую большинством из них. Каково наименьшее m , при котором вероятность ошибки не превосходит 0,01?

9. Найти условное математическое ожидание числа очков первого кубика при условии, что сумма числа очков первого и второго кубиков равна k ($k = 2, 3, \dots, 12$).

10. Вероятность попадания в цель при k -м выстреле равна $1 - 0,9^k$. Найти математическое ожидание числа попаданий при 7 выстрелах при условии, что при первых 4 выстрелах произошло хотя бы одно попадание.

11. Пять писем, предназначенных разным адресатам, вынуты из конвертов и снова вложены в случайном порядке по одному в каждый конверт. Независимы ли события {первое письмо вложено в свой конверт} и {второе письмо вложено в свой конверт}?

12. При хранении изделий каждое из них может в течение года проржаветь с вероятностью 0,01. Проверено 5 изделий с одинаковым временем хранения, которое может быть равно 1, 2, ..., 10 годам с равной вероятностью. Из проверенных изделий оказалось 2 проржавевших. Какова вероятность того, что время хранения изделий не менее 8 лет?

Глава 2

ПОСЛЕДОВАТЕЛЬНОСТЬ НЕЗАВИСИМЫХ ИСПЫТАНИЙ

§ 1. Независимые испытания с конечным числом исходов

Пусть производятся 1-е, 2-е, ... испытания с возможными исходами E_1, \dots, E_r . Если исходы различных испытаний независимы, то испытания называются *независимыми*. Так, при повторных бросаниях игрального кубика события {шестерка в первом бросании}, {двойка во втором бросании}, {двойка в третьем бросании}, {пятерка в четвертом бросании} независимы в совокупности. Рассмотрим лишь тот случай, когда вероятности исходов во всех испытаниях одинаковы, и обозначим через p_k вероятность k -го исхода в заданном заранее испытании. Итак, $p_k = P(E_k)$, $1 \leq k \leq r$. Поскольку исходы данного испытания несовместны, то

$$\sum_{k=1}^r p_k = \sum_{k=1}^n P(E_k) = P(\{E_1\} \cup \dots \cup \{E_r\}) = P(\Omega) = 1.$$

Пусть A_1, \dots, A_n — любые подмножества множества Ω . Тогда события {исход i -го испытания входит в A_i }, $1 \leq i \leq n$, незави-

симы. Принцип доказательства ясно виден на примере при $n = 2$:

$$\begin{aligned} & \mathbf{P} (\text{исход 1-го испытания входит в } A_1, \text{ исход 2-го} \\ & \text{испытания входит в } A_2) = \sum_{i \in A_1} \sum_{j \in A_2} \mathbf{P} (\text{исход 1-го} \\ & \text{испытания равен } E_i, \text{ исход 2-го испытания равен } E_j) = \\ & = \sum_{i \in A_1} \sum_{j \in A_2} p_i p_j = \sum_{i \in A_1} p_i \sum_{j \in A_2} p_j = \mathbf{P} (\text{исход 1-го испытания} \\ & \text{входит в } A_1) \mathbf{P} (\text{исход 2-го испытания входит в } A_2). \end{aligned}$$

Пусть e_i — исход i -го испытания, $\xi_i = f_i(e_i)$ — случайные величины. Если испытания независимы, то случайные величины, определяемые их результатами, независимы: если

$$B_i = \{e_i : f(e_i) \in A_i\},$$

то

$$\begin{aligned} \mathbf{P} (\xi_i \in A_i, 1 \leq i \leq n) &= \mathbf{P} (e_i \in B_i, 1 \leq i \leq n) = \\ &= \mathbf{P} (e_1 \in B_1) \dots \mathbf{P} (e_n \in B_n). \end{aligned}$$

Аналогично доказывается следующее утверждение. Если ξ_i — функция от исходов нескольких испытаний, причем номера этих испытаний не повторяются при различных i , то ξ_i — независимые случайные величины. Например, функция от первых $n - 1$ испытаний не зависит от функции n -го испытания.

§ 2. Последовательность испытаний Бернулли

Пусть каждое из независимых испытаний имеет два возможных исхода: успех с вероятностью p и неудача с вероятностью $q = 1 - p$. Эта последовательность называется *последовательностью испытаний Бернулли*. (Названия «успех» и «неудача» чисто условны и употребляются по традиции.) Закодировав успех единицей, а неудачу нулем, получим последовательность нулей и единиц, называемую *бернуллиевой последовательностью*; n -й член этой последовательности обозначим через ξ_n . Заметим, что ξ_n индикатор события {успех в n -м испытании}. Число успехов в испытаниях от 1-го до n -го обозначим через s_n . Тогда

$$s_n = \xi_1 + \dots + \xi_n. \quad (1)$$

На основании свойства, доказанного в § 1, ξ_i — независимые случайные величины. Ряд распределения случайной величины s_n обозначим так: $(k, p_n(k), 0 \leq k \leq n)$, т. е. $p_n(k) = \mathbf{P}(s_n = k)$.

Имеем *формулу Бернулли*:

$$p_n(k) = C_n^k p^k q^{n-k} = \frac{n!}{k! (n-k)!} p^k q^{n-k}. \quad (2)$$

Докажем ее тремя способами.

Д о к а з а т е л ь с т в о . 1. Событию $\{s_n = k\}$ благоприятствуют события $A(i_1, \dots, i_k)$, состоящие в том, что $\xi_j = 1$ при $j = i_1, \dots, j = i_k$; $\xi_j = 0$ при всех остальных j , $1 \leq j \leq n$, где i_1, \dots

..., i_k — различные целые числа от 1 до n . Например, событие $A(1, 2, \dots, k)$ состоит в том, что в первых k испытаниях был успех, в последующих $n - k$ испытаниях — неудача. Поскольку исходы испытаний независимы, то $P(A(i_1, \dots, i_k)) = p^k q^{n-k}$. Число способов выбора (i_1, \dots, i_k) равно C_n^k . Отсюда

$$p_n(k) = \sum_{i_1, \dots, i_k} p^k q^{n-k} = C_n^k p^k q^{n-k}.$$

2. Поскольку возможны лишь случаи $\xi_n = 1$, $\xi_n = 0$, то по формуле полной вероятности при $n \geq 2$ имеем

$$p_n(k) = P(s_n = k, \xi_n = 1) + P(s_n = k, \xi_n = 0).$$

Поскольку $s_n = s_{n-1} + \xi_n$, то

$$\{s_n = k, \xi_n = 1\} = \{s_{n-1} = k - 1, \xi_n = 1\}, \quad \{s_n = k, \xi_n = 0\} = \\ = \{s_{n-1} = k, \xi_n = 0\}.$$

События $\{s_{n-1} = k - 1\}$ и $\{\xi_n = 1\}$ независимы, события $\{s_{n-1} = k\}$ и $\{\xi_n = 0\}$ также независимы. Отсюда получаем равенство

$$p_n(k) = p_{n-1}(k - 1)p + p_{n-1}(k)q. \quad (3)$$

Введем обозначение

$$p_n(k) = u_n(k) p^k q^{n-k}.$$

Подставив его в формулу (3) и сократив на $p^k q^{n-k}$, получим

$$u_n(k) = u_{n-1}(k - 1) + u_{n-1}(k), \quad n \geq 2. \quad (4)$$

Если заданы $u_1(k)$, то рекуррентная формула (4) позволяет вычислить $u_2(k)$, $u_3(k)$ и т. д. При $n = 1$ имеем $p_1(1) = p$, $p_1(0) = q$, т. е. $u_1(0) = u_1(1) = 1$. В то же время, очевидно, $C_1^0 = C_1^1 = 1$ и C_n^k удовлетворяют тому же соотношению (4), а именно:

$$C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$$

(свойство, известное под названием «треугольник Паскаля»). Значит, по индукции получаем, что $u_n(k) = C_n^k$.

3. Обозначим $\pi_n(z) = p_n(0) + p_n(1)z + \dots + p_n(n)z^n$. Положив, что $p_n(k) = 0$ при $k > n$, можно записать

$$\pi_n(z) = \sum_{k=0}^{\infty} p_n(k) z^k. \quad (5)$$

Умножив обе части равенства (3) на z^k и просуммировав по k от 0 до ∞ , получим

$$\pi_n(z) = z\pi_{n-1}(z)p + \pi_{n-1}(z)q,$$

так как

$$p_{n-1}(-1) + zp_{n-1}(0) + z^2p_{n-1}(1) + \dots = z\pi_{n-1}(z)$$

ввиду того, что $p_{n-1}(-1) = 0$. Итак,

$$\pi_n(z) = (pz + q)\pi_{n-1}(z). \quad (6)$$

Поскольку $\pi_1(z) = p_1(0) + p_1(1)z = q + pz$, то из равенства (6), пользуясь биномом Ньютона, находим

$$\pi_n(z) = (pz + q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k} z^k. \quad (7)$$

Остается заметить, что вследствие равенства (5) число $p_n(k)$ является коэффициентом правой части (7) при z^k .

Положим в равенстве (7) $z = 1$. Тогда получим

$$(p + q)^n = 1 = \sum_{k=0}^n C_n^k p^k q^{n-k},$$

т. е. сумма вероятностей $p_n(k)$ по всем k от 0 до n равна 1.

Распределение случайной величины s_n называется *биномиальным распределением* с параметрами n, p , сама величина s_n называется *биномиальной* и символически обозначается $B(n, p)$.

Заметим, что если $B(n, p)$ и $B(m, p)$ независимы, то $B(n, p) + B(m, p) = B(n + m, p)$. Можно также записать

$$B(n, q) = n - B(n, p),$$

так как при замене p на q «успехи» и «неудачи» меняются местами.

§ 3. Теорема Пуассона (закон редких событий)

По формуле (2.2) вероятность того, что в n независимых испытаниях с вероятностью успеха p не будет ни одного успеха, определяется равенством

$$p_n(0) = q^n = (1 - p)^n. \quad (1)$$

Покажем, как изменяется выражение при больших n и малых p . Исследуем натуральный логарифм $p_n(0)$.

Справедливо неравенство

$$-p(1 + p) \leq \ln(1 - p) \leq -p, \quad 0 \leq p \leq 0,68. \quad (2)$$

По формуле Тейлора при $0 \leq p < 1$ $\ln(1 - p) = -p - p^2/2 \times \times (1 - \theta)^2$, $0 \leq \theta \leq p$, откуда следует оценка (2) при $0 \leq p < < 1 - 1/\sqrt{2} \approx 0,29$. На отрезке от 0,29 до 0,68 она доказывается численным методом. Потенцируя (2), умноженное на n , получаем неравенство

$$e^{-np(1+p)} \leq p_n(0) \leq e^{-np}, \quad p \leq 0,68. \quad (3)$$

Видим, что при большом n даже при весьма малых p число успехов может быть положительным с достаточно большой вероятностью.

Пример. Имеется $n = 10\,000$ микросхем, каждая из которых, независимо от других, может отказать в течение заданного времени с вероятностью $p = = 0,00025$. Вычислить вероятность $1 - p_n(0)$ того, что откажет хотя бы одна микросхема.

Имеем $np = 2,5$. Из неравенства (3) находим

$$1 - e^{-2,5} \leq 1 - p_n(0) \leq 1 - e^{-2,5 \cdot 1,000\,25}$$

или

$$0,917\,915 \leq 1 - p_n(0) \leq 0,917\,967.$$

Теорема Пуассона (закон редких событий). При $n \rightarrow \infty$, $np \rightarrow a$

$$p_n(k) \rightarrow e^{-a} a^k / k!, \quad k = 0, 1, 2, \dots \quad (4)$$

Доказательство. Предварительно исследуем поведение сомножителей правой части формулы Бернулли. При $n \rightarrow \infty$

$$\begin{aligned} C_n^k &= \frac{n!}{k! (n-k)!} = \frac{1}{k!} n(n-1) \dots (n-k+1) = \\ &= \frac{1}{k!} n^k \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) = \frac{1}{k!} n^k (1 + \alpha), \quad \alpha \rightarrow 0; \end{aligned} \quad (5)$$

$np = a(1 + \beta)$, $\beta \rightarrow 0$, откуда

$$p^k = n^{-k} a^k (1 + \beta)^k; \quad (6)$$

вследствие неравенства (3)

$$\exp\{-a(1 + \beta)(1 + a(1 + \beta)/n)\} \leq q^n \leq \exp\{-a(1 + \beta)\}. \quad (7)$$

Поскольку $\beta \rightarrow 0$, $n \rightarrow \infty$, то показатели экспонент стремятся к $(-a)$. А так как экспоненциальная функция непрерывна, то из (7) получаем

$$q^n = e^{-a} + \gamma, \quad \gamma \rightarrow 0. \quad (8)$$

Наконец,

$$q^{-k} = \left(1 - \frac{a}{n}(1 + \beta)\right)^{-k} = 1 + \delta, \quad \delta \rightarrow 0. \quad (9)$$

Перемножив равенства (5), (6), (8), (9) и перейдя к пределу, получим соотношение (4). Основной эффект здесь — сокращение на n^k при перемножении (5) и (6).

Теорема Пуассона используется при вычислении вероятностей редких событий (отказов сложной аппаратуры, преодоления большего числа препятствий и т. п.) в случае большого числа источников этих событий и малой вероятности события для каждого отдельного источника.

Пример. Управляющая вычислительная машина, собранная на микросхемах, отказывает лишь в случае отказа не менее 4 микросхем. Вычислить вероятность Q отказа вычислительной машины, пользуясь данными предыдущего примера. Имеем

$$\begin{aligned} Q &= \sum_{k \geq 4} p_n(k) = 1 - (p_n(0) + p_n(1) + p_n(2) + p_n(3)) \approx \\ &\approx 1 - e^{-a} (1 + a + a^2/2 + a^3/6), \end{aligned}$$

где $a = 2,5$. В результате вычислений имеем

$$Q \approx 0,2424.$$

Обобщенная теорема Пуассона. Пусть имеются независимые испытания; $p_i = 1 - q_i$ — вероятность успеха в i -м испытании. Если при $n \rightarrow \infty$

$$\sum_{i=1}^n p_i \rightarrow a, \quad (10)$$

$$\max\{p_1, \dots, p_n\} \rightarrow 0, \quad (11)$$

то вероятность $p_n(k)$ наступления k успехов в n испытаниях удовлетворяет соотношению (4).

Заметим, что p_i сами зависят от n , иначе (11) было бы возможно лишь при $p_i = 0$.

Доказательство проведем при $k = 2$ (в общем случае оно не сложнее, но более громоздко).

Вероятность того, что успехи наступят в i -м и j -м испытаниях ($i \neq j$) и не наступят во всех остальных, равна $u_i u_j R$, где $u_i = p_i/q_i$, $R = q_1 \dots q_n$.

По формуле полной вероятности

$$p_n(2) = R \sum_{i < j} u_i u_j.$$

Имеем

$$(u_1 + \dots + u_n)^2 = \sum_{i < j} u_i u_j + \sum_{i \geq j} u_i u_j + \sum_i u_i^2. \quad (12)$$

Первые две суммы в правой части равенства (12) совпадают, так что

$$\sum_{i < j} u_i u_j = \frac{1}{2} (u_1 + \dots + u_n)^2 - \frac{1}{2} (u_1^2 + \dots + u_n^2). \quad (13)$$

Из условия (11) имеем $p_i \leq \varepsilon$, где ε — бесконечно малая величина, а следовательно,

$$\sum p_i \leq \sum u_i \leq \sum p_i / (1 - \varepsilon);$$

так как к тому же $\sum p_i \rightarrow a$, то $\sum u_i \rightarrow a$, а значит, первое слагаемое правой части (12) стремится к $a^2/2$. Что касается второго слагаемого, то вследствие неравенства $u_i^2 \leq p_i^2 / (1 - \varepsilon)^2 \leq p_i \varepsilon / (1 - \varepsilon)^2$ имеем

$$\sum u_i^2 \leq \varepsilon (1 - \varepsilon)^{-2} \sum p_i \sim \varepsilon (1 - \varepsilon)^{-2} a \rightarrow 0.$$

Итак,

$$\sum_{i < j} u_i u_j \rightarrow a^2/2. \quad (14)$$

Оценим $\ln R = \sum_i \ln(1 - p_i)$. Из равенства (2) имеем

$$-\sum p_i (1 + \varepsilon) \leq \ln R \leq -\sum p_i;$$

приняв во внимание (10), из этого соотношения получаем, что $\ln R \rightarrow -a$, $R \rightarrow e^{-a}$. Этот последний результат в сочетании с (14) приводит к соотношению $p_n(2) \rightarrow e^{-a} a^2/2$ — частному случаю утверждения теоремы при $k = 2$.

Пример. Во время выполнения задания 150 блоков системы находятся в нагруженном режиме, 250 — в облегченном, 125 — в слабо нагруженном режиме. Вероятность отказа отдельного блока равна соответственно 0,01; 0,005; 0,002. Сколько нужно взять запасных блоков, чтобы заменить отказавшие блоки с вероятностью, не меньшей 0,99 (возможность отказа блоков, поставленных вместо отказавших, можно для простоты пренебречь).

Имеем

$$\sum p_i = 150 \cdot 0,01 + 250 \cdot 0,005 + 125 \cdot 0,002 = 3.$$

Последовательно вычисляем, пользуясь обобщенной теоремой Пуассона, предельные значения вероятностей $p_n(k)$ для $k = 0, 1, 2, \dots$ при $a = 3$ и суммируем их до тех пор, пока сумма станет не меньше 0,99. Имеем следующую таблицу:

k	0	1	2	...	7	8
$e^{-a} a^k / k!$	0,0498	0,1494	0,2240	...	0,0216	0,0081
$\sum_{i \leq k} e^{-a} a^i / i!$	0,0498	0,1992	0,4234	...	0,9880	0,9961

Видим, что минимальное число запасных блоков, удовлетворяющее заданному условию, равно 8.

Случайная величина ξ , имеющая ряд распределения $(k, e^{-a} \times \times a^k / k!, k = 0, 1, 2, \dots)$, называется *пуассоновской*, или *распределенной по закону Пуассона* с параметром a . Символическое обозначение

$$\xi = \text{Poiss}(a).$$

Вероятности $p_k(a) = e^{-a} a^k / k!$, суммы $P_k(a) = p_0(a) + p_1(a) + \dots + p_k(a)$, $Q_k(a) = p_k(a) + p_{k+1}(a) + \dots$ табулированы; они легко вычисляются также на микрокалькуляторе. Вычисление последовательных значений $p_n(k)$ легко производить по рекуррентной формуле

$$p_k(a) = p_{k-1}(a) a / k, \quad k \geq 1. \quad (15)$$

Приближение $p_n(k)$ или $p_n(n - k)$ с помощью вероятностей $p_n(a)$ называется *пуассоновской аппроксимацией*.

§ 4. Исследование поведения биномиальных вероятностей

Пусть p — фиксированное число из интервала $(0; 1)$, $n \rightarrow \infty$. Исследуем, как изменяются биномиальные вероятности $p_n(k) = C_n^k p^k q^{n-k}$. Обозначим $\alpha = k/n$, $\beta = 1 - \alpha$. Тогда $p_n(\alpha n)$ — вероятность того, что частота успехов в n независимых испытаниях в точности равна α . Нашей задачей и будет исследование поведения вероятности

$$p_n(\alpha n) = \frac{n!}{(\alpha n)! (\beta n)!} p^{\alpha n} q^{\beta n}. \quad (1)$$

Из математического анализа известна формула Стирлинга:

$$n! \sim \sqrt{2\pi n} n^n e^{-n}, \quad n \rightarrow \infty, \quad (2)$$

или с оценкой погрешности

$$n! \sim \sqrt{2\pi n} n^n e^{-n+\theta/12n}, \quad 0 < \theta < 1. \quad (3)$$

Из формулы (2)

$$n! / (\alpha n)! (\beta n)! \sim \frac{1}{\sqrt{2\pi\alpha\beta n}} \left(\frac{1}{\alpha}\right)^{\alpha n} \left(\frac{1}{\beta}\right)^{\beta n}, \quad (4)$$

а следовательно,

$$p_n(\alpha n) \sim \frac{1}{\sqrt{2\pi\alpha\beta n}} \left(\frac{p}{\alpha}\right)^{\alpha n} \left(\frac{q}{\beta}\right)^{\beta n}. \quad (5)$$

Введем функцию

$$g(\alpha) = \alpha \ln p - \alpha \ln \alpha + \beta \ln q - \beta \ln \beta. \quad (6)$$

Тогда

$$p_n(\alpha n) \sim \frac{1}{\sqrt{2\pi\alpha\beta n}} e^{ng(\alpha)}, \quad (7)$$

и весь вопрос сводится к исследованию поведения $g(\alpha)$. Имеем $g(p) = 0$; $g'(\alpha) = \ln(p\beta/q\alpha)$ и, в частности, $g'(p) = 0$; $g''(\alpha) = -1/\alpha\beta < 0$. Следовательно, $g(\alpha)$ имеет максимум в точке $\alpha = p$, возрастает при $\alpha < p$ и убывает при $\alpha > p$. Мы видим, что $g(\alpha) < 0$ при всех α , за исключением $\alpha = p$.

Формулой (7) можно пользоваться для практического расчета биномиальных вероятностей. Относительная погрешность формулы (7) вследствие (3) при больших k , $n - k$ по абсолютной величине не превосходит значения

$$\Delta_{nk} = \frac{1}{12k(n-k)}. \quad (8)$$

Пример. Вероятность рождения мальчика равна $1/2$. Вычислить вероятность рождения ровно половины и ровно 55 % мальчиков среди n родившихся детей при n , равном 10^2 , 10^4 , 10^6 , 10^8 .

Значения α , соответствующие 50 и 55 %, равны соответственно $\alpha_0 = 0,5$, $\alpha_1 = 0,55$. Как выше было указано, $g(\alpha_0) = 0$. В данном случае $g(\alpha) = -\ln 2 - \alpha \ln \alpha - \beta \ln \beta$, откуда

$$g(\alpha_1) = -\ln 2 - 0,55 \ln 0,55 - 0,45 \ln 0,45 \approx -0,005\,008.$$

По формуле (5) при указанных значениях n имеем

$$\begin{aligned} p_n(\alpha_0 n) &\approx \frac{2}{\sqrt{2\pi n}} \approx \frac{0,798}{\sqrt{n}} = 7,98 \cdot 10^{-2}; \quad 7,98 \cdot 10^{-3}; \quad 7,98 \cdot 10^{-4}; \\ &\quad 7,98 \cdot 10^{-5}; \\ p_n(\alpha_1 n) &\approx \frac{1}{\sqrt{2\pi \cdot 0,55 \cdot 0,45 n}} e^{-0,005\,008 n} \approx \frac{0,802}{\sqrt{n}} e^{-0,005\,008 n} = 0,0486; \\ &\quad 1,43 \cdot 10^{-24}; \quad \text{порядка } 10^{-2178}; \quad \text{порядка } 10^{-217\,498}. \end{aligned}$$

§ 5. Вероятности отклонений. Закон больших чисел

Рассмотрим отношение $p_n(k+1)$ к $p_n(k)$:

$$\frac{p_n(k+1)}{p_n(k)} = \frac{n! p^{k+1} q^{n-k-1}}{(k+1)! (n-k-1)!} / \frac{n! p^k q^{n-k}}{k! (n-k)!} = \frac{(n-k)p}{(k+1)q}. \quad (1)$$

Это отношение убывает при возрастании k . При $k = \alpha n$ оно меньше $\beta p/\alpha q$; значит,

$$p_n(k+1)/p_n(k) \leq \beta p/\alpha q, \quad k \geq \alpha n. \quad (2)$$

Положим $\alpha = p + \varepsilon$, где $\varepsilon > 0$, $p + \varepsilon < 1$. Тогда

$$\beta p/\alpha q = (pq - \varepsilon p)/(pq + \varepsilon q) = \theta < 1. \quad (3)$$

Из неравенства (2) последовательно получаем

$$p_n(k+1) \leq p_n(\alpha n) \theta, \quad p_n(k+2) = p_n(\alpha n) \frac{p_n(\alpha n+1)}{p_n(\alpha n)} \times \\ \times \frac{p_n(\alpha n+2)}{p_n(\alpha n+1)} \leq p_n(\alpha n) \theta^2$$

и т. д.

Событие $\{s_n \geq \alpha n\}$, где $s_n = B(n, p)$, есть объединение событий $\{s_n = k\}$ по $k \geq \alpha n$. Отсюда

$$P(s_n \geq (p + \varepsilon)n) < p_n((p + \varepsilon)n)(1 + \theta + \theta^2 + \dots) = \\ = p_n((p + \varepsilon)n)/(1 - \theta) = p_n((p + \varepsilon)n)(pq + \varepsilon q)/\varepsilon. \quad (4)$$

Аналогично при $0 < \varepsilon < p$ выводится неравенство

$$P(s_n \leq (p - \varepsilon)n) < p_n((p - \varepsilon)n)(pq + \varepsilon p)/\varepsilon. \quad (5)$$

Правые части соотношений (4) и (5) дают не только верхнюю оценку левых частей, но и эквивалентны им при целых $(p \pm \varepsilon)n$: в формуле (2) при ограниченном $k - \alpha n$ и $n \rightarrow \infty$ имеем $p_n(k+1)/p_n(k) \sim \beta p/\alpha q$, так что вначале идут слагаемые, приблизительно равные членам геометрической прогрессии, затем — меньшие слагаемые, но их суммой уже можно пренебречь.

Пример. Оценить вероятность того, что из 100 родившихся детей не менее 55 мальчиков.

Вероятность рождения ровно 55 мальчиков вычислена в примере предыдущего параграфа и равна 0,0486. По формуле (4) это число нужно умножить на $(pq + \varepsilon q)/\varepsilon = (0,55 \cdot 0,45 + 0,05 \cdot 0,45)/0,05 = 5,4$. В результате получаем $P(s_{100} \geq 55) \leq 0,0486 \cdot 5,4 \approx 0,262$. Из (4.7) следует, что $p_n((p \pm \varepsilon)n)$ убывают к нулю экспоненциально быстро при $n \rightarrow \infty$. С учетом этого факта из (4), (5) следует, что

$$P(s_n \geq (p + \varepsilon)n) \leq c_0 \rho^n, \quad P(s_n \leq (p - \varepsilon)n) \leq c_0 \rho^n, \quad (6)$$

где c_0 и ρ — постоянные, зависящие от p и ε , $0 < \rho < 1$. Поскольку

$$P(|s_n - pn| \geq \varepsilon n) = P(s_n \geq (p + \varepsilon)n) + P(s_n \leq (p - \varepsilon)n),$$

то, положив $c = 2c_0$, из (6) получим оценку

$$P\left(\left|\frac{s_n}{n} - p\right| \geq \varepsilon\right) \leq c \rho^n. \quad (7)$$

Заметим, что s_n/n есть частота успеха в n независимых испытаниях. Разность $(s_n/n) - p$ называют отклонением, $|(s_n/n) - p|$ — абсолютным отклонением частоты от вероятности. Таким образом, оценку (7) можно словесно выразить так: вероятность того, что абсолютное отклонение частоты от вероятности больше или равно ε , при $n \rightarrow \infty$ убывает к нулю экспоненциально быстро при любом заданном $\varepsilon > 0$.

В теории и практике большую роль играет закон больших чисел. Под этим названием объединяется целый ряд теорем о том, что при тех или иных условиях абсолютное отклонение среднего арифметического случайных величин от их математического ожидания может быть не больше ε лишь с бесконечно малой вероятностью при $n \rightarrow \infty$ вероятностью. Простейшая из этих теорем называется законом больших чисел в форме Бернулли и выражается двумя эквивалентными

соотношениями при $\varepsilon > 0$:

$$P\left(\left|\frac{s_n}{n} - p\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0, \quad (8)$$

$$P\left(\left|\frac{s_n}{n} - p\right| < \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1. \quad (9)$$

Первое из соотношений следует из оценки (7); (9) следует из (8) (стоит лишь перейти к противоположному событию).

Закон больших чисел служит теоретическим обоснованием принципиальной возможности вычисления вероятности по частоте сколь угодно высокой точностью при достаточно большом числе наблюдений.

Пример предыдущего параграфа иллюстрирует значения вероятности 5 %-го отклонения частоты от вероятности при больших значениях n . Если взять не 5 %, а, например, 0,01 %, то качественный эффект тот же — при достаточно больших n вероятность отклонения, большего данного числа, становится пренебрежимо малой.

Теорема Бореля (у с и л е н н ы й з а к о н б о л ь ш и х ч и с е л)

$$P\left(\frac{s_n}{n} \xrightarrow{n \rightarrow \infty} p\right) = 1. \quad (10)$$

Объясним соотношение (10). Событие $\left\{\frac{s_n}{n} \rightarrow p\right\}$ равно произведению событий $A_r = \left\{\left|\frac{s_n}{n} - p\right| < \frac{1}{r}\right\}$ начиная с некоторого n . Поскольку $A_{r+1} \subset A_r$, то установив, что $P(A_r) = 1$ для всех r , по аксиоме непрерывности (гл. 1) получим, что

$$P\left(\frac{s_n}{n} \rightarrow p\right) = P\left(\bigcap_{r=1}^{\infty} A_r\right) = \lim_{r \rightarrow \infty} P(A_r) = 1.$$

Итак, достаточно установить, что $P(A_r) = 1$.

Имеем $A_r = \bigcap_{N=1}^{\infty} \left\{\left|\frac{s_n}{n} - p\right| < \frac{1}{r}, n \geq N\right\}$. На этот раз получим объединение событий, N -е из которых включается в $(N+1)$ -е. Снова по аксиоме непрерывности находим

$$P(A_r) = \lim_{N \rightarrow \infty} P\left(\left|\frac{s_n}{n} - p\right| < \frac{1}{r}, n \geq N\right). \quad (11)$$

Согласно неравенству (7),

$$\begin{aligned} P\left(\bigcap_{n=N}^{\infty} \left\{\left|\frac{s_n}{n} - p\right| \geq \frac{1}{r}\right\}\right) &\leq \sum_{n=N}^{\infty} P\left(\left|\frac{s_n}{n} - p\right| \geq \frac{1}{r}\right) \leq \\ &\leq c \sum_{n=N}^{\infty} \rho^n = c\rho^N/(1-\rho), \quad c = c(r), \quad \rho = \rho(r) < 1. \end{aligned}$$

Отсюда, переходя к противоположному событию, получим

$$P\left(\left|\frac{s_n}{n} - p\right| < \frac{1}{r}, n \geq N\right) \geq 1 - cr^N/(1 - p) \xrightarrow{n \rightarrow \infty} 1$$

и из (11) следует, что $P(A_r) = 1$, что и требовалось доказать.

В данном случае вероятностная мера может быть определена только на континуальном множестве элементарных событий: таковыми являются последовательности $\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots)$, где ε_n равно 1 или 0 в зависимости от успеха или неудачи в n -м испытании. Сопоставив последовательности $\bar{\varepsilon}$ число $\omega = \sum_{n=1}^{\infty} 2^{-n} \varepsilon^n$ (по ω можно вос-

становить $\bar{\varepsilon}$, если только ω не двоично-рационально: для десятичных чисел это случай типа $0,500\dots = 0,499\dots$), получим пространство Ω в виде отрезка $\{0 \leq \omega \leq 1\}$. В частности, если вероятностная мера на этом отрезке задается условием, что вероятность любого отрезка равна его длине, то ε_i независимы и $P(\varepsilon_i = 1) = P(\varepsilon_i = 0) = 1/2$. Действительно, двоичные числа, у которых первые n цифр фиксированы, занимают интервал длины 2^{-n} . При вероятности успеха, отличной от $1/2$, построение вероятностной меры на отрезке делается с помощью иного соответствия между $\bar{\varepsilon}$ и ω . Так, событию $\{\varepsilon_1 = 0\}$ сопоставляется интервал $(0; q)$, событию $\{\varepsilon_1 = 1\}$ — интервал $(q; 1)$. Каждый из полученных двух интервалов снова делится в отношении $q : p$ и т. д.

§ 6. Теорема Муавра—Лапласа

Из формулы (4.7) видно, что при $n \rightarrow \infty$ вероятности $p_n(\alpha n)$ экспоненциально быстро убывают при любом фиксированном α , кроме $\alpha = p$. В этом последнем случае $p_n(pn) \sim 1/\sqrt{2\pi npq}$ ¹. Важную роль в практическом и теоретическом отношении играют уклонения частоты от вероятности порядка $1/\sqrt{n}$: для них $p_n(k)$ имеют порядок $1/\sqrt{n}$. Назовем *нормированным уклонением* величину

$$x_k = (\alpha - p) \sqrt{n/pq} = (k - np)/\sqrt{npq}, \quad (1)$$

где k — возможные значения случайной величины s_n .

Теорема Муавра — Лапласа. При $n \rightarrow \infty$

$$p_n(k) \sim \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2} \quad (2)$$

равномерно относительно x_k из интервала $|x_k| < T$, где T — любое фиксированное число.

Приближение $p_n(k)$ выражением, задаваемым правой частью (2), называется *нормальной аппроксимацией*.

¹ Эффект дробности pn пренебрежимо мал: поскольку $g'(p) = 0$, то смещение α на $1/n$ от значения $\alpha = p$ ведет к смещению g порядка $1/n^2$, а следовательно, к смещению порядка $1/n$ показателя экспоненты в (4.7).

Иначе теорему Муавра — Лапласа можно сформулировать так: *относительная погрешность нормальной аппроксимации $p_n(k)$ стремится к нулю при $n \rightarrow \infty$ в любом фиксированном заранее интервале нормированных отклонений.*

Доказательство. Формулу (1) можно записать так:

$$\alpha - p = x_k \sqrt{pq/n}. \quad (3)$$

В то же время, как уже отмечалось, $g(p) = 0$, $g'(p) = 0$, $g''(\alpha) = -1/\alpha\beta$. Поэтому по формуле Тейлора

$$g(\alpha) = -(\alpha - p)^2/2p'q', \quad (4)$$

где p' — промежуточное значение между p и α , $q' = 1 - p'$. Подставив в (4) значение $\alpha - p$ из (3) и умножив результат на n , получим

$$ng(\alpha) = -x_k^2 pq/2p'q'. \quad (5)$$

Вследствие (3), $\alpha - p$ равномерно стремится к нулю, а следовательно, $p'q'$ равномерно стремится к pq . Отсюда заключаем, что $ng(\alpha)$ равно $-x_k^2/2$ с точностью до величины, равномерно стремящейся к нулю при условии $|x_k| < T$. Таким образом, $e^{ng(\alpha)} \sim e^{-x_k^2/2}$. Согласно формуле (5.7), замечаем, что $\alpha\beta \sim pq$.

При фиксированном, хотя и весьма большом, n нормальная аппроксимация теряет точность при больших $|x_k|$. Сделаем приближенную оценку погрешности. По формуле Тейлора

$$g(\alpha) \approx (\alpha - p)^2/2 + (\alpha - p)^3(p - q)/6p^2q^2 + (\alpha - p)^4(1 - 3pq)/12p^3q^3 \quad (6)$$

Подставив вместо $\alpha - p$ значение из (3) и умножив результат на n , получим погрешность показателя экспоненты:

$$\Delta = ng(\alpha) + \frac{x_k^2}{2} \approx \frac{x_k^3(p - q)}{6\sqrt{npq}} + \frac{x_k^4(1 - 3pq)}{12npq}. \quad (7)$$

Если $p \neq q$, то второе слагаемое правой части (7) пренебрежимо мало по сравнению с первым слагаемым; условие $|\Delta| < \varepsilon$ можно приближенно записать в виде

$$|x_k^3| \leq 6\varepsilon \sqrt{npq}/(p - q); \quad (8)$$

если $p = q$, т. е. $p = 1/2$, то остается лишь второе слагаемое, откуда получаем

$$x_k^4 \leq 12\varepsilon npq/(1 - 3pq). \quad (9)$$

В обоих случаях при данных условиях замена $\alpha\beta$ на pq в выражении $\sqrt{2\pi n\alpha\beta}$ внесет пренебрежимо малую погрешность.

При практических расчетах следует контролировать погрешность, вычисляя ее по формуле (7) для максимального $|x_k|$, необходимого в расчете; если точность неудовлетворительная, следует переходить к универсальной оценке (5.7). Надо иметь в виду и то, что при

$p \rightarrow 0$ и $p \rightarrow 1$ нормальная аппроксимация также ухудшается, но на этот случай имеется пуассоновская аппроксимация (см. § 3).

Пример. Пусть $n = 1200$, $p = 0,25$. Тогда $np = 300$, $q = 0,75$, $\sqrt{npq} = 15$. Для 9 равноотстоящих значений k сведем в таблицу значения $p_n(k)$, вычисленные по формуле (5.7), их нормальные аппроксимации, а также значения отношений нормальных аппроксимаций к $p_n(k)$:

k	x_k	$p_n(k)$	α	Нормальные аппроксимации	Значения отношений
180	-8	$9,93 \cdot 10^{-18}$	0,15	$3,37 \cdot 10^{-16}$	33,9
210	-6	$1,08 \cdot 10^{-10}$	0,175	$4,05 \cdot 10^{-10}$	3,76
240	-4	$6,47 \cdot 10^{-6}$	0,2	$8,92 \cdot 10^{-6}$	1,38
270	-2	$3,580 \cdot 10^{-3}$	0,225	$3,599 \cdot 10^{-3}$	1,0054
300	0	0,0265961	0,25	0,0265961	1,0000
330	2	$3,64 \cdot 10^{-3}$	0,275	$3,599 \cdot 10^{-3}$	0,9883
360	4	$1,16 \cdot 10^{-5}$	0,3	$8,92 \cdot 10^{-6}$	0,769
390	6	$1,04 \cdot 10^{-9}$	0,325	$4,05 \cdot 10^{-10}$	0,389
420	8	$3,05 \cdot 10^{-15}$	0,35	$3,37 \cdot 10^{-16}$	0,11

§ 7. Интегральная теорема Лапласа

Исследуем предельное поведение при $n \rightarrow \infty$ вероятности $P_n(A, B)$ события, состоящего в том, что $A \leq s_n < B$, где $A < B$ — заданные числа.

Использование полуинтервала $A \leq s_n < B$ вместо интервала $A < s_n < B$ или отрезка $A \leq s_n \leq B$ удобнее тем, что, в отличие от них, объединение примыкающих полуинтервалов не содержит общей точки и само является полуинтервалом, так что по формуле полной вероятности при $A \leq B \leq C$

$$P_n(A, B) + P_n(B, C) = P_n(A, C). \quad (1)$$

По формуле полной вероятности

$$P_n(A, B) \leq \sum_{A \leq k < B} p_n(k). \quad (2)$$

Неравенство $A \leq k < B$ равносильно неравенству

$$a = (A - np)/\sqrt{npq} \leq (k - np)/\sqrt{npq} < (B - np)/\sqrt{npq} = b.$$

Поэтому (2) можно записать в виде

$$P_n(A, B) = \sum_{a \leq x_k < b} p_n(np + x_k \sqrt{npq}). \quad (3)$$

По теореме Муавра — Лапласа

$$p_n(np + x_k \sqrt{npq}) \sim \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2},$$

а следовательно, и

$$P_n(A, B) \sim \frac{1}{\sqrt{2\pi}} \sum_{a \leq x_k < b} e^{-x_k^2/2} \frac{1}{\sqrt{npq}}. \quad (4)$$

Поскольку $\Delta x = x_{k+1} - x_k = (k+1 - np)/\sqrt{npq} - (k - np)/\sqrt{npq} = 1/\sqrt{npq}$, то правая часть (4) есть интегральная сумма интеграла функции

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad (5)$$

называемой *функцией Гаусса*. Правда, крайний слева и крайний справа интервалы при этом приходится увеличить или уменьшить, но не более, чем на Δx . Поскольку $\varphi(x) < 1$, то вносимая этим погрешность пренебрежимо мала. В пределе получаем

$$P_n(A, B) = P\left(a \leq \frac{s_n - np}{\sqrt{npq}} < b\right) \xrightarrow{n \rightarrow \infty} \int_a^b \varphi(x) dx. \quad (6)$$

Можно доказать, что стремление к пределу равномерно относительно a, b . Это и составляет содержание *интегральной теоремы Лапласа*. В частности, положив $a = -\infty$, $b = x$, получим утверждение

$$P(s_n - np/\sqrt{npq} < x) \xrightarrow{n \rightarrow \infty} \int_{-\infty}^x \varphi(t) dt. \quad (7)$$

Функция

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (8)$$

называется *интегральной функцией Гаусса*. Эта функция не выражается через элементарные; она табулирована и вычисляется с помощью стандартной процедуры на любой ЭВМ. Приведем таблицу значений интегральной функции Гаусса, используемых при решении примеров:

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
$-\infty$	0,0000	-1,0	0,1587	1,5	0,9332
-3,0	0,0014	-0,5	0,3085	2,0	0,9772
-2,5	0,0062	0,0	0,5000	2,5	0,9938
-2,0	0,0228	0,5	0,6915	3,0	0,9986
-1,5	0,0668	1,0	0,8413	∞	1,0000

Имеем

$$\int_a^b \varphi(x) dx = \Phi(b) - \Phi(a).$$

Поскольку при больших n все $p_n(k) \leq 1/\sqrt{2\pi npq}$, то в приближенных расчетах по формуле (6) выражение « $a \leq \dots$ » можно заменить на выражение « $a < \dots$ », а выражение « $\dots < b$ » — на

«... ≤ b». В частности,

$$P(|(s_n - np)/\sqrt{npq}| \leq z) \xrightarrow{n \rightarrow \infty} \int_{-z}^z \varphi(x) dx = \\ = \Phi(z) - \Phi(-z) = 2\Phi(z) - 1. \quad (9)$$

Пример 1. С борта самолета передано кодированное сообщение, состоящее из 150 сигналов, в условиях сильных помех. Вероятность искажения отдельного сигнала составляет 0,4, причем искажения различных сигналов независимы. Оценить вероятность f того, что искаженными окажутся не более 72 сигналов.

Имеем $n = 150$, $p = 0,4$, $q = 0,6$, $\sqrt{npq} = 6$; нормированное уклонение $(72 - 150 \cdot 0,4) : 6 = 2$. Отсюда

$$f \approx \Phi(2) \approx 0,977.$$

Пример 2. Управляющая ЭВМ обслуживает 1000 производственных установок. Вероятность того, что данная установка в данный момент времени требует обслуживания, равна 0,12. Сколько каналов обслуживания r необходимо иметь, чтобы вероятность наличия канала для каждой установки, требующей обслуживания в данный момент времени, была не меньше 0,99?

Имеем $n = 1000$, $p = 0,12$, $np = 120$, $\sqrt{npq} \approx 10,3$;

$$P(s_n \leq r) = P((s_n - np)/\sqrt{npq} \leq (r - np)/\sqrt{npq}) \approx \Phi((r - np)/\sqrt{npq}) \geq 0,99.$$

По таблице находим z , для которого $\Phi(z) \geq 0,99$. Возьмем значение $z = 2,5$, так как $\Phi(2,5) \approx 0,994$. Отсюда получаем неравенство

$$\frac{r - np}{\sqrt{npq}} \geq z, \quad \frac{r - 120}{10,3} \geq 2,5, \quad r \geq 120 + 2,5 \cdot 10,3 = 145,75.$$

Следовательно, условие задачи удовлетворяется при $r \geq 146$.

Пример 3. Проектируется система из $n = 5000$ узлов связи, способная исправно работать при отказе любых $k = 200$ из них. Допустимая вероятность отказа системы 0,0015. Чему равна максимально допустимая вероятность p отказа отдельного узла (отказы узлов независимы)?

По таблице находим $z = 3$, так как $\Phi(3) \approx 0,9986 > 1 - 0,0015$. Имеем $(200 - 5000p)/\sqrt{5000pq} = 3$. Возведя это равенство в квадрат, получим квадратное уравнение $25045p^2 - 2045p + 40 = 0$, имеющее два корня: $p_1 = 0,0324791$, $p_2 = 0,041739$. При $p = p_1$ имеем $(k - np)/\sqrt{npq} = 3$, при $p = p_2$ имеем $(k - np)/\sqrt{npq} = -3$, что не соответствует условию задачи. Итак, $p \approx 0,03$. Можно проверить, что при $p < p_1$ вероятность отказа меньше $\Phi(3)$.

§ 8. Испытания с несколькими исходами

Пусть имеется n независимых испытаний с исходами E_1, \dots, E_r ; для каждого испытания $P(E_i) = p_i$, $1 \leq i \leq r$, $p_1 + \dots + p_r = 1$. Найдем вероятность $P_n(k_1, \dots, k_r)$ того, что число исходов E_i равно k_i при $i = 1, \dots, r$. (Например, при $r = 2$, $p_1 = p$, $p_2 = q$ $P_n(k, n - k) = p_n(k)$) Подчиним k_i естественному условию $k_1 + \dots + k_r = n$. Если фиксированы номера испытаний для 1-го, ..., r -го исходов, то вероятность соответствующих исходов равна $p_1^{k_1} \dots p_r^{k_r}$. Номера испытаний для 1-го исхода можно выбрать $C_n^{k_1}$ способами. Из оставшихся $n - k_1$ номеров испытаний можно выбрать k_2 номеров для 2-го исхода $C_{n-k_1}^{k_2}$ способами. Продолжая этот процесс, в результате получаем число $C_n(k_1, \dots, k_r) = C_n^{k_1} C_{n-k_1}^{k_2} \dots$

... $C_n^{k_1 + \dots + k_{r-1}}$ способов выбора номеров всех исходов. Записав числа сочетаний через факториалы, можно заметить, что ряд сомножителей сокращается «наискось», и в результате получаем

$$P_n(k_1, \dots, k_r) = C_n(k_1, \dots, k_r) p_1^{k_1} \dots p_r^{k_r} = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}. \quad (1)$$

Набор вероятностей вида (1) называется *полиномиальным распределением*: разложив полином $(p_1 z_1 + \dots + p_r z_r)^n$ по степеням z_1, \dots, z_r , получим, что $P_n(k_1, \dots, k_r)$ равно коэффициенту при $z_1^{k_1} \dots z_r^{k_r}$. Ниже предполагается, что $p_i > 0$, $1 \leq i \leq r$.

Пример. Футбольной команде предстоит сыграть 5 матчей с другими командами. Считая, что результаты матчей независимы, причем вероятность выигрыша 0,2, ничьей 0,6, проигрыша 0,2, найти вероятность того, что данная команда будет иметь 2 выигрыша, 1 проигрыш и 2 ничьих.

Требуется вычислить $P_5(2, 2, 1)$ при $p_1 = 0,2$, $p_2 = 0,6$, $p_3 = 0,2$. По формуле (1) имеем

$$P_5(2, 2, 1) = \frac{5!}{2! 2! 1!} \cdot 0,2^2 \cdot 0,6^2 \cdot 0,2 \approx 0,0864.$$

Пусть $s_n^{(i)}$ — число исходов E_i ($1 \leq i \leq r$). Тогда $s_n^{(i)}$ — биномиальная случайная величина. При различных i эти величины зависимы. Действительно, $P(s_n^{(i)} = n) = p_i^n > 0$ и в то же время $P(s_n^{(1)} = n, s_n^{(2)} = n) = 0$.

Обозначим $k_i = \alpha_i n$. Тогда $\alpha_i \geq 0$, $\alpha_1 + \dots + \alpha_r = 1$. Аналогично (4.4), (4.5) выводятся соотношения

$$C_n(k_1, \dots, k_r) \sim (2\pi n)^{-(r-1)/2} (\alpha_1 \dots \alpha_r)^{-1/2} \times \times \exp \left\{ n \left(\alpha_1 \ln \frac{1}{\alpha_1} + \dots + \alpha_r \ln \frac{1}{\alpha_r} \right) \right\}, \quad (2)$$

$$P_n(k_1, \dots, k_r) \sim (2\pi n)^{-(r-1)/2} (\alpha_1 \dots \alpha_r)^{-1/2} e^{ng(\alpha_1, \dots, \alpha_r)}, \quad (3)$$

где

$$g(\alpha_1, \dots, \alpha_r) = \sum_{i=1}^r \alpha_i \ln \frac{p_i}{\alpha_i}. \quad (4)$$

Функция $g(\alpha_1, \dots, \alpha_r)$ на симплексе $\{\alpha_i \geq 0, \Sigma \alpha_i = 1\}$ имеет максимум, равный нулю, в точке $(\alpha_i = p_i, 1 \leq i \leq r)$ и отрицательна во всех остальных точках. Отсюда подобно (5.7) получаем оценку

$$\sum_{|\bar{\alpha} - \bar{p}| > \varepsilon} P_n(\alpha_1 n, \dots, \alpha_r n) \leq C \rho^n, \quad (5)$$

где $\bar{\alpha} = (\alpha_1, \dots, \alpha_r)$, $\bar{p} = (p_1, \dots, p_r)$, $|\bar{\alpha} - \bar{p}| = \left(\sum_{i=1}^r (\alpha_i - p_i)^2 \right)^{1/2}$

— расстояние между r -мерными точками $\bar{\alpha}$, \bar{p} ; $C = C(r, \varepsilon)$; $\rho = \rho(r, \varepsilon) < 1$.

Заметив, что в m -мерный шар можно вписать гиперкуб, оценку (4) можно вывести из (5.7) также путем мажорирования левой части (4) суммами по полосам $\{|\alpha_i - p_i| > \varepsilon/\sqrt{r}\}$, несколько увеличив \bar{p} для компенсации числа сочетаний. Итак, сумма вероятностей по α вне произвольно малой окрестности \bar{p} экспоненциально убывает при $n \rightarrow \infty$.

Исследование вероятности $P_n(\alpha_1 n, \dots, \alpha_r n)$ в окрестности точки \bar{p} приводит к локальной (т. е. относящейся к индивидуальным $\bar{\alpha}$) предельной теореме, обобщающей теорему Муавра — Лапласа.

Локальная предельная теорема. Пусть $q_i = 1 - p_i$, $k_i = n p_i + x_i \sqrt{n p_i q_i}$, $1 \leq i \leq r$. Тогда при условии $|x_i| < T$, $1 \leq i \leq r$, при $n \rightarrow \infty$ равномерно относительно k_1, \dots, k_r выполняется соотношение

$$P_n(k_1, \dots, k_r) \sim (2\pi n)^{-(r-1)/2} (p_1 \dots p_r)^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^r q_i x_i^2 \right\}. \quad (6)$$

Интегральная предельная теорема. Возьмем в m -мерном пространстве любую область V , ограниченную конечным числом гладких гиперповерхностей. Тогда при $n \rightarrow \infty$

$$\sum_{(x_1, \dots, x_r) \in V} P_n(k_1, \dots, k_r) \sim (2\pi n)^{-(r-1)/2} (p_1 \dots p_r)^{-1/2} \times \\ \times \int \dots \int_V \exp \left\{ -\frac{1}{2} \sum_{i=1}^r q_i x_i^2 \right\} dx_1 \dots dx_r. \quad (7)$$

§ 9. Применение вероятностей уклонений к теории информации

Предположим, что источник сообщений вырабатывает сообщения в виде слов $\bar{e} = (e_1, \dots, e_n)$, где e_i — независимые случайные символы алфавита $\{E_1, \dots, E_r\}$, причем $P(e_i = E_j) = p_j$, $1 \leq i \leq n$, $p_1 + \dots + p_r = 1$. Таким образом, имеем n независимых испытаний с r исходами, как и в § 8. Перед тем как посылать сообщение в канал связи, его надлежит закодировать, т. е. сопоставить (e_1, \dots, e_n) некоторое слово $\bar{f} = (f_1, \dots, f_N)$ в алфавите $\{F_1, \dots, F_m\}$, где N может зависеть от \bar{e} : $N = N(\bar{e})$.

Средней длиной кодового слова назовем величину

$$\bar{N} = \sum_{\bar{e}} N(\bar{e}) P(\bar{e}). \quad (1)$$

Энтропией сообщения на один символ назовем величину

$$H = \sum_{i=1}^r p_i \log_m (1/p_i). \quad (2)$$

Пусть $\varepsilon > 0$ — сколь угодно малое фиксированное число. Учитывая результаты § 8, можно доказать, что при достаточно большом

n возможен способ кодирования, при котором

$$\bar{N} \leq H(1 + \varepsilon). \quad (3)$$

Разобьем сообщения на «высоковероятные» и «маловероятные», отнеся к первым те, для которых частоты символов E_j отличаются от p_j не более, чем на δ . Таким образом, если k_j — число символов E_j в данном сообщении, то для $\alpha_j = k_j/n$ имеем

$$|\alpha_j - p_j| \leq \delta, \quad 1 \leq j \leq r. \quad (4)$$

Обозначим через T число возможных высоковероятных сообщений. Имеем

$$T \leq K \max_{|\alpha_j - p_j| \leq \delta} C_n(k_1, \dots, k_r), \quad (5)$$

где K — число наборов (k_1, \dots, k_r) , для которых выполняются неравенства (4). Каждый такой набор можно задать числами k_1, \dots, k_{r-1} , так как $k_1 + \dots + k_r = n$. Отсюда

$$K \leq (2n\delta)^{r-1}. \quad (6)$$

Формулу (8.2) можно записать в следующем эквивалентном виде:

$$C_n(k_1, \dots, k_r) \sim (2\pi n)^{-(r-1)/2} (\alpha_1 \dots \alpha_r)^{-1/4} m^{n(\alpha_1 \log_m (1/\alpha_1) + \dots)}. \quad (7)$$

При условии (4) показатель степени m отличается от nH не более, чем на $n\varepsilon/2$, если $\delta \leq \delta_0$. Тогда из (5) — (7) при $n \geq n_0$ получим

$$T \leq m^{n(H+\varepsilon)}, \quad (8)$$

так как взятый «для избытка» множитель $m^{-n\varepsilon/2}$ намного превышает множитель порядка $n^{(r-1)/2}$, получаемый при перемножении (6) и (7). Из (8) следует, что высоковероятные сообщения можно закодировать словами длины $n(H + \varepsilon)$: их число равно $m^{n(H+\varepsilon)}$. (Погрешность, вносимая тем, что $n(H + \varepsilon)$ не обязательно целое число, пренебрежимо мала).

Число маловероятных слов меньше, чем $r^n = m^{n \log_m r}$, и все их можно закодировать словами m -значного алфавита длины $n \log_m r$. В результате данного способа кодирования получаем

$$\begin{aligned} \bar{N} &\leq \sum_{\text{высоковер. } \bar{e}} n(H + \varepsilon) \mathbf{P}(\bar{e}) + \sum_{\text{маловер. } \bar{e}} n \log_m r \mathbf{P}(\bar{e}) = \\ &= n(H + \varepsilon) + n(\log_m r - H - \varepsilon) \sum_{\text{маловер. } \bar{e}} \mathbf{P}(\bar{e}). \end{aligned} \quad (9)$$

По оценке (8.5) последняя сумма пренебрежимо мала, следовательно, получаем оценку (3).

Обычно принимаемое требование, чтобы ни одно кодовое слово не было продолжением другого (как в случае телефонных номеров), не изменяет существа дела: код всех маловероятных сообщений может начинаться с одного и того же слова длины $n(H + \varepsilon)$.

Докажем теперь, что N не может быть меньше $H(1 - \varepsilon)$. В самом деле, вероятность каждого высоковероятного сообщения

$$m^{-n(H+\varepsilon)} \leq p_1^{k_1} \dots p_r^{k_r} = m^{n(\alpha_1 \log_m p_1 + \dots)} \leq m^{-n(H-\varepsilon)}, \quad (10)$$

если δ достаточно мало.

Кодовых слов длины, меньшей $n(H - 2\varepsilon)$, не больше

$$1 + m + m^2 + \dots + m^{n(H-2\varepsilon)} < 2m^{n(H-2\varepsilon)}.$$

Из (10) следует, что сумма вероятностей тех сообщений \bar{e} , которые закодированы словами длины, меньшей $n(H - 2\varepsilon)$, не больше $2m^{-n(H-\varepsilon)+n(H-2\varepsilon)} = 2m^{-n\varepsilon}$. Отсюда

$$\bar{N} \geq n(H - 2\varepsilon) \sum_{N(\bar{e}) \geq n(H-2\varepsilon)} P(\bar{e}) \geq n(H - 2\varepsilon)(1 - 2m^{-n\varepsilon}), \quad n \geq n'.$$

Ввиду произвольности $\varepsilon > 0$ последнее выражение в принципе не отличается от $n(H - \varepsilon)$.

§ 10. События, связанные с исходами независимых испытаний

Часто приходится рассматривать более сложные функции исходов независимых испытаний, чем число успехов или частоты различных исходов. Приведем примеры.

Пример 1. В кольцевой системе связи (рис. 4) каждый из пунктов связан линиями с двумя соседними и двумя пунктами за ними. Обрывы линий — независимые события с вероятностью p . Требуется найти вероятность Q того, что хотя бы двумя пунктами связь окажется нарушенной.

Пример 2. Проложены линии связи между любыми двумя из n абонентов. В результате сильного повреждающего воздействия линии независимым образом обрываются с вероятностью $q = 1 - p$. Найти вероятность P того, что каждый абонент сохранит связь хотя бы с одним из остальных.

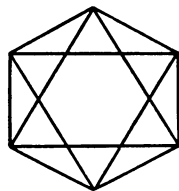


Рис. 4

Пример 3. Радиолокационное поле с двойным перекрытием (рис. 5) обеспечивает рубеж между точками O_1 и O_2 . РЛС может находиться в состоянии отказа с вероятностью q ; отказы различных РЛС — независимые события. Требуется найти вероятность Q наличия интервала, не покрытого радиолокационным полем.

Пример 4. Судовая энергетическая установка состоит из двух идентичных подсистем A и B , одна из которых изображена на рис. 6. Каждый из блоков a_1, a_2, a_3 подсистем позволяет развить мощность, равную 1, а каждый из блоков b_1, b_2 — мощность, равную 1,5. Если исправны k блоков из числа a_i и l блоков из числа b_i , то мощность подсистемы равна $\min\{k; 1,5l\}$ единиц. Мощность установки равна сумме мощностей подсистем. Найти вероятность P того, что мощность установки не менее 4, если блоки a_i исправны с вероятностью $p_1 = 1 - q_1$, а блоки b_i — с вероятностью $p_2 = 1 - q_2$, причем отказы всех 10 блоков обеих подсистем есть независимые события.

Пример 5. По линии связи передаются сигналы о текущих значениях параметров технологической установки, управляемой ЭВМ. При нормальном состоянии установки последовательность сигналов есть бернуллиева последователь-

ность символов 0, 1; $P(1) = P(0) = 1/2$. Имеется система защиты, подающая в линию связи сигнал о разладке установки в виде кодового слова

$$\overbrace{1 \ 0 \ \dots \ 0 \ 1}^{N-2}$$

Найти вероятность Q_n того, что такое же слово появится хотя бы один раз среди n переданных символов при нормальном состоянии установки (вероятность ложной тревоги).

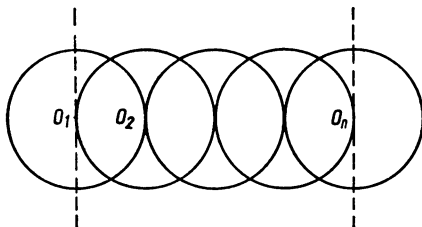


Рис. 5

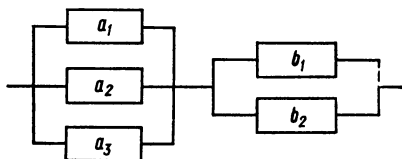


Рис. 6

Пример 6. n дробинки независимо бросаются в N ячеек; каждая дробинка может попасть в любую ячейку с вероятностью $1/N$. Требуется найти вероятность p_k того, что число ячеек, в которые не попала ни одна дробинка, равно k .

Пусть

$$\eta = \xi_1 + \dots + \xi_n, \quad (1)$$

где ξ_i — случайные величины, принимающие значения 0 и 1. Рассмотрим число

$$\beta_r = M\eta(\eta - 1) \dots (\eta - r + 1), \quad (2)$$

называемое *факториальным моментом* порядка r случайной величины η , $r = 1, 2, \dots$. Если $\eta = k$, то выражение $\xi_{i_1} \dots \xi_{i_r}$ равно 1 для C_k^r наборов $1 \leq i_1 < \dots < i_r \leq n$, для остальных же наборов оно равно 0. Отсюда

$$\sum_{1 \leq i_1 < \dots < i_r \leq n} \xi_{i_1} \dots \xi_{i_r} = C_n^r = \frac{1}{r!} \eta(\eta - 1) \dots (\eta - r + 1). \quad (3)$$

Взяв математическое ожидание левой и правой частей (3), получим

$$M \sum_{i_1 < \dots < i_r} \xi_{i_1} \dots \xi_{i_r} = \frac{1}{r!} \beta_r = \gamma_r. \quad (4)$$

Выражение $\gamma_r = MC_n^r$ называется *биномиальным моментом* случайной величины η .

Введем событие $A_i = \{\xi_i = 1\}$. Тогда ξ_i — индикатор события A_i , и $M\xi_{i_1} \dots \xi_{i_r} = P(A_{i_1} \dots A_{i_r})$. Можно записать

$$\gamma_r = \sum_{i_1 < \dots < i_r} P(A_{i_1} \dots A_{i_r}), \quad (5)$$

$$\beta_r = \sum_{i_k \neq i_l, k \neq l} P(A_{i_1} \dots A_{i_r}). \quad (6)$$

Теорема. Пусть $\eta_n = \xi_{n1} + \dots + \xi_{nn}$, где ξ_{nj} принимают значения 0, 1, и

$$\beta_{nr} = M\eta_n(\eta_n - 1) \dots (\eta_n - r + 1) \xrightarrow{n \rightarrow \infty} a^r, \quad r = 1, 2, \dots, \quad (7)$$

или, что эквивалентно,

$$\gamma_{nr} = MC_{\eta}^r \xrightarrow{n \rightarrow \infty} \frac{1}{r!} a^r, \quad r = 1, 2, \dots. \quad (8)$$

Тогда

$$P(\eta_n = k) \xrightarrow{n \rightarrow \infty} e^{-a} a^k / k!, \quad k = 0, 1, 2, \dots. \quad (9)$$

В случае, если выполняется (9), говорят, что случайная величина η_n — асимптотически пуассоновская с параметром a . Пусть η — пуассоновская случайная величина с параметром a . Тогда $\eta = k$ с вероятностью $e^{-a} a^k / k!$; с этой же вероятностью $\eta(\eta - 1) \dots (\eta - r + 1)$ равна $k(k - 1) \dots (k - r + 1)$. Отсюда

$$\begin{aligned} M\eta(\eta - 1) \dots (\eta - r + 1) &= \sum_{k=0}^{\infty} k \dots (k - r + 1) e^{-a} a^k / k! = \\ &= \sum_{k=0}^{\infty} e^{-a} a^k / (k - r)! = a^r. \end{aligned} \quad (10)$$

Итак, чтобы случайная величина была асимптотически пуассоновской, достаточно, чтобы факториальные (биномиальные) моменты случайной величины η_n сходились при $n \rightarrow \infty$ к соответствующим моментам пуассоновской случайной величины.

Доказательство теоремы приведем лишь для случая $k = 0$. Неравенства Бонферони, перейдя к противоположному событию, можно записать так:

$$\sum_{r=0}^{2m-1} (-1)^r \gamma_{nr} \leq P(\eta_n = 0) \leq \sum_{r=0}^{2m} (-1)^r \gamma_{nr}. \quad (11)$$

Вследствие (8) для любого $\varepsilon > 0$ при достаточно большом n

$$\sum_{r=0}^{2m-1} (-1)^r a^r / r! - \frac{\varepsilon}{2} \leq P(\eta_n = 0) \leq \sum_{r=0}^{2m} (-1)^r a^r / r! + \frac{\varepsilon}{2}. \quad (12)$$

Если заранее выбрать m таким образом, чтобы суммы в (12) отличались от e^{-a} меньше, чем на $\varepsilon/2$, то получим, что при достаточно большом n

$$|P(\eta_n = 0) - e^{-a}| < \varepsilon, \quad (13)$$

что и требовалось доказать.

Заметим, что от событий A_i (случайных величин ξ_i) требуется не независимость, а лишь слабая зависимость, выражаемая соотношениями (8) или (9). Наиболее удобное в приложениях условие асимптотической пуассоновости дается следующей предельной теоремой Севастьянова.

Теорема. Пусть при любом $r = 2, 3, \dots$ множество наборов $(i_1 < \dots < i_r)$ можно разбить на подмножества U_r и V_r , так, что

$$\Sigma P(A_{i_r}) \xrightarrow{n \rightarrow \infty} a, \quad (14)$$

$$P(A_{i_1} \dots A_{i_r}) / P(A_{i_1}) \dots P(A_{i_r}) \xrightarrow{n \rightarrow \infty} 1 \quad (15)$$

равномерно по $(i_1, \dots, i_r) \in U_r$ при любом $r = 2, 3, \dots$,

$$\sum_{(i_1, \dots, i_r) \in V_r} P(A_{i_1} \dots A_{i_r}) \xrightarrow{n \rightarrow \infty} 0, \quad r = 2, 3, \dots \quad (16)$$

и

$$\sum_{(i_1, \dots, i_r) \in V_r} P(A_{i_1}) \dots P(A_{i_r}) \xrightarrow{n \rightarrow \infty} 0, \quad r = 2, 3, \dots \quad (17)$$

Тогда случайная величина η_n асимптотически пуассоновская с параметром a .

Заметим, что для независимых событий A_{i_1}, \dots, A_{i_r} условие (15) выполняется, так как левая часть отношения (15) равна 1.

Покажем применение теоремы Севастьянова к приближенному решению задач из приведенных выше примеров.

Пример 1. Пусть ξ_i — индикатор разрыва кольца связи непосредственно за пунктом i по ходу часовой стрелки (событие A_i). Имеем $P(A_i) = p^3$. Обозначим

$$a = \sum_{i=1}^n P(A_i) = np^3$$

и предположим, что a — ограниченная величина. При $r \geq 2$ множество U_r определим, как множество наборов r вершин, из которых никакие две не являются соседними на кольце. Тогда для любого набора $(i_1, \dots, i_r) \in U_r$ события A_{i_1}, \dots, A_{i_r} независимы, а следовательно, (15) выполняется. Оценку (16) покажем на примере $r = 3$. Для $(i, j, k) \in V_3$ либо $j = i + 1, k > j + 1$ (суммирование по модулю n), либо $j > i + 1, k = j + 1$, либо $j = i + 1, k = j + 1$. В первом случае $P(A_i A_j A_k) = P(A_i A_j) P(A_k) = p^8$, в то время как (i, j, k) можно выбрать меньше, чем $n^2/2$ способами, откуда $\Sigma A_{ijk} < n^2 p^8 / 2 \sim n^{-2/3} \rightarrow 0$. Второй случай аналогичен первому. В третьем случае $P(A_i A_j A_k) = p^7$ и $(A_i A_j A_k)$ можно выбрать n способами (при $n \geq 3$). Отсюда $\Sigma P(A_i A_j A_k) = a^{2/3} n^{-1/3} \rightarrow 0$. В общем случае ($r \geq 2$) при $(i_1, \dots, i_r) \in V_r$ событие $A_{i_1 \dots i_r}$ имеет следствием событие $A_{j_0} A_{j_0+1} A_{j_1} \dots A_{j_t}$, причем индексы j_1, \dots, j_t не примыкают ни друг к другу, ни к j_0 , ни к $j_0 + 1$. Вероятность такого события равна p^{5+3t} , а число способов выбора индексов j_0, \dots, j_t меньше $n^{t+1}/t!$. Отсюда $\Sigma P(A_{i_1} \dots A_{i_r}) \leq \sum_t a^{t+1} \left(\frac{a}{n}\right)^{-2/3} / t! \xrightarrow{n \rightarrow \infty} 0$, а следовательно, выполняется условие (16). Число элементов U_r не меньше $n(n-3) \times \dots \times (n-6) \dots (n-3(r-1))/r!$: когда выбраны i_1, \dots, i_{k-1} , то i_k может быть любым, кроме i_1, \dots, i_{k-1} , а также кроме соседних с ними индексов; $r!$ появляется за счет упорядочения. Отсюда находим, что число элементов V_r не больше, чем

$$(n(n-1) \dots (n-r+1) - n(n-3) \dots (n-3(r-1))) / r! = O(n^{r-1});$$

так как $P(A_{i_1}) \dots P(A_{i_r}) = p^{3r}$, то видим, что выполняется и (17). Итак, число разрывов кольца есть асимптотически пуассоновская случайная величина с параметром $a = np^3$. Одни пункты отрезаны от других, если кольцо имеет хотя бы два разрыва. Отсюда

$$Q \approx 1 - e^{-a} (1 + a), \quad a = np^3.$$

Пример 2. Пусть A_i — событие, состоящее в том, что связь i -го абонента со всеми остальными оборвана. Тогда $P(A_i) = q^{n-1}$. Если η — число абонентов, оторванных от всех остальных, то искомая вероятность $P = P(\eta = 0)$. Докажем асимптотическую пуассоновость η , полагая, что

$$\sum_{i=1}^n P(A_i) = nq^{n-1} = a,$$

или, что эквивалентно,

$$q = \left(\frac{a}{n}\right)^{1/(n-1)} = \exp\left\{\frac{1}{n-1} \ln \frac{a}{n}\right\} \xrightarrow{n \rightarrow \infty} 1.$$

Положим $U_r = \{1, \dots, n\}$, $V_r = \emptyset$. Оценим $P(A_{i_1} \dots A_{i_r})$. Между абонентами i_1, \dots, i_r существовало $r(r-1)/2$ линий связи и, кроме того, $r(n-r)$ линий соединяли этих абонентов с другими $n-r$. Событию $A_{i_1} \dots A_{i_r}$ благоприятствует обрыв всех этих линий, откуда

$$P(A_{i_1} \dots A_{i_r}) = q^{r(r-1)/2 + r(n-r)} = q^{r(n-r+1)/2}.$$

Тогда

$$P(A_{i_1} \dots A_{i_r}) / P(A_{i_1}) \dots P(A_{i_r}) = q^{-r(r+1)/2} \rightarrow 1.$$

Итак, (15) выполнено и левые части (16) и (17) в данном случае равны 0. По теореме Севастьянова имеем, что при большом n

$$P \approx \exp\{-nq^{n-1}\}.$$

Пример 3. Имеем

$$Q = P\left(\bigcup_{i=1}^{n-1} B_i B_{i+1}\right),$$

где B_i — отказ i -й РЛС. Далее, $P(B_i B_{i+1}) = q^2$, $a = (n-1)q^2$. При ограниченном a

$$Q \approx e^{-a};$$

вывод почти не отличается от рассмотренного в примере 1.

Пример 4. Возможные значения мощности каждой из подсистем A, B : 1; 1,5; 2; 3. Благоприятствующие значения мощности подсистем A и B задаются парами: (1; 3), (1,5; 3), (2; 2), (2; 3), (3; 1), (3; 1,5), (3; 2), (3; 3).

Видим, что если мощность хотя бы одной подсистемы равна 3, рассматриваемое событие наступает. Вероятность такого исхода

$$2p_1^3 p_2^2 + p_1^6 p_2^4.$$

Остался еще один благоприятный исход: (2; 2). Нужно, чтобы в каждой подсистеме блоки b_1 и b_2 были исправны, а из a_1, a_2, a_3 были исправны ровно два. Вероятность такого события для одной подсистемы по формуле включения и исключения равна

$$p_2^2(3p^2 - 3p^3 + p^3) = p_2^2(3p^2 - 2p^3).$$

В результате находим

$$P = 2p_1^3 p_2^3 + p_1^6 p_2^4 + p_2^4(3p^2 - 2p^3)^2.$$

Пример 5. Пусть A_i — случайное появление кодового слова с началом на i -м такте ($1 \leq i \leq n - N + 1$). Тогда $P(A_i) = 2^{-N}$. Считаем, что

$$a = \Sigma P(A_i) = (n - N + 1) \cdot 2^{-N}$$

— ограниченная величина. При $r \geq 2$ набор индексов (i_1, \dots, i_r) отнесем к множеству U , в том и только том случае, если i_1, \dots, i_r отличаются друг от друга

больше, чем на n . Тогда, очевидно, $P(A_{i_1} \dots A_{i_r}) = P(A_{i_1}) \dots P(A_{i_r}) = 2^{-rN}$ при $(i_1, \dots, i_r) \in U_r$. При $(i_1, \dots, i_r) \in V_r$ имеем $P(A_{i_1} \dots A_{i_r}) = 0$. Итак, (15) и (16) выполняются. Число элементов множества V_r не больше, чем $n(N-1)C_{n-2}^{r-2} < n^{r-1}N/(r-2)!$. Действительно, $n(N-1)$ способами фиксируется самая левая пара (i_j, i_{j+1}) , для которой $i_{j+1} - i \leq n$, и C_{n-2}^{r-2} способами — остальные $r-2$ индексов. Следовательно, левая часть (17) не больше, чем

$$n^{r-1}2^{-rN}N/(r-2)! = a^r N/n(r-2)!.$$

Поскольку $N \sim \log_2 n$, то последнее выражение бесконечно мало при $n \rightarrow \infty$. Итак,

$$Q_n \approx 1 - e^{-a}.$$

Пример 6. Пусть A_i — событие, состоящее в том, что в i -ю ячейку не попала ни одна дробинка. Тогда $P(A_i) = \left(1 - \frac{1}{N}\right)^n$; $\sum_{i=1}^N P(A_i) = N \left(1 - \frac{1}{N}\right)^n = a$ считаем ограниченной величиной. Имеем

$$n = \ln \frac{a}{N} / \ln \left(1 - \frac{1}{N}\right).$$

Аналогично при различных i_1, \dots, i_r имеем

$$P(A_{i_1} \dots A_{i_r}) = \left(1 + \frac{r}{N}\right)^n.$$

При больших N и заданном r находим

$$\ln \left(1 - \frac{1}{N}\right) = -\frac{1}{N} + O\left(\frac{1}{N^2}\right), \quad \ln \left(1 + \frac{r}{N}\right) = \frac{r}{N} + O\left(\frac{1}{N^2}\right),$$

откуда

$$\begin{aligned} \ln P(A_{i_1} \dots A_{i_r}) - \ln (P(A_{i_1}) \dots P(A_{i_r})) &= n \left(-\frac{r}{N} + O\left(\frac{1}{N^2}\right) + \frac{r}{N} + \right. \\ &\quad \left. + O\left(\frac{1}{N^2}\right) \right) = O(n/N^2) = O(\ln N/N), \end{aligned}$$

а следовательно, (15) выполняется равномерно относительно (i_1, \dots, i_r) . По теореме Севастьянова при больших N

$$p_k \approx e^{-a} a^k / k!, \quad k = 0, 1, 2, \dots$$

В рассмотренных примерах используется выражение «величина a ограничена». Смысл этого выражения вполне точен в математической предельной схеме (при $n \rightarrow \infty$), но не определен в применении к конкретному значению параметра в практических расчетах. В конкретных исследованиях (например, теории расчета систем связи) вырабатываются рекомендации, касающиеся использования расчетных формул в четко указанных интервалах изменения параметров. Существуют также теоретические способы оценки погрешностей приближенных формул; наиболее универсальный из них — оценка остаточного члена в формуле Тейлора (например, для $\ln(1 - 1/N)$), с помощью которой эта формула получена.

§ 11. Неограниченное случайное блуждание

Пусть имеется испытание с элементарными исходами ω и каждому ω сопоставлен вектор $f(\omega) = (f^{(1)}(\omega), \dots, f^{(m)}(\omega))$, где $f^{(i)}(\omega)$ — целые числа. Функция $\xi = (\xi^{(1)}, \dots, \xi^{(m)}) = f(\omega)$ называется *m-мерной целочисленной случайной величиной*. Пусть $x = (x^{(1)}, \dots, x^{(m)})$ — вектор из множества \mathbb{Z}^m *m*-мерных векторов с целочисленными компонентами. Обозначим

$$p_{\xi}(x) = \mathbf{P}(\xi = x) = \mathbf{P}(\{\omega : f(\omega) = x\}) \quad (1)$$

и назовем $p_{\xi}(x)$ как функцию x *распределением случайной величины* ξ . (Здесь и далее в местах, где это не может вызвать неоднозначности толкования, слова «*m*-мерная целочисленная» будут опускаться.) Случайные величины ξ_1, \dots, ξ_n назовем *независимыми*, если

$$\mathbf{P}(\xi_1 = x_1, \dots, \xi_n = x_n) = p_{\xi_1}(x_1) \dots p_{\xi_n}(x_n) \quad (2)$$

для любых x_1, \dots, x_n . Из (2) путем суммирования выводится формула

$$\begin{aligned} \mathbf{P}(\xi_1 \in A_1, \dots, \xi_n \in A_n) &= \mathbf{P}(\xi_1 \in A_1) \dots \mathbf{P}(\xi_n \in A_n) = \\ &= \sum_{x_1 \in A_1} p_{\xi_1}(x_1) \dots \sum_{x_n \in A_n} p_{\xi_n}(x_n), \end{aligned} \quad (3)$$

где A_i — произвольные множества *m*-мерного пространства. Если ω_i — исходы независимых испытаний, то случайные величины $f_i(\omega_i)$ независимы.

Бесконечная последовательность (ξ_n) называется *последовательностью независимых случайных величин*, если ξ_1, \dots, ξ_n — независимые случайные величины при любом n .

Случайные величины ξ_i называются *одинаково распределенными*, если $\mathbf{P}(\xi_i \in A)$ не зависит от i . Для этого необходимо и достаточно, чтобы распределение $p_{\xi_i}(x)$ не зависело от i .

Неограниченным случайным блужданием называется последовательность $(s_n, n \geq 0)$, где s_0 — заданный вектор (чаще всего нулевой); $s_n = \xi_1 + \dots + \xi_n, n \geq 1, (\xi_n)$ — последовательность независимых одинаково распределенных случайных величин. При $m = 1$ случайное блуждание можно изобразить графиком (рис. 7) или случайной ломаной (рис. 8). При $m > 1$ можно построить график для каждой из m компонент (s_n) .

Обозначим $p(x) = p_{\xi_1}(x), i \geq 1; u_n(x) = P_{s_n}(x) = \mathbf{P}(s_n = x), x \in \mathbb{Z}^m$. По формуле полной вероятности

$$\begin{aligned} \mathbf{P}(s_n = x) &= \sum_{y \in \mathbb{Z}^m} \mathbf{P}(s_{n-1} = y, s_n = x) = \\ &= \sum_{y \in \mathbb{Z}^m} \mathbf{P}(s_{n-1} = y, \xi_n = x - y); \end{aligned}$$

так как s_{n-1} и ξ_n , очевидно, независимы, то отсюда

$$u_n(x) = \sum_{y \in \mathbb{Z}^m} u_{n-1}(y) p(x - y), \quad (4)$$

или, что то же самое,

$$u_n(x) = \sum_{y \in \mathbb{Z}^m} u_{n-1}(x-y) p(y). \quad (5)$$

При $n = 0$ имеем

$$u_0(x) = \delta_{x, s_0}, \quad (6)$$

где δ_{ab} — символ Кронекера, равный 1 при $a = b$ и 0 при $a \neq b$. По формуле (4), с учетом (6), можно вычислить $u_n(x)$ при любом n .

Пусть $(x, t) = x^{(1)}t^{(1)} + \dots + x^{(m)}t^{(m)}$ — скалярное произведение векторов x и t . Характеристической функцией m -мерной цело-

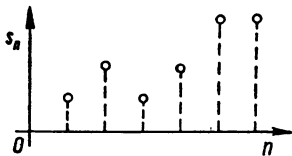


Рис. 7

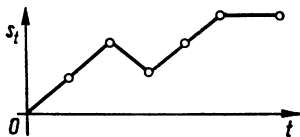


Рис. 8

численной случайной величины ξ называется функция переменного $t = (t^{(1)}, \dots, t^{(m)})$:

$$\varphi_{\xi}(t) = \sum_{x \in \mathbb{Z}^m} p_{\xi}(x) e^{i(x,t)}. \quad (7)$$

Эта функция — периодическая по $t^{(1)}, \dots, t^{(m)}$ с периодом 2π . Умножим обе части (4) на $e^{i(x,t)}$ и просуммируем по x . В результате получим

$$\begin{aligned} \sum_{x \in \mathbb{Z}^m} u_n(x) e^{i(x,t)} &= \sum_{x \in \mathbb{Z}^m} e^{i(x,t)} \sum_{y \in \mathbb{Z}^m} u_{n-1}(y) p(x-y) = \\ &= \sum_{y \in \mathbb{Z}^m} e^{i(y,t)} u_{n-1}(y) \sum_{x \in \mathbb{Z}^m} e^{i(x-y,t)} p(x-y) = |x-y=z| = \\ &= \sum_{y \in \mathbb{Z}^m} e^{i(y,t)} u_{n-1}(y) \sum_{z \in \mathbb{Z}^m} e^{i(z,t)} p(z), \end{aligned}$$

или

$$\varphi_{s_n}(t) = \varphi_{s_{n-1}}(t) \varphi_{\xi_n}(t), \quad n \geq 1. \quad (8)$$

Изменение порядка суммирования законно в силу абсолютной сходимости: $|e^{i\alpha}| = 1$, $\sum_x u_{n-1}(x) = 1$, $\sum_x p(x) = 1$. При $n = 0$

$$\varphi_{s_0}(t) = \sum_{z \in \mathbb{Z}^m} \delta_{z, s_0} e^{i(z,t)} = e^{i(s_0,t)}. \quad (9)$$

Обозначим для краткости

$$\psi(t) = \varphi_{\xi_n}(t), \quad \varphi_n(t) = \varphi_{s_n}(t), \quad n \geq 1.$$

Тогда из равенств (8) и (9) получаем

$$\varphi_n(t) = e^{i(s_0,t)} \psi^n(t), \quad n = 0, 1, 2, \dots \quad (10)$$

Из равенства (7) следует, $\varphi_{\xi}(t)$ — абсолютно сходящийся ряд Фурье, коэффициентами которого являются числа $p_{\xi}(x)$. Для этих последних выводится формула обращения

$$p_{\xi}(x) = (2\pi)^{-m} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e^{-i(x,t)} \varphi_{\xi}(t) dt_1 \dots dt_m. \quad (11)$$

В частности, из (10) следует

$$u_n(x + s_0) = (2\pi)^{-m} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e^{-i(x,t)} \psi^n(t) dt_1 \dots dt_m. \quad (12)$$

Эффективным методом анализа $u_n(x)$ является сведение выражения (11) к контурному интегралу. Пусть, например, $m = 1$ и для аналитической функции $f(z)$, $|z| < R + \varepsilon$, где $R \geq 1$, имеем $f(e^{it}) = \varphi_{\xi}(t)$ при любых действительных значениях t . Тогда заменой переменных $e^{it} = z$ формула (11) преобразуется к виду

$$p_{\xi}(x) = \frac{1}{2\pi i} \int_{|z|=R} \frac{f(z)}{z^{x+1}} dz. \quad (13)$$

Так, например, при $p(1) = p(-1) = \frac{1}{2}$ $\psi(t) = \frac{1}{2}(e^{it} + e^{-it})$, и функция $f(z)$, соответствующая $\varphi_n(t) = \psi^n(t)$, равна $2^{-n}(z + z^{-1})^n$.

Случайное блуждание удобно использовать для представления функции времени, введя шаг времени h_0 и шаги h_1, \dots, h_m по координатам $x^{(1)}, \dots, x^{(m)}$ пространства. Тогда

$$X(t) = (h_1 s^{(1)}(t/h_0), \dots, h_m s^{(m)}(t/h_0)), \quad t = 0, h_0, 2h_0, \dots \quad (14)$$

В интервалах между точками $h_0 n$ функцию $X(t)$ часто определяют по линейному закону.

Неограниченное случайное блуждание — наиболее простая вероятностная модель случайного движения безынерционной частицы — послужило основой для разработки математической теории диффузии. Величину s_n по традиции называют *положением частицы* в момент n . Разнообразие применений неограниченного случайного блуждания видно из следующих примеров.

Пример 1. Параметр технического устройства вследствие случайных толчков изменяет свое значение. Полагая, что в интервале длины h_0 параметр сохраняет свое значение с вероятностью $p(0)$ и может сместиться на $\pm h_1$ с вероятностями $p(\pm 1)$, получим неограниченное случайное блуждание в форме (14). Такая же модель используется в случае векторного параметра $(X^{(1)}(t), \dots, X^{(m)}(t))$. Представляет интерес вычисление вероятности события $\{X(t) \in A\}$, где A — область работоспособности устройства: например, A задается условием $\sum_{i=1}^m (X^{(i)}(t))^2 \leq R^2$.

Пример 2. Маневрирующая подводная лодка, находясь в данном квадрате, в течение определенного отрезка времени может перейти в один из соседних квадратов либо остаться на месте.

Пример 3. На склад в течение дня поступает случайное число комплектов оборудования. Заявки на оборудование также случайны. Обозначим через ξ_n раз-

ность между числом комплектов оборудования, поступивших на склад, и числом выданных комплектов в течение n -го дня. Тогда число s_n находящихся на складе к концу n -го дня комплектов изменяется по закону неограниченного случайного блуждания (если $s_n < 0$, то это означает наличие задолженности в поставке оборудования).

Положение блуждающей частицы s_n можно выразить через частоты исходов n независимых испытаний. В самом деле,

$$s_n - s_0 = \sum_{x \in \mathbb{Z}^m} x N_n(x), \quad (15)$$

$$s_n^{(i)} - s_0^{(i)} = \sum_{x^{(i)} \in \mathbb{Z}^1} x^{(i)} N_n^{(i)}(x^{(i)}), \quad (16)$$

где $N_n(x)$ — число испытаний с исходом $\{\xi_j = x\}$, $N_n(x^{(i)})$ — число испытаний с исходом $\{\xi_j^{(i)} = x^{(i)}\}$, $1 \leq j \leq n$.

Равенства (15), (16) можно записать в виде

$$\frac{1}{n} (s_n - s_0) = \sum_{x \in \mathbb{Z}^m} x \hat{p}_n(x), \quad (17)$$

$$\frac{1}{n} (s_n^{(i)} - s_0^{(i)}) = \sum_{x^{(i)} \in \mathbb{Z}^1} x^{(i)} \hat{p}_n^{(i)}(x^{(i)}), \quad (18)$$

где \hat{p}_n , $\hat{p}_n^{(i)}$ — частоты соответствующих исходов.

Пусть $(s_n, n \geq 0)$ — неограниченное случайное блуждание. Предположим, что $\sum_{x \in \mathbb{Z}^m} |x^{(i)}| p(x) < \infty$, $1 \leq i \leq m$, и обозначим

$$a = (a^{(1)}, \dots, a^{(m)}) = \sum_{x \in \mathbb{Z}^m} x p(x), \quad (19)$$

назвав этот вектор *математическим ожиданием* m -мерной случайной величины ξ с распределением $p(x)$, $x \in \mathbb{Z}^m$. Имеем

$$a^{(i)} = \sum_{x \in \mathbb{Z}^m} x^{(i)} p(x) = M \xi^{(i)}, \quad 1 \leq i \leq m, \quad (20)$$

так как по формуле полной вероятности сумма $p(x)$ по всем x , для которых i -я компонента равна $x^{(i)}$, составляет $P(\xi^{(i)} = x^{(i)})$.

Докажем для случайных блужданий с конечным числом возможных значений случайных величин ξ_n закон больших чисел: если $U_\varepsilon(a)$ произвольно малая окрестность вектора a , то

$$P(s_n/n \in U_\varepsilon(a)) \xrightarrow{n \rightarrow \infty} 1. \quad (21)$$

Действительно, из формулы (18) следует, что нахождение s_n вне окрестности a означает нахождение частоты $\hat{p}_n^{(i)}(x^{(i)})$ вне достаточно малой окрестности $p^{(i)}(x^{(i)})$ при некоторых i , $x^{(i)}$. Таким образом,

$$P(s_n/n \notin U_\varepsilon(a)) \leq \sum_{i, x^{(i)}} P(\hat{p}_n^{(i)}(x^{(i)}) \notin U_\delta(p^{(i)}(x^{(i)})). \quad (22)$$

Каждое слагаемое правой части (22) бесконечно мало в силу закона больших чисел для испытаний Бернулли: $\hat{p}_n^{(i)}(x^{(i)}) = B(n, p^{(i)}(x^{(i)}))/n$. Подобным же образом из (5.7) вытекает более сильное соотношение

$$P(s_n/n \notin U_\varepsilon(a)) \leq c\rho^n, \quad (23)$$

где $c > 0$, $\rho < 1$ зависят от окрестности и вероятностей исходов, а также усиленный закон больших чисел

$$P(s_n/n \xrightarrow{n \rightarrow \infty} a) = 1. \quad (24)$$

Последнее утверждение является некоторым обобщением теоремы Бореля (§ 5).

§ 12. Ограниченное случайное блуждание. Производящие функции

Многие задачи физики, техники, военного дела и других областей приводят к исследованию последовательности случайных величин, изменяющихся в определенной области и ведущих себя внутри этой области так же, как неограниченное случайное блуждание, но изменяющих характер поведения при столкновении с границей области. Так, система с защитой характеризуется в момент t параметром $s(t)$, совершающим случайное блуждание до момента попадания на границу допустимой области. В этот момент срабатывает устройство защиты, в результате чего параметр возвращается в исходное состояние. В математических моделях рассматриваются поглощающая граница (при ее достижении частица навсегда останавливается), отражающая граница (частица выбрасывается внутрь области) и другие. Ниже будет рассмотрено одномерное ($m = 1$) случайное блуждание в $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ с задерживающей границей, определяемое следующим образом.

Пусть s_0 — заданное число; $s_n \in \mathbb{Z}^+$, $n \in \mathbb{Z}^+$, — положение частицы на n -м шаге (в момент n), $(\xi_n, n \geq 1)$ — независимые случайные величины, принимающие значения из \mathbb{Z}^+ ,

$$f(k) = P(\xi_n = k), \quad f(0) + f(1) + \dots = 1.$$

Тогда для $n \geq 1$

$$s_n = \begin{cases} s_{n-1} + \xi_n - 1 & \text{при } s_{n-1} > 0, \\ \xi_n & \text{при } s_{n-1} = 0. \end{cases} \quad (1)$$

Пример. Управляющая вычислительная машина последовательно обрабатывает заявки. Время обработки одной заявки назовем циклом. Обозначим: s_n — число ожидающих заявок после окончания n -го цикла, ξ_n — число заявок, поступающих в течение n -го цикла. Тогда если $s_{n-1} > 0$, то за время n -го цикла обрабатывается одна заявка и поступает ξ_n новых заявок, в результате чего $s_n = s_{n-1} + \xi_n - 1$; если $s_{n-1} = 0$, то в течение n -го цикла заявка не обрабатывается и тогда $s_n = \xi_n$.

Соотношение (1) можно записать так:

$$\{s_n = k\} = \left\{ \bigcup_{i=1}^{k+1} \{s_{n-1} = i, \xi_n = k - i + 1\} \right\} \cup \{s_{n-1} = 0, \xi_n = k\}. \quad (2)$$

Введем обозначение $p_n(k) = \mathbf{P}(s_n = k)$, $k \in \mathbb{Z}^+$. Тогда из (2) получаем рекуррентное соотношение

$$p_n(k) = \sum_{i=1}^{k+1} p_{n-1}(i) f_{k-i+1} + p_{n-1}(0) f_k, \quad k \in \mathbb{Z}^+. \quad (3)$$

Наибольший интерес представляет стационарное распределение ($p(k)$) случайной величины s_n , т. е. такое распределение, которое не зависит от n . В предположении, что $\mathbf{P}(s_{n-1} = k) = \mathbf{P}(s_n = k) = p(k)$, $k \in \mathbb{Z}^+$, соотношение (3) превращается в уравнение

$$p(k) = \sum_{i=1}^{k+1} p(i) f_{k-i+1} + p(0) f_k, \quad k \in \mathbb{Z}^+. \quad (4)$$

Дальнейший анализ удобно вести методом производящих функций. Если ξ — случайная величина со значениями из \mathbb{Z}^+ , то *производящей функцией* ξ (ее распределения) называется функция комплексного переменного z вида

$$\varphi_\xi(z) = \sum_{k=0}^{\infty} z^k p_\xi(k), \quad (5)$$

где $p_\xi(k) = \mathbf{P}(\xi = k)$, $k \in \mathbb{Z}^+$. Поскольку $p_\xi(k) \geq 0$, $\sum_k p_\xi(k) = 1$, то ряд (5) сходится абсолютно при $|z| \leq 1$. Внутри круга сходимости производящая функция имеет производные всех порядков, причем допускается почленное дифференцирование ряда (5). Имеем

$$p_\xi(k) = \frac{1}{k!} \varphi_\xi^{(k)}(0), \quad k \geq 1; \quad p_\xi(0) = \varphi_\xi(0). \quad (6)$$

Таким образом, по производящей функции распределение случайной величины восстанавливается однозначно.

Очевидно,

$$\varphi_\xi(1) = 1. \quad (7)$$

Продифференцировав (5) почленно внутри круга сходимости, получим

$$\varphi'_\xi(z) = \sum_{k=0}^{\infty} k z^{k-1} p_\xi(k). \quad (8)$$

Если ξ имеет математическое ожидание

$$\mathbf{M}\xi = \sum_{k=0}^{\infty} k p_\xi(k), \quad (9)$$

то из (8) получаем

$$\mathbf{M}\xi = \lim_{\substack{z \rightarrow 1 \\ |z| < 1}} \varphi'_\xi(z). \quad (10)$$

В частности, если $\varphi_{\xi}(z)$ определена при $|z| < 1 + \varepsilon$ для некоторого $\varepsilon > 0$, то

$$M\xi = \varphi'_{\xi}(1), \quad (11)$$

ввиду того что в этом случае $\varphi'_{\xi}(z)$ непрерывна в точке $z = 1$.

Возвращаясь к уравнению (4), примем без доказательства предположение, что при условии

$$\tau = M\xi_n < 1 \quad (12)$$

существует распределение $(p(k), k \in \mathbb{Z}^+)$, и найдем его производящую функцию

$$\varphi(z) = \sum_{k=0}^{\infty} p(k) z^k$$

через производящую функцию

$$\psi(z) = \sum_{k=0}^{\infty} f_k z^k.$$

Для этого умножим левую и правую части (4) на z^k и просуммируем по $k \in \mathbb{Z}^+$. Сумму $p(i) f_{k-i+1} z^k$, которая при этом встретится, представим в виде

$$\sum_{i=1}^{\infty} p_{n-1,i} z^i \sum_{k=i-1}^{\infty} f_{k-i+1} z^{k-i+1} = \varphi(z) \psi(z).$$

В результате получим уравнение, которое решается относительно $\varphi(z)$:

$$\varphi(z) = p(0) \psi(z) (1 - z) / (\psi(z) - z). \quad (13)$$

По теореме Лагранжа при действительном $z < 1$

$$\psi(z) = 1 + (z - 1) \psi'(\theta), \quad z < \theta < 1,$$

откуда, вследствие (10), $1 - \psi(z) \sim (1 - z) \tau$;

$$\psi(z) - z = (1 - z) - (1 - \psi(z)) \sim (1 - z) (1 - \tau).$$

При $z \rightarrow 1$ в левой части (13) получаем 1, в правой $p(0)/(1 - \tau)$. Итак, $p(0) = 1 - \tau$ и окончательно

$$\varphi(z) = (1 - \tau) \psi(z) (1 - z) / (\psi(z) - z). \quad (14)$$

Формула (14) называется *формулой Поллачека — Хинчина* и играет важную роль в теории массового обслуживания.

§ 13. Рекуррентные события

Пусть производятся 0-е, 1-е, ... испытания, в каждом из которых может наступить некоторый исход E . Обозначим $n_0 < n_1 < n_2 < \dots$ номера испытаний, закончившихся исходом E . Данный исход называется *рекуррентным событием*, если $n_0 = 0$ и $(n_i - n_{i-1}, i \geq 1)$ есть последовательность независимых одинаково распределенных

случайных величин со значениями $1, 2, \dots$. Обозначив через I_n индикатор события $\{n\text{-е испытание имеет исход } E\}$, получим следующее эквивалентное предыдущему определение. Событие E называется *рекуррентным*, если $I_0 = 1$ и расстояния между последовательными единицами в последовательности $(I_n, n \geq 0)$ — независимые одинаково распределенные случайные величины. Обозначим $\xi_i = n_i - n_{i-1}$, $f_k = P(\xi_i = k)$, $k = 1, 2, \dots$. По условию $\sum_{k=1}^{\infty} f_k = 1$.

Пример. Через интервалы времени, равные 1 в некотором масштабе, производится проверка технического устройства на работоспособность. При обнаружении отказа устройство заменяется новым, идентичным ему по вероятностным свойствам. Моменты замены устройства естественно считать моментами наступления рекуррентного события.

Обозначим $u_n = M I_n$. Тогда по условию $u_0 = 1$. Возьмем любое $n \geq 1$. Поскольку $I_0 = 1$, то $\sum_{k=0}^{n-1} I_k \geq 1$. Возможен один из n несовместных случаев: $I_k = 1$, $I_{k+1} = \dots = I_{n-1} = 0$ ($0 \leq k \leq n-1$). Следовательно,

$$\begin{aligned} u_n = P(I_n = 1) &= \sum_{k=0}^{n-1} P(I_k = 1; I_l = 0, k < l < n; I_n = 1) = \\ &= \sum_{k=0}^{n-1} u_k f_{n-k} = \sum_{k=1}^n u_{n-k} f_k. \end{aligned} \quad (1)$$

Пусть $U(z) = \sum_{n=0}^{\infty} u_n z^n$, $\psi(z) = \sum_{n=1}^{\infty} f_n z^n$. Умножим левую и правую части (1) на z^n и просуммируем по $n \geq 1$. Слева получим $U(z) - 1$. В правой части

$$\sum_{n=1}^{\infty} z^n \sum_{k=0}^{n-1} u_k f_{n-k} = \sum_{k=0}^{\infty} u_k z^k \sum_{n=k+1}^{\infty} f_{n-k} z^{n-k} = U(z) \psi(z).$$

Итак, получена формула

$$U(z) = 1/(1 - \psi(z)). \quad (2)$$

Деление законно, так как при $|z| = \rho < 1$ $|\psi(z)| \leq \sum_{n=1}^{\infty} f_n \rho^n \leq \rho \sum_{n=1}^{\infty} f_n = \rho$, а следовательно, $1 - \psi(z)$ не обращается в нуль.

В то же время при $z \rightarrow 1$, как легко видеть, $\psi(z) \rightarrow 1$, а значит, $U(z) \rightarrow \infty$. Допустим, что $\tau = M \xi_n < \infty$. В предыдущем параграфе было выяснено, что $1 - \psi(z) \sim (1 - z) \tau$ при $z \rightarrow 1$. Если предположить, что $u_n \xrightarrow{n \rightarrow \infty} a$, т. е. $u_n = a + \varepsilon_n$, где $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$, то

$$U(z) = \sum_{n=0}^{\infty} (a + \varepsilon_n) z^n = \frac{a}{1-z} + \sum_{n=0}^{N-1} \varepsilon_n z^n + \sum_{n=N}^{\infty} \varepsilon_n z^n.$$

Выберем N так, чтобы $|\varepsilon_n| < \varepsilon$ при $n \geq N$. Тогда при действительных z , $0 < z < 1$, третье слагаемое меньше $\varepsilon/(1-z)$, второе же меньше $\varepsilon_0 + \dots + \varepsilon_{N-1}$. Умножив числитель и знаменатель (2) на $1-z$ и перейдя к пределу при $z \rightarrow 1$, справа получим $1/\tau$, слева число, отличающееся от a меньше, чем на ε . Поскольку ε произвольно мало, то $a = 1/\tau$.

Итак, если $\lim_{n \rightarrow \infty} u_n$ существует, то этот предел равен обратной величине к времени между наступлениями рекуррентного события. Однако вопрос о существовании указанного предела представляет собой более сложную задачу; ее решение рассмотрено в следующем параграфе.

§ 14. Теорема восстановления

Примем следующие условия:

$$\tau = M\xi_n < \infty \quad (1)$$

и

$$\text{НОД}(k : f_k > 0) = 1, \quad (2)$$

где НОД — наибольший общий делитель. В частности, если $f_1 > 0$, то (2) выполняется.

Пример. Пусть $f_2 > 0$, $f_5 > 0$, $f_k = 0$ для остальных k . Тогда, поскольку НОД чисел 2 и 5 равен 1, то (2) выполняется, а так как возможных значений ξ_n конечное число (2 и 5), то выполняется и (1).

Поскольку $0 \leq u_n \leq 1$, то

$$0 \leq \alpha \leq \lim_{n \rightarrow \infty} u_n \leq \overline{\lim}_{n \rightarrow \infty} u_n = \beta_n \leq 1 \quad (3)$$

и существуют последовательности (u_n) , сходящиеся к α и β . Пусть, например, $u_n \xrightarrow{r \rightarrow \infty} \beta$. Докажем, что тогда

$$u_{n-k} \xrightarrow{r \rightarrow \infty} \beta \quad (4)$$

для любого k , для которого $f_k > 0$.

Предположим противное. Тогда найдется такое ε , что

$$u_{n-k_0} f_{k_0} < \beta f_{k_0} - \varepsilon \quad (5)$$

для бесконечного множества значений r и некоторого k_0 . Возьмем

такое $N \geq k_0$, чтобы $\sum_{k=N+1}^{\infty} f_k < \frac{\varepsilon}{2}$. Тогда и $\sum_{k=N+1}^{\infty} u_{n-k} f_k < \frac{\varepsilon}{2}$, так как $u_{n-k} \leq 1$. Из (3.1) получаем

$$u_n \leq \sum_{k=1}^N u_{n-k} f_k + \frac{\varepsilon}{2}. \quad (6)$$

Положим $n = n_r$ и рассмотрим верхний предел левой и правой частей (6) при $r \rightarrow \infty$. По условию $u_{n_r} \rightarrow \beta$.

Согласно (5),

$$\overline{\lim}_{r \rightarrow \infty} u_{n_r - k_0} f_{k_0} \leq \beta f_{k_0} - \varepsilon. \quad (7)$$

Поскольку $\beta = \overline{\lim}_{n \rightarrow \infty} u_n$, то

$$\overline{\lim}_{r \rightarrow \infty} u_{n_r - k} f_k \leq \beta f_k. \quad (8)$$

Из (6) — (8) приходим к противоречию:

$$\beta \leq \beta f_{k_0} - \varepsilon + \sum_{k \neq k_0} \beta f_k + \frac{\varepsilon}{2} = \beta \sum f_k - \frac{\varepsilon}{2} = \beta - \frac{\varepsilon}{2}.$$

Соотношение (4) доказано.

Пусть теперь k_1, k_2, \dots, k_m — такие числа, для которых $f_{k_i} > 0, \dots, f_{k_m} > 0$. Тогда, используя соотношение (4) m раз, получим

$$u_{n_r} - (l_1 k_1 + \dots + l_m k_m) \xrightarrow{r \rightarrow \infty} \beta. \quad (9)$$

Имеется следующее элементарное утверждение теории чисел. Пусть НОД $(k_i) = 1$. Тогда любое достаточно большое число s можно представить в виде

$$s = \sum l_i k_i. \quad (10)$$

Пример. Пусть $k_1 = 3, k_2 = 5$. Тогда $8 = 5 + 3, 9 = 3 \cdot 3, 10 = 2 \cdot 5, 11 = 2 \cdot 3 + 5, 12 = 4 \cdot 3$ и т. д., т. е. любое $s \geq 8$ можно представить в виде $3a + 5b$, где $a, b \in \mathbb{Z}^+$.

Вспомнив условие (2) и соотношение (9), заключаем, что

$$u_{n_r - s} \xrightarrow{r \rightarrow \infty} \beta \quad (11)$$

для любого $s \geq s_0$.

Обозначим $F_k = f_{k+1} + f_{k+2} + \dots$. Тогда $f_k = F_{k-1} - F_k$; $F_0 = 1$;

$$\sum_{k=0}^{\infty} F_k = \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} f_n = \sum_{n=1}^{\infty} f_n \sum_{k=0}^{n-1} 1 = \sum_{n=0}^{\infty} n f_n = \tau. \quad (12)$$

Из (13.1) получаем

$$u_n F_0 = \sum_{k=0}^{n-1} u_k (F_{n-k-1} - F_{n-k})$$

или

$$s_n = \sum_{k=0}^n u_k F_{n-k} = \sum_{k=0}^{n-1} u_k F_{n-1-k} = s_{n-1}. \quad (13)$$

Поскольку $s_0 = u_0 F_0 = 1$, то из (13) находим, что

$$\sum_{k=0}^n u_{n-k} F_k = 1, \quad n \geq 0. \quad (14)$$

Найдем такое t , чтобы $\sum_{k=t+1}^{\infty} F_k < \varepsilon$. Тогда из (14) получаем неравенство

$$\sum_{k=0}^t u_{n-k} F_k < 1 + \varepsilon. \quad (15)$$

Взяв $n = n_r + s$, получим, что в (15) все $u_{n-k} \rightarrow \beta$ при $r \rightarrow \infty$. Отсюда

$$\beta \sum_{k=0}^t F_k < 1 + \varepsilon,$$

так что

$$\beta < (1 + \varepsilon)/(\tau - \varepsilon).$$

В силу произвольности $\varepsilon > 0$ находим $\beta \leq 1/\tau$.

Аналогично, рассматривая последовательность $u_{n_r} \rightarrow \alpha$, приходим к выводу, что $\beta \geq 1/\tau$. Из совпадения верхнего предела с нижним следует соотношение

$$\lim_{n \rightarrow \infty} u_n = 1/\tau^1. \quad (16)$$

Можно доказать, что равенство (16) имеет место и при $\tau = \infty$.

Пример. В системе может произойти сбой на n -м такте с вероятностью $1 - p$, независимо от подобных событий на других тактах. В момент каждого k -го сбоя посылается вызов контрольного устройства. Найти $\lim_{n \rightarrow \infty} u_n$, где u_n — вероятность посылки вызова на n -м такте.

Обозначим через ξ_n случайную величину, равную числу отказов между $(n-1)$ -м и n -м вызовами. Можно записать, что $\xi_n = \alpha_1 + \dots + \alpha_k$, где α_i — число тактов между двумя сбоями. Имеем

$$P(\alpha_i = m) = (1 - p) p^{m-1}, \quad m \geq 1,$$

откуда

$$\begin{aligned} M\alpha_i &= (1 - p) \sum_{m=0}^{\infty} m p^{m-1} = \\ &= (1 - p) \frac{d}{dp} \sum_{m=0}^{\infty} p^m = (1 - p) \frac{d}{dp} \frac{1}{1 - p} = \frac{1}{1 - p}. \end{aligned}$$

Таким образом,

$$\tau = M\xi_n = k/(1 - p),$$

а следовательно,

$$\lim_{n \rightarrow \infty} u_n = (1 - p)/k.$$

В качестве одного из следствий предельного соотношения (16) решим задачу о *коэффициенте готовности* из теории надежности. Технический элемент исправен в течение α_k тактов, затем неисправен (восстанавливается) в течение β_k тактов, где k — номер цикла ($k = 1, 2, \dots$); случайные величины $\alpha_1, \beta_1, \alpha_2, \beta_2, \dots$ независимы;

¹ Приведенный нами способ доказательства взят из книги В. Феллера «Введение в теорию вероятностей и ее приложения» (М., 1964. Т. 1).

$a_i = P(\alpha_k = i)$, $b_i = P(\beta_k = i)$, $i = 1, 2, \dots$; $M\alpha_k = T_0$, $M\beta_k = T_1$, причем $\text{НОД}(i : P(\alpha_k + \beta_k = i) > 0) = 1$. Обозначим через f_n вероятность того, что на такте n элемент исправен. Если существует $\lim_{n \rightarrow \infty} f_n = f$, то f называется *коэффициентом готовности элемента*.

Началом цикла будем считать первый такт после восстановления. Если положить, что нулевой цикл начинается на такте 0, то k -й будет начинаться в момент $\alpha_1 + \beta_1 + \dots + \alpha_k + \beta_k$. Вероятность того, что в момент n начинается некоторый цикл, обозначим u_n , по формуле полной вероятности

$$f_n = \sum_{k=0}^n u_k A_{n-k}, \quad (17)$$

где $A_m = a_{m+1} + a_{m+2} + \dots$. Например, если в момент n начался некоторый цикл, то в момент n элемент исправен с вероятностью $A_0 = a_1 + a_2 + \dots = 1$; этому случаю соответствует вероятность $u_n A_0$. Имеем

$$u_n \xrightarrow{n \rightarrow \infty} 1/M(\alpha_k + \beta_k) = 1/(T_0 + T_1). \quad (18)$$

Поскольку $\sum_{k=0}^{\infty} A_k = T_0$ (подобная формула выводилась в процессе доказательства (16)), то из (17)

$$f_n = \sum_{k=0}^n u_{n-k} A_k \xrightarrow{n \rightarrow \infty} \lim_{n \rightarrow \infty} u_n \sum_{k=0}^{\infty} A_k = T_0/(T_0 + T_1).$$

Переход к пределу обосновывается тем, что ряд сходится равномерно в силу сходимости $\sum A_k$ и ограниченности n . Таким образом,

$$f = T_0/(T_0 + T_1). \quad (19)$$

Задачи

1. Производится испытание с возможными исходами $\omega_1, \omega_2, \omega_3, \omega_4$; все исходы равновероятны. Указать все возможные пары независимых событий, определяемых исходом данного испытания. В частности, будут ли независимыми события $\{\omega_1, \omega_2\}$ и $\{\omega_1, \omega_3\}$?

2. Рассмотрим схему испытаний Бернулли с вероятностью успеха $p = 1/2$. Пусть ξ_i — индикатор успеха в i -м испытании. Каким условиям должны удовлетворять a_0, \dots, a_n , чтобы случайная величина $a_0 + a_1 \xi_1 + \dots + a_n \xi_n$ имела биномиальное распределение?

3. Как изменится решение предыдущей задачи, если $p = 1/3$?

4. Производится 10 независимых испытаний. Вероятность неудачи в первом из них равна q , в каждом из остальных $(1 - q)/9$. Вывести формулу для вероятности $Q(q)$ неудачи хотя бы в одном испытании. С помощью калькулятора составить график функции $Q(q)/(1 - e^{-1})$ при $0 \leq q \leq 1$. Интерпретировать результат в свете обобщенной теоремы Пуассона.

5. Вероятность попадания в цель при одном выстреле равна 0,3. Вычислить с помощью калькулятора вероятности 0, 1, 2, ..., 10 попаданий при 10 выстрелах, а также соответствующие вероятности для распределения Пуассона с параметром $\lambda = 3$. Каково процентное отклонение пуассоновских вероятностей от биномиальных?

6. Пусть γ_n — пуассоновская случайная величина с параметром $a = n$.
Найти

$$\lim_{n \rightarrow \infty} \sqrt{n} P(\gamma_n = n).$$

7. Воспользовавшись теоремой Муавра — Лапласа, с помощью калькулятора приближенно вычислить $p_n(k)$ при следующих числовых данных:

- а) $n = 100$, $p = 0,5$, $k = 72$; б) $n = 160$, $p = 0,48$, $k = 80$;
в) $n = 1600$, $p = 0,52$, $k = 875$; г) $n = 10000$, $p = 0,25$, $k = 2400$.

8. Воспользовавшись интегральной теоремой Лапласа, приближенно вычислить вероятность того, что в посылке из 200 радиосигналов будет подавлено не более 50, если вероятность подавления отдельного радиосигнала равна 0,2.

9. В независимых испытаниях возможны три равновероятных исхода. При каких k_1, k_2, k_3 , в сумме равных 12, значение $p_{12}(k_1, k_2, k_3)$ будет: а) максимальным; б) минимальным?

10. Пусть в 120 независимых испытаниях возможны три равновероятных исхода; s_1, s_2 и s_3 — числа испытаний с 1-м, 2-м, 3-м исходами. С помощью формулы (8.3) приближенно оценить вероятность того, что s_1, s_2 и s_3 отличаются друг от друга не более, чем на 2.

11. Найти максимальное значение H в выражении (9.2). При каких p_1, \dots, p_r оно достигается?

12. Из семи членов жюри за А голосовало четверо, за Б трое. Найти вероятность того, что при подсчете голосов в случайном порядке Б ни разу не «перевесил» А (перевес Б над А произошел бы, например, в том случае, если бы первой попалась карточка «за Б» или первая «за А», а вторая и третья — «за Б»).

13. Частица совершает случайное блуждание по целочисленной решетке плоскости, исходя из начала координат и за единицу времени независимо сдвигаясь по каждой координате на ± 1 с равными вероятностями. Записать формулу для математического ожидания A_n числа возвращений частицы в начало координат за первые n шагов. На основании теоремы Муавра — Лапласа исследовать поведение A_n при больших n .

14. Решить задачу, аналогичную предыдущей, при блуждании по целочисленной решетке трехмерного пространства.

15. Монета бросается до тех пор, пока она не выпадет пять раз подряд гербом. Найти ряд распределения числа бросаний монеты.

16. Пусть $(\xi_1, \xi_2, \dots, \xi_{10})$ — отрезок бернуллиевой последовательности. При $1 \leq k \leq 7$ обозначим I_k, J_k — индикаторы событий $\{\xi_k = \xi_{k+1} = \xi_{k+2} = \xi_{k+3} = 1\}$ и $\{\xi_k = \xi_{k+3} = 1, \xi_{k+1} = \xi_{k+2} = 0\}$. Найти распределение случайной величины $J_1 + \dots + J_7$ и доказать, что величина $I_1 + \dots + I_7$ имеет другое распределение.

17. В независимой случайной равновероятной последовательности цифр каждый член последовательности — одна из цифр 0, 1, ..., 9. Обозначим σ число выпадений слова «1990» на отрезке последовательности длины 1990. Вычислить факториальные моменты величины σ порядков 1, 2, 3, 4.

Глава 3

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 1. Случайная величина. Функция распределения.

Математическое ожидание

В гл. 1 случайная величина ξ была определена как числовая функция $f(\omega)$ элементарного события, т. е. величина, значение которой определяется исходом ω случайного испытания. Предполагается, что событие $\{\xi < x\}$, или, что то же самое, множество тех

ω , для которых $f(\omega) < x$, имеет определенную вероятность для любого x .

Пусть ξ_1, \dots, ξ_m — случайные величины, заданные на одном и том же вероятностном пространстве, т. е. $\xi_1 = f_1(\omega)$, \dots , $\xi_m = f_m(\omega)$. Векторная функция $\xi = (\xi_1, \dots, \xi_m)$ элементарного события ω называется m -мерной случайной величиной (случайным вектором).

Пример. При стрельбе из артиллерийского орудия элементарное событие ω есть совокупность параметров, определяющих точность траектории данного снаряда: погрешности в определении координат орудия, цели и наблюдательного пункта, масса снаряда, масса заряда и многие другие. Значение ω определяет две величины: ξ_1 — ошибку по дальности и ξ_2 — ошибку по направлению; их совокупность (ξ_1, ξ_2) есть двумерная случайная величина.

Многие свойства одномерных ($m = 1$) и многомерных ($m \geq 2$) случайных величин не различаются, поэтому целесообразно изучить их с единых позиций.

В гл. I была определена функция распределения $F_\xi(x)$ одномерной случайной величины ξ как $P(\xi < x)$, $x \in \mathbb{R}$.

Обозначим m -мерное пространство через \mathbb{R}^m , его точки будем обозначать $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$ и т. д. Символическое неравенство $x < y$ будем понимать как систему неравенств $x_1 < y_1, \dots, x_m < y_m$.

Функцией распределения $F_\xi(x) = F_\xi(x_1, \dots, x_m)$ m -мерной случайной величины ξ называется $P(\xi < x) = P(\xi_1 < x_1, \dots, \xi_m < x_m)$ как функция $x \in \mathbb{R}^m$.

Покажем, что $F_\xi(x)$ существует. Имеем $\{\xi < x\} = \{\xi_1 < x_1\} \times \{\xi_2 < x_2\} \dots \{\xi_m < x_m\}$, т. е. пересечение m событий. По свойству σ -алгебры пересечение событий есть также событие, т. е. оно обладает некоторой вероятностью.

Математическим ожиданием $M\xi$ одномерной случайной величины $\xi = f(\omega)$ в гл. I назван интеграл по вероятностной мере

$$M\xi = \int f(\omega) P(d\omega). \quad (1)$$

Если ξ — случайная величина с конечным множеством возможных значений x_n , то

$$M\xi = \sum_n x_n P(\xi = x_n). \quad (2)$$

Формула (2) сохраняется и в случае, когда случайная величина принимает счетное (т. е. представимое в виде последовательности) множество значений, если ряд в правой части (2) сходится абсолютно. Для произвольных измеримых функций $f(\omega)$ смысл интеграла (1) устанавливается с помощью монотонного предельного перехода (см. «Дополнение»).

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — m -мерная случайная величина. Тогда ее математическим ожиданием называется вектор

$$M\xi = (M\xi_1, \dots, M\xi_m) \quad (3)$$

в предположении, что $M\xi_1, \dots, M\xi_m$ существуют.

Большинство свойств случайных величин не зависят от природы множества Ω , порождающего эти величины, и выражаются через свойства функции распределения. Так, опыт с двумя равновозможными исходами можно представить и как бросание монеты, и как бросание игрального кубика с фиксированием четности числа выпавших очков, и как бесконечную последовательность бросаний монеты с фиксированием исхода бросания, при котором частоты герба и надписи впервые сравниваются. Поэтому вместо Ω берут пространство \mathbb{R}^m , вместо вероятностной меры на Ω — вероятностную меру на множествах A пространства \mathbb{R}^m , порождаемую функцией распределения $F(x)$. При таком подходе Ω отождествляется с \mathbb{R}^m , а элементарными событиями ω становятся сами значения случайной величины ξ . Так, в приведенном выше примере реализацию случайного вектора ошибок (ξ_1, ξ_2) можно считать элементарным событием, независимо от конкретных значений параметров, породивших эти ошибки.

Рассмотрим все σ -алгебры подмножеств \mathbb{R}^m , содержащие все подмножества вида $\{\xi < x\}$. Если некоторое подмножество \mathbb{R}^+ входит в каждую такую σ -алгебру, назовем его *борелевским множеством* \mathbb{R}^m . Совокупность борелевских множеств есть σ -алгебра \mathcal{B} борелевских множеств. Следующее утверждение имеет фундаментальное значение.

Функция распределения $F_\xi(x)$ определяет вероятностную меру $P(A) = P(\xi < A)$, где A — любое борелевское множество, т. е. $A \in \mathcal{B}$.

Насколько широк класс борелевских множеств, можно судить по перечислению некоторых множеств, являющихся борелевскими.

1. Гиперпараллелепипеды $\{a_i < \xi_i < b_i, 1 \leq i \leq m\}$, где $-\infty \leq a_i \leq b_i \leq \infty$; вместо знака $<$ внутри фигурных скобок может быть знак \leq в одном или обоих случаях.

2. Открытые множества, содержащие вместе с каждой своей точкой некоторую ее окрестность. (Пусть, например, множество A определяется условиями $f_i(\xi_1, \dots, \xi_m) > 0, 1 \leq i \leq N$, где f_i — непрерывные функции. Тогда, если $\xi = (\xi_1, \dots, \xi_m) \in A$, то $f_i(\xi_1, \dots, \xi_m) = \varepsilon_i > 0$. Вследствие непрерывности найдется такое $\delta > 0$, что при расстоянии ξ' от ξ , меньшем δ , будет $|f_i(\xi') - f_i(\xi)| < \varepsilon_i$, т. е. $f_i(\xi') > 0$. Это и означает, что окрестность точки ξ радиуса δ принадлежит A , т. е. A — открытое множество).

3. Замкнутые множества, т. е. множества, содержащие все свои предельные точки.

4. Множества A , имеющие m -мерный объем, определяемый как интеграл Римана

$$V(A) = \int_A \dots \int dx_1 \dots dx_m.$$

С помощью функции распределения можно определить вероятность попадания ξ в гиперпараллелепипед $\{a \leq \xi < b\} = \{a_i \leq$

$\leq \xi_i < b_i, 1 \leq i \leq m\}$. Для доказательства заметим, что

$$\{\xi < b\} = \{a \leq \xi < b\} \cup \left(\bigcup_{i=1}^m \{\xi_i < a_i, \xi < b\} \right),$$

причем событие $\{a \leq \xi < b\}$ несовместно с $\bigcup_{i=1}^m \{\dots\}$. Отсюда

$$P_{\xi}(b) = P(a \leq \xi < b) + P\left(\bigcup_{i=1}^m \{\xi_i < a_i, \xi < b\}\right).$$

В силу формулы включения и исключения

$$\begin{aligned} & P\left(\bigcup_{i=1}^m \{\xi_i < a_i, \xi < b\}\right) = \\ &= \sum_{r=1}^m (-1)^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq m} P(\xi_{i_1} < a_{i_1}, \dots, \xi_{i_r} < a_{i_r}; \xi < b) = \\ &= \sum_{r=1}^m (-1)^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq m} F_{\xi}(a_{i_1}, \dots, a_{i_m}; b), \end{aligned}$$

где $F_{\xi}(a_{i_1}, \dots, a_{i_m}; b)$ — символическое обозначение значения $F_{\xi}(x_1, \dots, x_m)$ при $x_{i_1} = a_{i_1}, \dots, x_{i_r} = a_{i_r}, x_j = b_j$ для всех остальных j . Следовательно,

$$P(a \leq \xi < b) = F_{\xi}(b) + \sum_{r=1}^m (-1)^r \sum_{1 \leq i_1 < \dots < i_r \leq m} F_{\xi}(a_{i_1}, \dots, a_{i_m}; b). \quad (4)$$

Так, при $m = 1$

$$P(a \leq \xi < b) = F_{\xi}(b) - F_{\xi}(a); \quad (5)$$

при $m = 2$

$$\begin{aligned} P(a_1 \leq \xi_1 < b_1, a_2 \leq \xi_2 < b_2) &= F_{\xi}(b_1, b_2) - F_{\xi}(a_1, b_2) - \\ &- F_{\xi}(b_1, a_2) + F_{\xi}(a_1, a_2). \end{aligned}$$

Пусть $f(x)$ — функция, заданная при $x \in \mathbb{R}^m$. Интеграл этой функции по вероятностной мере в \mathbb{R}^m , порожденной функцией распределения $F_{\xi}(x)$, называется *интегралом Стильеса* по F_{ξ} и обозначается

$$\int f(x) dF_{\xi}(x) = \int \dots \int f(x_1, \dots, x_m) dF_{\xi}(x_1, \dots, x_m). \quad (6)$$

Этот интеграл есть в то же время математическим ожиданием случайной величины $\eta = f(\xi)$:

$$Mf(\xi) = \int f(x) dF_{\xi}(x) \quad (7)$$

(о замене переменных в интеграле, приводящей к этой формуле, см. «Дополнение»).

Пусть $\xi = (\xi_1, \dots, \xi_m)$, $\eta = (\eta_1, \dots, \eta_n)$ — случайные величины. Тогда $\zeta = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n)$ есть $(m+n)$ -мерная случайная

величина. Ее функцию распределения обозначим $F_{\xi}(x, y)$, где $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_n)$: $F_{\xi}(x, y) = \mathbf{P}(\xi < x, \eta < y)$.

Случайные величины ξ и η называются *независимыми*, если

$$F_{\xi}(x, y) = F_{\xi}(x) F_{\eta}(y), \quad x \in \mathbb{R}^m, \quad y \in \mathbb{R}^n. \quad (8)$$

Справедливо следующее свойство. Если ξ и η независимы, то

$$\mathbf{P}(\xi \in A, \eta \in B) = \mathbf{P}(\xi \in A) \mathbf{P}(\eta \in B) \quad (9)$$

для любых борелевских множеств $A \subset \mathbb{R}^m$, $B \subset \mathbb{R}^n$. Формула (9) устанавливается стандартным приемом теории меры: сначала для гиперпараллелепипедов с помощью формулы (4), а затем распространяется на все множества из σ -алгебры борелевских множеств.

Подобным же образом вводится понятие независимости N случайных величин ξ_i любых размерностей: если $\xi = (\xi_1, \dots, \xi_N)$, то говорят, что ξ_i *независимы*, если

$$F_{\xi}(x_1, \dots, x_N) = F_{\xi_1}(x_1) \dots F_{\xi_N}(x_N). \quad (10)$$

Последовательность случайных величин (ξ_n) любых размерностей называется *последовательностью независимых случайных величин*, если ξ_1, \dots, ξ_N независимы для любого N .

Важнейшее свойство независимых случайных величин выражается равенством

$$\mathbf{M}(f_1(\xi_1) \dots f_N(\xi_N)) = \mathbf{M}f_1(\xi_1) \dots \mathbf{M}f_N(\xi_N) \quad (11)$$

для любых f_1, \dots, f_N , для которых $\mathbf{M}f_k(\xi_k)$ существуют. Равенство (11) очевидно для случайных величин, принимающих конечное множество возможных значений, для остальных же устанавливается монотонным предельным переходом (см. «Дополнение»). Аналитически оно выражается равенством

$$\begin{aligned} & \int f(x_1) \dots f(x_N) dF(x_1, \dots, x_N) = \\ & = \int f(x_1) dF_1(x_1) \dots \int f(x_N) dF_N(x_N) \end{aligned} \quad (12)$$

для совместной функции распределения $F(x_1, \dots, x_N)$ независимых случайных величин с функциями распределения $F_1(x_1), \dots, F_N(x_N)$.

Пусть ξ_1, \dots, ξ_N — независимые случайные величины; η_1, \dots, η_m — функции от них, причем множества аргументов этих функций не пересекаются. (Например, в случае $N = 14$ можно взять $\eta_1 = f_1(\xi_1, \xi_2)$, $\eta_2 = f_2(\xi_4, \xi_8, \xi_{14})$, $\eta_3 = f_3(\xi_3, \xi_9)$, $\eta_4 = f_4(\xi_5, \xi_6, \xi_{10}, \xi_{12})$.) Если η_1, \dots, η_m — случайные величины, то они независимы. Действительно, события $\{\eta_i \in A_i\}$ являются событиями, связанными со значениями аргументов этих функций.

§ 2. Свойства функции распределения

1. Поскольку значения $F_{\xi}(x)$ — вероятности событий, то

$$0 \leq F_{\xi}(x) \leq 1. \quad (1)$$

2. Выражение, записанное в правой части (1.4), неотрицательно.

Действительно, это выражение — вероятность события. Докажем свойство: $F_{\xi}(x_1, \dots, x_m)$ монотонно не убывает по переменным x_1, \dots, x_m , т. е. при $x_1 < y_1, \dots, x_m < y_m$

$$F_{\xi}(x_1, \dots, x_m) \leq F_{\xi}(y_1, \dots, y_m). \quad (2)$$

В самом деле, имеет место соотношение $\{\xi < x\} \subset \{\xi < y\}$; остается заметить, что при $A \subset B$ выполняется неравенство $P(A) \leq P(B)$.

Поскольку события $A_n = \{\xi_i < n, 1 \leq i \leq m\}$ удовлетворяют условиям $A_n \subset A_{n+1}, \bigcup_n A_n = \Omega$, то по аксиоме непрерывности $P(A_n) = F_{\xi}(n, \dots, n) \xrightarrow{n \rightarrow \infty} 1$. Если теперь переменные $x_i \rightarrow \infty, 1 \leq i \leq m$, то для любого n , начиная с некоторого момента, $x_i \geq n, 1 \leq i \leq m$, а тогда в силу (2) $F_{\xi}(x_1, \dots, x_n) \geq F_{\xi}(n, \dots, n)$. Получаем

$$\lim_{x_1, \dots, x_m \rightarrow +\infty} F_{\xi}(x_1, \dots, x_m) = 1. \quad (3)$$

3. Если некоторое $x_i \rightarrow -\infty$, остальные x_j постоянны или изменяются произвольным образом, то

$$F_{\xi}(x_1, \dots, x_n) \xrightarrow{x_i \rightarrow -\infty} 0. \quad (4)$$

Имеем

$$F_{\xi}(x_1, \dots, x_n) = P(\xi_i < x_i, 1 \leq i \leq m) \leq P(\xi_i < x_i).$$

Обозначив $B_n = \{\xi_i < -n\}$, получим $B_{n+1} \subset B_n, \bigcap_n B_n = \emptyset$, откуда $\lim_{n \rightarrow \infty} P(\xi_i < -n) = 0$. Поскольку начиная с некоторого места $x_i < -n$, то $\{\xi_i < x_i\} \subset \{\xi_i < -n\}$, а следовательно, выполняется (4).

$$4. F_{\xi}(x_1, \dots, x_m) = \lim_{z_i \uparrow x_i} F_{\xi}(z_1, \dots, z_m) = F_{\xi}(x_1 - 0, \dots, x_m - 0), \quad (5)$$

т. е. функция распределения непрерывна слева по совокупности своих аргументов (случай, когда $m = 1$, изображен на рис. 9).

Для доказательства возьмем последовательность $x^{(n)} = (x_1^{(n)}, \dots, x_m^{(n)})$, для которой $x_i^{(n)} \uparrow x_i, 1 \leq i \leq m$. Тогда $\{\xi < x^{(n)}\} \subset \{\xi < x^{(n+1)}\}$ и $\{\xi < x\} = \bigcup_n \{\xi < x^{(n)}\}$. По аксиоме непрерывности

$$P(\xi < x) = \lim_{n \rightarrow \infty} P(\xi < x^{(n)}),$$

что и выражено соотношением (5).

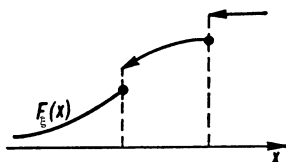
5. Если некоторые переменные x_i устремить к бесконечности, оставляя остальные неизменными, то в силу свойства (2) $F_{\xi}(x_1, \dots, x_n)$ будет сходиться к пределу, который обозначим тем же сим-

волом F_{ξ} , отметив символом ∞ переменные, устремленные к бесконечности. Например,

$$F_{\xi}(\infty, y) = \lim_{x \rightarrow \infty} F_{\xi}(x, y), \quad F_{\xi}(\infty, y, \infty, u) = \lim_{x, z \rightarrow \infty} F(x, y, z, u).$$

В таком случае $F_{\xi}(\dots, \infty, \dots, x_i, \dots)$ — функция распределения случайной величины ξ' , компоненты которой есть компоненты исходной величины ξ , соответствующие переменным, оставшимся конечными. Так, например, при $\xi = (\alpha, \beta, \gamma, \delta)$ с функцией распределения $F_{\xi}(x, y, z, u)$ имеем $F_{\xi}(\infty, y, \infty, u) = P(\beta < y, \delta < u)$. Доказательство, как и в случае предыдущих свойств, можно провести с помощью аксиомы непрерывности.

Если $\xi = (\xi_1, \dots, \xi_m)$, то $F_{\xi}(x_1, \dots, x_m)$ называют также *совместной функцией распределения* случайных величин ξ_1, \dots, ξ_m , а $F_{\xi_i}(x_i) = F_{\xi}(\infty, \dots, \infty, x_i, \infty, \dots)$,



где x_i — i -й аргумент F_{ξ} , — *частными (маргинальными) функциями распределения* случайных величин ξ_i . Частные функции распределения не определяют совместную функцию распределения однозначно.

6. Пусть $a = (a_1, \dots, a_m) \in \mathbb{R}^m$. Тогда по аксиоме непрерывности $P(\xi = a) = \lim_{\varepsilon \rightarrow 0} P(a \leq \varepsilon < a + \varepsilon)$ и по формуле (5)

$$P(\xi = a) = F_{\xi}(a + 0) + \sum_{r=1}^m (-1)^r \sum_{1 \leq i_1 < \dots < i_r \leq m} F_{\xi}(a_{i_1}, \dots, a_{i_r}; a + 0) \quad (6)$$

При $m = 1$ и $m = 2$ из (6) получаем

$$P(\xi = a) = F_{\xi}(a + 0) - F_{\xi}(a), \quad (7)$$

$$P(\xi_1 = a, \xi_2 = b) = F_{\xi}(a + 0, b + 0) - F_{\xi}(a, b + 0) - F_{\xi}(a + 0, b) + F_{\xi}(a, b).$$

Из формулы (6) вытекает следствие. Если $F_{\xi}(x)$ — непрерывная функция, то вероятность события $\{\xi = a\}$ равна 0 при любом $a \in \mathbb{R}^m$.

7. Если случайная величина ξ принимает значения $x_n = (x_{n1}, \dots, x_{nm})$ с вероятностями p_n , $\sum_n p_n = 1$, то

$$F_{\xi}(x) = \sum_{n: x_n < x} p_n. \quad (8)$$

Пример. Пусть $m = 2$ и ряд распределения случайной величины $\xi = (\alpha, \beta)$ задается таблицей

Значение ξ	(0,0)	(1,0)	(0,1)	(1,1)
Вероятность	0,6	0,2	0,08	0,12

Тогда $F_{\xi}(x)$ — ступенчатая функция, значения которой показаны на рис. 10.

Пусть в \mathbb{R}^m задана область D , имеющая m -мерный объем $V(D)$. Зададим распределение случайной величины ξ следующим условием.

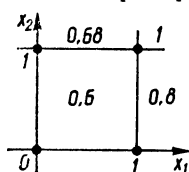


Рис. 10

Если A — любая область \mathbb{R}^m , то $P(A) = V(AD)/V(D)$, т. е. вероятность попадания ξ в область A пропорциональна объему части области D , входящей в A . Случайная величина ξ называется *равномерно распределенной* в области D . В таком случае

$$F_{\xi}(x) = V(\{\xi < x\} D)/V(D). \quad (9)$$

Пример 1. Одномерная случайная величина ξ равномерно распределена в интервале $(a; b)$. Одномерный объем — это длина; следовательно, $V(D) = b - a$. При $x < a$ множество $\{\xi < x\}$ не пересекается с D ; следовательно, $F_{\xi}(x) = 0$. При $a \leq x \leq b$ имеем $\{\xi < x\} D = (a, x)$, откуда

$$V(\{\xi < x\} D) = x - a, \quad F_{\xi}(x) = (x - a)/(b - a).$$

Если, наконец, $x > b$, то $\{\xi < x\} D = (a, b)$, откуда $F_{\xi}(x) = (b - a)/(b - a) = 1$ (рис. 11).

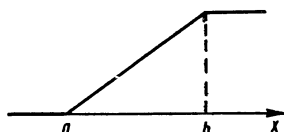


Рис. 11

Пример 2. Область D — треугольник плоскости (ξ_1, ξ_2) с вершинами $(0, \pm 1)$ и $(-1, 0)$. При различных значениях (x, y) пересечение множества $\{\xi_1 < x, \xi_2 < y\}$ с областью D — одна из фигур, показанных на рис. 12 штриховкой. Имеем $V(D) =$

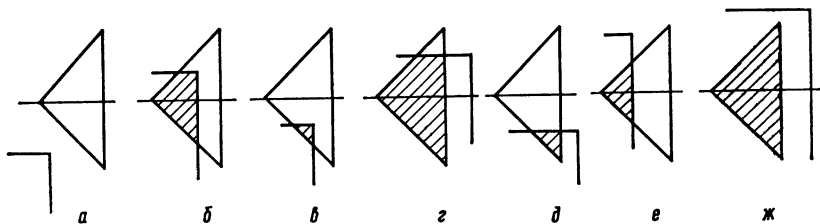


Рис. 12

$= 1$. Площади фигур вычисляются из геометрических соображений. Эти площади, а следовательно, и $F_{\xi}(x, y)$ задаются выражениями:

0	в случае а	$(1 + y)^2/2$	в случае д
$(1 + x)^2 - (1 + x - y)^2/2$	в случае б	$(1 + x)^2$	в случае е
$(1 + x + y)^2/2$	в случае в	1	в случае ж
$1 - (1 - y)^2/2$	в случае г		

Области плоскости (x, y) , в которых имеет место тот или иной случай, показаны на рис. 13.

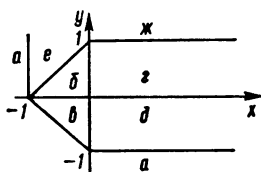


Рис. 13

Случайная величина $\xi = (\xi_1, \dots, \xi_m)$ может принимать значения на некоторой гиперповерхности S пространства \mathbb{R}^m (например, на единичной сфере трехмерного пространства). В этом случае будем называть ξ *равномерно распределенной* на S , если вероятность события $\{\xi \in S'\}$ равна

$V(S')/V(S)$, где V — объем соответствующей размерности на гиперповерхности S (например, площадь сферической области).

§ 3. Непрерывные распределения

Случайная величина ξ со значениями в \mathbb{R}^m называется *непрерывной (имеющей непрерывное распределение)*, если для любой области $D \in \mathbb{R}^m$

$$\begin{aligned} P(\xi \in D) &= \int \dots \int_D dF_{\xi}(x_1, \dots, x_m) = \\ &= \int \dots \int_D p_{\xi}(x_1, \dots, x_m) dx_1 \dots dx_m, \end{aligned} \quad (1)$$

где $p_{\xi}(x_1, \dots, x_m)$ — неотрицательная функция, называемая *плотностью вероятности (плотностью распределения)* случайной величины ξ . Сокращенное обозначение (1)

$$P(\xi \in D) = \int_D dF_{\xi}(x) = \int_D p_{\xi}(x) dx. \quad (2)$$

Пусть плотность $p_{\xi}(x)$ непрерывна в данной точке x , u_x — элементарная область, включающаяся в бесконечно малую окрестность точки x . Тогда

$$P(\xi \in u_x) = p_{\xi}(x) V(u_x) + o(V(u_x)), \quad (3)$$

где $V(u_x)$ — объем u_x .

Так, в случае $m = 1$, взяв $\Delta > 0$, получим для точек x , в которых $p_{\xi}(x)$ непрерывна,

$$P(x < \xi < x + \Delta) = p_{\xi}(x) \Delta + o(\Delta), \quad (4)$$

$$P(x - \Delta < \xi < x) = p_{\xi}(x) \Delta + o(\Delta); \quad (5)$$

при $\Delta \geq 0$, $\Delta' \geq 0$, $\Delta + \Delta' > 0$

$$P(x - \Delta' < \xi < x + \Delta) = p_{\xi}(x) (\Delta + \Delta') + o(\Delta + \Delta'). \quad (6)$$

Тот же результат имеет место, если в условиях относительно ξ знак $<$ заменить на \leq .

Возьмем в (1) область D в виде $\{t_i < x_i, 1 \leq i \leq m\}$, где t_i — переменные интегрирования, переобозначенные для отличия от x_i . Получим

$$P(\xi_i < x_i, 1 \leq i \leq m) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} p_{\xi}(t_1, \dots, t_m) dt_1 \dots dt_m, \quad (7)$$

или в сокращенном виде

$$P(\xi < x) = \int_{-\infty}^x p_{\xi}(t) dt. \quad (8)$$

Продифференцируем (7) по x_1 , затем по x_2 , ... и, наконец, по x_m . При весьма общих условиях

$$p_{\xi}(x) = \frac{\partial^m}{\partial x_1 \dots \partial x_m} F_{\xi}(x). \quad (9)$$

Для этого достаточно, чтобы кратный интеграл (7) можно было преобразовать к повторному, а именно:

$$\int_{-\infty}^{x_1} dt_1 \int_{-\infty}^{x_2} dt_2 \dots \int_{-\infty}^{x_m} p_{\xi}(t_1, \dots, t_m) dt_m,$$

и чтобы функции $p_{\xi}(x_1, \dots, x_m)$, $\int_{-\infty}^{x_m} p_{\xi}(x_1, \dots, x_{m-1}, t_m) dt_m, \dots$, \dots , $\int_{-\infty}^{x_2} dt_2 \dots \int_{-\infty}^{x_m} p_{\xi}(x_1, t_2, \dots, t_m) dt_m$ были непрерывны в окрестности точки (x_1, \dots, x_m) .

Так, при $m = 1$ из (4) и (5) имеем

$$F_{\xi}(x + \Delta) - F_{\xi}(x) = \int_x^{x+\Delta} p_{\xi}(t) dt, \quad (10)$$

$$F_{\xi}(x) - F_{\xi}(x - \Delta) = \int_{x-\Delta}^x p_{\xi}(t) dt, \quad (11)$$

а следовательно,

$$F'_{\xi}(x) = p_{\xi}(x) \quad (12)$$

для любой точки x , в которой $p_{\xi}(x)$ непрерывна.

Примером непрерывной случайной величины является случайная величина ξ , равномерно распределенная в области V пространства \mathbb{R}^m . Для нее

$$p_{\xi}(x) = \begin{cases} 1/V(D), & x \in D, \\ 0, & x \notin D. \end{cases} \quad (13)$$

Пусть $\xi = (\xi_1, \dots, \xi_m)$, $\xi' = (\xi_1, \dots, \xi_{m'})$, где $1 \leq m' \leq m$. В равенстве (8) положим $x_i = \infty$, $m' < i \leq m$. Получим

$$P(\xi_i < x_i, 1 \leq i \leq m') = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_{m'}} g(t_1, \dots, t_{m'}) dt_1 \dots dt_{m'}, \quad (14)$$

где

$$g(x_1, \dots, x_{m'}) = \int_{-\infty}^{\infty} dt_{m'+1} \dots \int_{-\infty}^{\infty} p_{\xi}(x_1, \dots, x_{m'}, t_{m'+1}, \dots, t_m) dt_m. \quad (15)$$

Формула (14) означает, что ξ' — непрерывная случайная величина с плотностью

$$p_{\xi'}(x_1, \dots, x_{m'}) = g(x_1, \dots, x_{m'}). \quad (16)$$

Итак, чтобы найти плотность вероятности случайной величины меньшей размерности, достаточно проинтегрировать плотность вероятности величины большей размерности в пределах $(-\infty; \infty)$ по

излишним переменным. Например, для получения частной плотности вероятности ξ_1 интегрируем по переменным t_2, \dots, t_m :

$$p_{\xi_1}(x) = \int_{-\infty}^{\infty} dt_2 \dots \int_{-\infty}^{\infty} p_{\xi}(x, t_2, \dots, t_m) dt_m. \quad (17)$$

Пример. Случайная величина $\xi = (\alpha, \beta)$ равномерно распределена в единичном круге с центром в начале координат. Найти плотность вероятности случайной величины α .

Имеем $p_{\xi}(x, y) = 1/\pi$ при $x^2 + y^2 < 1$. Отсюда $p_{\alpha}(x) = \int p_{\xi}(x, y) dy$ — это деленная на π длина отрезка AB (рис. 14). Находя ее по теореме Пифагора, получим

$$p_{\alpha}(x) = \frac{2}{\pi} \sqrt{1-x^2}, \quad -1 \leq x \leq 1.$$

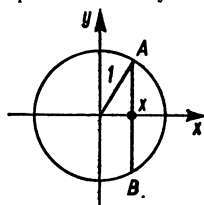


Рис. 14

При $|x| > 1$ искомая функция, очевидно, равна нулю.

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — случайная величина с плотностью $p(x) = p(x_1, \dots, x_m)$ ¹. Пусть задана функция $f(x)$, для которой $\int_{-\infty}^{\infty} f(x) p(x) dx$ сходится абсолютно. Тогда (см. «Дополнение»)

$$Mf(\xi) = \int_{-\infty}^{\infty} f(x) p(x) dx. \quad (18)$$

В частности,

$$\int_{-\infty}^{\infty} p(x) dx = M1 = 1.$$

Пример 1. Ущерб цели от разрыва осколочного снаряда на расстоянии r от нее равен $e^{-\gamma r^2}$. Найти математическое ожидание величины ущерба, если распределение точки разрыва равномерно в шаре радиуса R с центром в цели (цель предполагается точечной).

Если ξ — величина ущерба, то

$$M\xi = \left(\frac{4}{3} \pi R^3 \right)^{-1} I(R),$$

где

$$I(R) = \iiint_{x^2+y^2+z^2 < R^2} e^{-\gamma(x^2+y^2+z^2)} dx dy dz.$$

Поскольку разность объемов шаров радиусов R и $R + dR$ есть $4\pi R^2 dR$, то

$$I'(R) = 4\pi R^2 e^{-\gamma R^2}.$$

Отсюда

$$I(R) = 4\pi \int_0^R x^2 e^{-\gamma x^2} dx.$$

Данный интеграл не выражается через элементарные функции.

Пример 2. Одномерная случайная величина ξ имеет плотность вероятности $p(x) = 1 - |x|$, $-1 \leq x \leq 1$. Найти $M \ln(1/|\xi|)$.

¹ Символ ξ в обозначении плотности часто опускается.

По основной формуле (18)

$$\begin{aligned} M \ln (1/|\xi|) &= \int_{-\infty}^{\infty} \ln (1/|x|) p(x) dx = \int_{-1}^1 \ln (1/|x|) (1 - |x|) dx = \\ &= -2 \int_0^1 (\ln x) (1 - x) dx. \end{aligned}$$

Подстановкой $\ln x = -t$ этот интеграл приводится к виду

$$2 \int_0^{\infty} t (e^{-t} - e^{-2t}) dt$$

и сведением к гамма-функции либо интегрированием по частям получим

$$M \ln (1/|\xi|) = 3/2.$$

Пример 3. Случайная величина $\xi = (\xi_1, \dots, \xi_m)$ равномерно распределена в гиперкубе $\{0 < \xi_1 < 1, 1 \leq i \leq m\}$. Найти $P(\max(\xi_1, \dots, \xi_m) < z)$ и $P(\min(\xi_1, \dots, \xi_m) < z)$, где $0 \leq z \leq 1$.

Область $D = \{0 < \xi_0 < 1, 1 \leq i \leq m\}$ имеет объем, равный 1. Событию $\{\max(\xi_1, \dots, \xi_m) < z\}$ благоприятствуют те и только те (ξ_1, \dots, ξ_m) , для которых $\xi_1 < z, \dots, \xi_m < z$. Объем области, включающей такие (ξ_1, \dots, ξ_m) , равен $\int_0^z \dots \int_0^z dx \dots dx_m = z^m$. Итак,

$$P(\max(\xi_1, \dots, \xi_m) < z) = z^m.$$

Во втором случае удобнее вычислить вероятность противоположного события

$\{\min(\xi_1, \dots, \xi_m) \geq z\} = \{\xi_0 \geq z, 1 \leq i \leq m\}$. Его вероятность равна $\int_z^1 \dots \int_z^1 dx_1 \dots$
 $\dots dx_m = (1 - z)^m$. Отсюда

$$P(\min(\xi_1, \dots, \xi_m) < z) = 1 - (1 - z)^m.$$

Непрерывные случайные величины $\xi = (\xi_1, \dots, \xi_m)$ и $\eta = (\eta_1, \dots, \eta_n)$ независимы в том и только том случае, если плотность $p_{\xi}(x, y)$ случайной величины $\zeta = (\xi, \eta) = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n)$ существует и имеет вид

$$p_{\zeta}(x, y) = p_{\xi}(x) p_{\eta}(y), \quad x \in \mathbb{R}^m, y \in \mathbb{R}^n. \quad (19)$$

Достаточность этого условия доказывается интегрированием (19) по $x \in A, y \in B$, где A и B — борелевские множества. Необходимость (19) докажем для точек (x, y) , в которых существует непрерывная плотность ζ . Для элементарных объемов dV_1 и dV_2 пространств $\mathbb{R}^m, \mathbb{R}^n$ в силу независимости ξ, η имеем

$$P(\xi \in dV_1, \eta \in dV_2) = P(\xi \in dV_1) P(\eta \in dV_2).$$

Правая часть этого равенства эквивалентна $p_{\xi}(x) dV_1 p_{\eta}(y) dV_2$, а левая эквивалентна $p_{\zeta}(x, y) dV_1 dV_2$. Отсюда следует равенство (19).

Непрерывные случайные величины ξ_1, \dots, ξ_N любой размерности независимы в том и только том случае, если случайная величина

$\xi = (\xi_1, \dots, \xi_N)$ непрерывна и выполняется равенство

$$p_{\xi}(x_1, \dots, x_N) = p_{\xi_1}(x_1) \dots p_{\xi_N}(x_N). \quad (20)$$

Заметим, что значение плотности в отдельных точках можно изменить, что не сказывается на функции распределения. Поэтому с позиций теории меры нужно сказать, что распределению $F_{\xi}(x)$ соответствуют не одна плотность $p_{\xi}(x)$, а целое семейство вариантов, различающихся значениями на множествах лебеговой меры нуль. В таком случае, конечно, можно изменить в отдельных точках левую или правую части (20) так, чтобы в этих точках равенство нарушалось. Однако утверждение о необходимости (20) сохраняет смысл, если его понимать так: существует вариант плотности $p_{\xi}(x_1, \dots, x_N)$, а именно: $p_{\xi_1}(x_1) \dots p_{\xi_N}(x_N)$, для которого равенство (20) выполняется при любых x_1, \dots, x_N . В практических случаях плотности кусочно непрерывны, и этот вопрос никогда не возникает.

§ 4. Дисперсия. Ковариация. Моменты распределения

Пусть ξ — случайная величина со значениями из \mathbb{R}^m . Тогда случайная величина $\xi^0 = \xi - M\xi$ называется *центрированной случайной величиной*. Имеем $M\xi^0 = 0$ (нулевой вектор). Для ξ со значениями из \mathbb{R} число

$$\alpha_k = M\xi^k \quad (1)$$

назовем *начальным моментом* порядка k ($k = 1, 2, \dots$) случайной величины ξ . В частности, $\alpha_1 = M\xi$. Число

$$\mu_k = M(\xi^0)^k \quad (2)$$

назовем *центральный момент* порядка k ($k = 1, 2, \dots$) случайной величины ξ . В частности, $\mu_1 = 0$. Между начальными и центральными моментами существует соответствие. Заметив, что $\alpha_1 = M\xi$, найдем

$$\xi^2 = ((\xi - \alpha_1) + \alpha_1)^2 = (\xi - \alpha_1)^2 + 2\alpha_1(\xi - \alpha_1) + \alpha_1^2,$$

$$\xi^3 = ((\xi - \alpha_1) + \alpha_1)^3 = (\xi - \alpha_1)^3 + 3\alpha_1(\xi - \alpha_1)^2 + 3\alpha_1^2(\xi - \alpha_1) + \alpha_1^3.$$

Приравняв математические ожидания левой и правой частей этих равенств, получим

$$\alpha_2 = \mu_2 + \alpha_1^2, \quad (3)$$

$$\alpha_3 = \mu_3 + 3\alpha_1\mu_2 + \alpha_1^3. \quad (4)$$

Аналогично

$$\mu_2 = \alpha_2 - \alpha_1^2, \quad \mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3. \quad (5)$$

Подобные же соотношения существуют для α_k, μ_k при произвольном k . Если необходимо отметить, что момент относится к данной величине ξ , записывают

$$\alpha_k = \alpha_k(\xi), \quad \mu_k = \mu_k(\xi).$$

Наибольшее применение имеет центральный момент порядка 2, называемый *дисперсией* случайной величины ξ и обозначаемый $D\xi$. Таким образом,

$$D\xi = \mu_2 = M(\xi^0)^2 = M(\xi - M\xi)^2 = M\xi^2 - (M\xi)^2. \quad (6)$$

По определению полагаем $D\xi = \infty$ в случаях, когда $M\xi = \infty$, $M\xi$ не существует, либо $M\xi^2 = \infty$. В остальных случаях $D\xi < \infty$. Корень из дисперсии $D\xi$ называется *среднеквадратическим отклонением* случайной величины ξ и обозначается $\sigma(\xi)$. Итак,

$$\sigma(\xi) = (D\xi)^{1/2}.$$

Случайная величина ξ может быть физической величиной, имеющей некоторую размерность (расстояние, сила и т. п.). Тогда среднеквадратическое отклонение ξ имеет ту же размерность.

Пусть ξ, η — одномерные случайные величины с конечными дисперсиями. Математическое ожидание случайной величины $\xi^0\eta^0 = (\xi - M\xi)(\eta - M\eta)$ называется *ковариацией* ξ и η и обозначается

$$\text{cov}(\xi, \eta) = M\xi^0\eta^0. \quad (7)$$

Раскрыв скобки в выражении $(\xi - M\xi)(\eta - M\eta)$ и перейдя к математическим ожиданиям, получим формулу

$$\text{cov}(\xi, \eta) = M\xi\eta - M\xi M\eta.$$

Из неравенства Коши — Буняковского для абстрактных интегралов, которое на вероятностном языке формулируется в виде

$$(M\alpha\beta)^2 \leq M\alpha^2 M\beta^2, \quad (8)$$

имеем

$$(\text{cov}(\xi, \eta))^2 \leq M(\xi^0)^2 M(\eta^0)^2 = D\xi D\eta,$$

или в более употребляемом виде

$$|\text{cov}(\xi, \eta)| \leq \sigma(\xi) \sigma(\eta). \quad (9)$$

Нормированная величина

$$r(\xi, \eta) = \text{cov}(\xi, \eta) / (\sigma(\xi) \sigma(\eta)) \quad (10)$$

называется *коэффициентом корреляции* случайных величин ξ, η . Коэффициент корреляции определяется только для величин ξ, η с конечными дисперсиями и в силу (9) удовлетворяет неравенству

$$|r(\xi, \eta)| \leq 1. \quad (11)$$

Дисперсия и среднеквадратическое отклонение используются в качестве меры рассеяния значений случайной величины относительно математического ожидания. Для оправдания их понятий как меры рассеяния рассмотрим два примера.

Пример 1. Случайная величина ξ равномерно распределена в интервале $(a - \Delta/2; a + \Delta/2)$, т. е. имеет в этом интервале плотность $1/\Delta$.

Имеем

$$\begin{aligned}
 M\xi &= (1/\Delta) \int_{a-\Delta/2}^{a+\Delta/2} x dx = (1/\Delta) \int_{a-\Delta/2}^{a+\Delta/2} ((x-a) + a) dx = \\
 &= (1/\Delta) \int_{-\Delta/2}^{\Delta/2} (t + a) dt = a; \\
 D\xi &= (1/\Delta) \int_{a-\Delta/2}^{a+\Delta/2} (x-a)^2 dx = (1/\Delta) \int_{-\Delta/2}^{\Delta/2} t^2 dt = \Delta^2/12, \quad \sigma(\xi) = \Delta/2 \sqrt{3}.
 \end{aligned} \tag{12}$$

Пример 2. Случайная величина ξ имеет плотность

$$\frac{1}{2\Delta} e^{-|x-a|/\Delta}, \quad x \in \mathbb{R}.$$

Как и в предыдущем случае, получаем $M\xi = a$. Далее,

$$\begin{aligned}
 D\xi &= \frac{1}{2\Delta} \int_{-\infty}^{\infty} (x-a)^2 e^{-|x-a|/\Delta} dx = \frac{1}{\Delta} \int_0^{\infty} (x-a)^2 e^{-(x-a)/\Delta} dx = \\
 &= \int_0^{\infty} t^2 e^{-t} dt = 2\Delta^2;
 \end{aligned} \tag{13}$$

$$\sigma(\xi) = \Delta \sqrt{2}.$$

По формуле для $Mf(\xi)$

$$\alpha_k = \int x^k dF_{\xi}(x). \tag{14}$$

Если $\xi = (\xi_1, \dots, \xi_m)$, где ξ_i — одномерные величины, то начальный момент порядка k величины ξ_i определяется интегралом

$$\alpha_k(\xi_i) = \int x_i^k dF_{\xi}(x_1, \dots, x_m).$$

Центральный момент

$$\mu_k(\xi) = \int (x-a)^k dF_{\xi}(x),$$

$$\mu_k(\xi_i) = \int (x_i-a)^k dF_{\xi}(x_1, \dots, x_m)$$

и, в частности,

$$D\xi = \int (x-a)^2 dF_{\xi}(x), \tag{15}$$

$$D\xi_i = \int (x_i-a)^2 dF_{\xi}(x_1, \dots, x_m). \tag{16}$$

Ковариация случайных величин ξ, η с совместной функцией распределения $F(x, y)$ имеет вид

$$\text{cov}(\xi, \eta) = \int (x-a)(y-b) dF(x, y), \tag{17}$$

где $a = M\xi, b = M\eta$;

$$\text{cov}(\xi_i, \xi_j) = \int (x_i-a_i)(x_j-a_j) dF(x_1, \dots, x_m), \tag{18}$$

где $a_i = M\xi_i, a_j = M\xi_j$.

Очевидно,

$$\text{cov}(\xi, \xi) = D\xi.$$

Поэтому формула (18) справедлива и при $j = i$, превращаясь в этом случае в формулу (16). Для непрерывных случайных величин формулы (14) — (18) выражаются через обычные интегралы: нужно лишь заменить $dF(x_1, \dots, x_m)$ на $p_\xi(x_1, \dots, x_m) dx_1 \dots dx_m$. Например,

$$\begin{aligned} D\xi &= \int_{-\infty}^{\infty} (x - M\xi)^2 p_\xi(x) dx, \\ D\xi_i &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - M\xi_i)^2 p_\xi(x_1, \dots, x_m) dx_1 \dots dx_m, \\ \text{cov}(\xi_i, \xi_j) &= \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - M\xi_i)(x_j - M\xi_j) p_\xi(x_1, \dots, x_m) dx_1 \dots dx_m. \end{aligned}$$

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — случайная величина, для которой $D\xi_i < \infty$, $1 \leq i \leq m$. Корреляционной матрицей ξ называется матрица

$$K_\xi = \begin{pmatrix} \text{cov}(\xi_1, \xi_1) & \text{cov}(\xi_1, \xi_2) & \dots & \text{cov}(\xi_1, \xi_m) \\ \text{cov}(\xi_2, \xi_1) & \text{cov}(\xi_2, \xi_2) & \dots & \text{cov}(\xi_2, \xi_m) \\ \dots & \dots & \dots & \dots \\ \text{cov}(\xi_m, \xi_1) & \text{cov}(\xi_m, \xi_2) & \dots & \text{cov}(\xi_m, \xi_m) \end{pmatrix}.$$

Назовем случайной матрицей набор заданных на одном и том же вероятностном пространстве случайных величин ξ_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$, расположенных в виде матрицы:

$$(\xi_{ij}) = \begin{pmatrix} \xi_{11} & \xi_{12} & \dots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \dots & \xi_{2n} \\ \dots & \dots & \dots & \dots \\ \xi_{m1} & \xi_{m2} & \dots & \xi_{mn} \end{pmatrix}.$$

Можно определить сумму случайных матриц одинаковых порядков:

$$(\xi_{ij}) + (\eta_{ij}) = (\xi_{ij} + \eta_{ij});$$

произведение случайной матрицы (ξ_{ij}) порядка $m \times n$ на случайную матрицу (η_{ij}) порядка $n \times r$:

$$(\xi_{ij})(\eta_{ij}) = \left(\sum_{k=1}^n \xi_{ik} \eta_{kj} \right).$$

Это произведение — случайная матрица порядка $m \times r$.

Математическое ожидание случайной матрицы (ξ_{ij}) определяется как числовая матрица $(M\xi_{ij})$ и обозначается $M(\xi_{ij})$.

Если (ξ_{ij}) — случайная матрица, A и B — постоянные матрицы соответствующего порядка, то

$$MA(\xi_{ij}) = AM(\xi_{ij}), \quad (20)$$

$$M((\xi_{ij})B) = (M\xi_{ij})B. \quad (21)$$

Так, например, если $A = (a_{ij})$, то (i, j) -м элементом матрицы A (ξ_{ij}) будет $\sum_k a_{ik} \xi_{kj}$, а следовательно, (i, j) -й элемент матрицы M ($A \times \times (\xi_{ij})$) равен $\sum_k a_{ik} M \xi_{kj}$, т. е. (i, j) -му элементу произведения матриц A и M (ξ_{ij}).

При матричных операциях случайные и числовые векторы понимаются как векторы-столбцы. Например,

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Соответствующие векторы-строки обозначаются с помощью штриха — знака транспонирования:

$$\xi' = (\xi_1, \dots, \xi_m), \quad \eta' = (\eta_1, \dots, \eta_m), \quad x' = (x_1, \dots, x_m), \\ y' = (y_1, \dots, y_m).$$

Там, где это не может привести к недоразумениям, используется обычная запись $\xi = (\xi_1, \dots, \xi_m)$ без штриха.

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — многомерная случайная величина. Как выше было принято, символом $\xi^0 = (\xi_1^0, \dots, \xi_m^0)$ обозначается центрированная величина: $\xi^0 = \xi - M\xi = (\xi_1 - M\xi_1, \dots, \xi_m - M\xi_m)$. Имеем

$$\xi^0 \xi^{0'} = \begin{pmatrix} \xi_1^0 \\ \vdots \\ \xi_m^0 \end{pmatrix} (\xi_1^0 \dots \xi_m^0) = \begin{pmatrix} \xi_1^0 \xi_1^0 & \dots & \xi_1^0 \xi_m^0 \\ \vdots & \ddots & \vdots \\ \xi_m^0 \xi_1^0 & \dots & \xi_m^0 \xi_m^0 \end{pmatrix} = (\xi_i^0 \xi_j^0). \quad (22)$$

Приравняв математические ожидания начального и конечного выражений (22), получим

$$M \xi^0 \xi^{0'} = K_\xi = (\text{cov} (\xi_i, \xi_j)). \quad (23)$$

Полезно также понятие корреляционной матрицы $K_{\xi\eta}$ двух многомерных случайных величин ξ и η , а именно:

$$K_{\xi\eta} = (\text{cov} (\xi_i, \eta_j)).$$

По аналогии с ковариацией одномерных величин эту матрицу также обозначают $\text{cov} (\xi, \eta)$.

§ 5. Свойства дисперсии. Неравенство Чебышева

1. При умножении случайной величины на постоянную дисперсия умножается на квадрат этой постоянной:

$$D(c\xi) = c^2 D\xi.$$

Действительно,

$$D(c\xi) = M(c^2 \xi^2) - (M(c\xi))^2 = c^2 M\xi^2 - c^2 (M\xi)^2 = c^2 D\xi.$$

2. Дисперсия постоянной равна нулю (более подробно: дисперсия случайной величины, с вероятностью 1 принимающей един-

ственное значение, равна 0):

$$Dc = 0.$$

Действительно,

$$Dc = M(c^2) - (Mc)^2 = c^2 - c^2 = 0.$$

3. При смещении случайной величины на постоянную дисперсия не изменяется:

$$D(\xi + c) = D\xi. \quad (3)$$

В самом деле,

$$(\xi + c)^0 = (\xi + c) - M(\xi + c) = \xi + c - M\xi - c = \xi^0,$$

и формула (3) получается вследствие того, что $D\xi = M(\xi^0)^2$.

Назовем случайные величины ξ и η *некоррелированными*, если $\text{cov}(\xi, \eta) = 0$ (имеется в виду нулевая матрица).

Если ξ, η независимы, то они некоррелированы.

Для доказательства положим $\xi = (\xi_1, \dots, \xi_m)$, $\eta = (\eta_1, \dots, \eta_n)$, где ξ_i, η_i — одномерные случайные величины. Обозначим:

$$M\xi_i = a_i, \quad M\eta_i = b_i, \quad F_{ij}(x, y) = P(\xi_i < x, \eta_i < y).$$

Имеем

$$\begin{aligned} \text{cov}(\xi_i, \eta_j) &= \iint (x - a_i)(y - b_j) dF_{ij}(x, y) = \\ &= \iint (x - a_i)(y - b_j) dF_{\xi_i}(x) dF_{\eta_j}(y) = \\ &= \int (x - a_i) dF_{\xi_i}(x) \int (y - b_j) dF_{\eta_j}(y) = M\xi_i^0 M\eta_j^0 = 0. \end{aligned}$$

Заметим, что не любые некоррелированные случайные величины независимы. Так, равномерная в интервале $(-1/2; 1/2)$ величина ξ и величина $\eta = \xi^2$ некоррелированы:

$$M\xi = \int_{-1/2}^{1/2} x dx = 0, \quad M\xi^2 = \int_{-1/2}^{1/2} x^2 dx = 1/12, \quad M\xi^3 = \int_{-1/2}^{1/2} x^3 dx = 0,$$

откуда

$$M\xi^0 \eta^0 = M\xi(\xi^2 - 1/12) = M\xi^3 - (1/12)M\xi = 0.$$

Тем не менее, ξ и η зависимы: $P(\xi > 0,2; \xi^2 > 0,04) = P(\xi > 0,2) = 0,3$, в то время как $P(\xi^2 > 0,04) = 1 - P(\xi^2 \leq 0,04) = 1 - P(-0,2 < \xi < 0,2) = 1 - 0,4 = 0,6$, так что

$$P(\xi > 0,2) P(\xi^2 > 0,04) = 0,3 \cdot 0,6 = 0,18.$$

Важно следующее свойство.

4. Дисперсия суммы некоррелированных одномерных случайных величин равна сумме их дисперсий:

$$D \sum_{i=1}^n \xi_i = \sum_{i=1}^n D\xi_i, \quad \text{cov}(\xi_i, \xi_j) = 0, \quad i \neq j. \quad (4)$$

Для доказательства заметим, что $(\sum \xi_i)^0 = \sum \xi_i^0$, откуда

$$\begin{aligned} D \sum_{i=1}^n \xi_i &= M \left(\sum_{i=1}^n \xi_i^0 \right)^2 = M \left(\sum_{i=1}^n (\xi_i^0)^2 + \sum_{i \neq j} \xi_i^0 \xi_j^0 \right) = \\ &= \sum_{i=1}^n M (\xi_i^0)^2 + \sum_{i \neq j} M \xi_i^0 \xi_j^0 = \sum_{i=1}^n D \xi_i + \sum_{i \neq j} \text{cov} (\xi_i, \xi_j) = \sum_{i=1}^n D \xi_i. \end{aligned}$$

Поскольку по доказанному независимые случайные величины ξ_i некоррелированы, то для них выполняется равенство (4).

Формула (4) имеет важное применение к теории погрешностей измерений.

Пусть требуется измерить некоторую физическую величину a . Предположим, что измерения делаются с помощью прибора и могут повторяться. В k -м измерении получаем результат x_i . Во многих случаях допустимо предположение, что x_i — некоррелированные случайные величины с математическим ожиданием a . Для увеличения точности измерения используется среднее арифметическое $\bar{x} = (x_1 + \dots + x_n)/n$.

Имеем

$$M\bar{x} = (Mx_1 + \dots + Mx_n)/n = na/n = a, \quad (5)$$

$$\begin{aligned} D\bar{x} &= (Dx_1 + \dots + Dx_n)/n^2 = n\sigma^2/n^2 = \sigma^2/n, \\ \sigma(\bar{x}) &= \sigma/\sqrt{n}. \end{aligned} \quad (6)$$

Видим, что с увеличением числа наблюдений в n раз среднеквадратическое отклонение среднего арифметического уменьшается в \sqrt{n} раз.

5. *Неравенство Чебышева.* При любом $\varepsilon > 0$

$$P(|\xi^0| \geq \varepsilon) = P(|\xi - M\xi| \geq \varepsilon) \leq D\xi/\varepsilon^2, \quad (7)$$

т. е. абсолютное отклонение одномерной случайной величины от ее математического ожидания больше или равно ε с вероятностью, не большей отношения дисперсии этой случайной величины к ε^2 .

Для доказательства неравенства Чебышева используем свойство монотонности математического ожидания: если $f(x) \leq g(x)$ для всех x , то $Mf(\xi) \leq Mg(\xi)$. Положим $f(x) = 0$ при $|x - a| < \varepsilon$, $f(x) = 1$ при $|x - a| \geq \varepsilon$; $g(x) = (x - a)^2/\varepsilon^2$, где $a = M\xi$. Тогда $f(x) \leq g(x)$ (рис. 15).

Имеем

$$Mf(\xi) = P(|\xi - a| \geq \varepsilon),$$

$$Mg(\xi) = \varepsilon^{-2} M(\xi - a)^2 = \varepsilon^{-2} D\xi.$$

Неравенство (7) доказано.

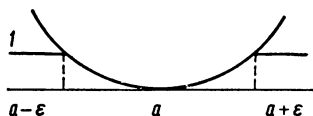


Рис. 15

Используя тот же принцип, подбором функций $f(x)$, $g(x)$ можно получить и другие полезные неравенства. Некоторые из них приведены в табл. 1.

Таблица 1

№ п/п	$f(x)$	$g(x)$	Неравенство
1	0 при $ x < \varepsilon$, 1 при $ x \geq \varepsilon$	$(1/\varepsilon) x $	$P(\xi \geq \varepsilon) \leq \frac{1}{\varepsilon} M \xi $ (неравенство Маркова)
2	0 при $ x < \varepsilon$, 1 при $ x \geq \varepsilon$	$(x/\varepsilon)^2$	$P(\xi \geq \varepsilon) \leq \frac{1}{\varepsilon^2} M\xi^2$ (неравенство Чебышева)
3	0 при $ x < \varepsilon$, 1 при $ x \geq \varepsilon$	$(x/\varepsilon)^{2n}$	$P(\xi \geq \varepsilon) \leq \frac{1}{\varepsilon^{2n}} M\xi^{2n}, \quad n = 1, 2, \dots$
4	0 при $ x - a < \varepsilon$, 1 при $ x - a \geq \varepsilon$	$(x - a)^{2n}/\varepsilon^{2n}$	$P(\xi - M\xi \geq \varepsilon) \leq \frac{1}{\varepsilon^{2n}} M(\xi - M\xi)^{2n},$ $n = 1, 2, \dots$
5	0 при $x < z$, 1 при $x \geq z$	$e^a(x-z)$	$P(\xi \geq z) \leq e^{-az} M e^{a\xi} \quad (a > 0)$
6	0 при $x > z$, 1 при $x \leq z$	$e^{-a}(x-z)$	$P(\xi \leq z) \leq e^{az} M e^{-a\xi} \quad (a > 0)$

Из неравенства Чебышева (7) следует закон больших чисел в форме Чебышева: вероятность того, что абсолютное отклонение среднего арифметического n некоррелированных случайных величин ξ_i с математическим ожиданием a и конечной дисперсией σ^2 от математического ожидания больше или равно ε , стремится к нулю при $n \rightarrow \infty$ при любом $\varepsilon > 0$:

$$P\left(\left|\frac{1}{n}(\xi_1 + \dots + \xi_n) - a\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0. \quad (8)$$

Закон больших чисел служит обоснованием принципиальной возможности вычисления математического ожидания случайной величины путем усреднения числа наблюдений.

Закон больших чисел в форме Бернулли (см. § 2.5) — частный случай закона в форме Чебышева: если ξ_i — индикатор успеха в i -м испытании, то ξ_i независимы (стало быть, и некоррелированы);

$$M\xi_i = p, \quad D\xi_i = pq.$$

Отсюда при любом $\varepsilon > 0$

$$\mathbf{P} \left(\left| \frac{1}{n} (\xi_1 + \dots + \xi_n) - p \right| \geq \varepsilon \right) \leq \frac{pq}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0,$$

что и утверждает закон больших чисел в форме Бернулли.

Из неравенства Чебышева выводятся еще два важных следствия.

1. Если $\mathbf{D}\xi = 0$, то $\xi = c$ с вероятностью 1. Действительно, обозначив $c = \mathbf{M}\xi$, имеем $\mathbf{P} \left(|\xi - c| \geq \frac{1}{n} \right) = 0$ для любого $n = 1, 2, \dots$.

Поскольку

$$\mathbf{P} (\xi \neq c) = \bigcup_{n=1}^{\infty} \mathbf{P} \left(|\xi - c| \geq \frac{1}{n} \right),$$

то по аксиоме непрерывности $\mathbf{P} (\xi \neq c) = 0$, что и требовалось доказать.

2. Если $\rho (\xi, \eta) = 1$, то $\eta = a\xi + b$, где $a > 0$; если $\rho (\xi, \eta) = -1$, то $\eta = a\xi + b$, где $a < 0$.

Имеем

$$\begin{aligned} \mathbf{D} \left(\frac{\xi}{\sigma(\xi)} \pm \frac{\eta}{\sigma(\eta)} \right) &= \mathbf{M} \left(\frac{\xi^0}{\sigma(\xi)} \pm \frac{\eta^0}{\sigma(\eta)} \right)^2 = \mathbf{M} (\xi^0)^2 / \sigma^2 (\xi) + \\ &+ \mathbf{M} (\eta^0)^2 / \sigma^2 (\eta) \pm 2\mathbf{M} (\xi^0 \eta^0) / (\sigma (\xi) \sigma (\eta)) = 2 (1 \pm r (\xi, \eta)). \end{aligned}$$

При $r (\xi, \eta) = 1$ выберем знак «—», тогда $\mathbf{D} \left(\frac{\xi}{\sigma(\xi)} - \frac{\eta}{\sigma(\eta)} \right) = 0$, откуда $\eta / \sigma (\eta) = \xi / \sigma (\xi) + c$, или $\eta = a\xi + b$, $a > 0$. При $r (\xi, \eta) = -1$ выберем знак «+», и тогда $\mathbf{D} \left(\frac{\xi}{\sigma(\xi)} + \frac{\eta}{\sigma(\eta)} \right) = 0$; $\eta / \sigma (\eta) = -\xi / \sigma (\xi) + c$; $\eta = a\xi + b$, $a < 0$.

§ 6. Линейное преобразование случайной величины и нормальное распределение

Пусть $F(x)$ и $G(x)$ — функции распределения m -мерных случайных величин. Если существуют такие случайные величины ξ и η со значениями в \mathbb{R}^m , вектор $b \in \mathbb{R}^m$ и невырожденная матрица A порядка $m \times m$, что

$$\eta = A\xi + b^1, \quad (1)$$

и при этом $F(x) = F_{\xi}(x)$, $G(x) = G_{\eta}(x)$, то говорят, что $F(x)$ и $G(x)$ принадлежат одному и тому же типу распределения.

Пусть распределение ξ непрерывно, т. е. существует плотность $p_{\xi}(x)$, $x \in \mathbb{R}^m$. При линейном преобразовании (1) элементарный объем dV пространства (ξ_1, \dots, ξ_m) переходит в элементарный объем $|\det A| dV$ пространства (η_1, \dots, η_m) ($\det A$ — обозначение определителя матрицы A). Вероятности попадания ξ (соответственно η) в оба объема совпадают. Таким образом, $p_{\xi}(x) dV = p_{\eta}(Ax +$

¹ Напомним, что при матричных операциях ξ , η , b — векторы-столбцы.

+ b) $|\det A| dV$, или после сокращения на dV

$$p_{\xi}(x) = |\det A| p_{\eta}(Ax + b). \quad (2)$$

Поскольку уравнение $Ax + b = y$ имеет решение $x = A^{-1}(y - b)$, где A^{-1} — матрица, обратная к A , то из (2) получаем обратное преобразование

$$p_{\eta}(x) = \frac{1}{|\det A|} p_{\xi}(A^{-1}(x - b)). \quad (3)$$

При $m = 1$ преобразование (1) сводится к виду

$$\eta = a\xi + b, \quad a \neq 0, \quad (4)$$

а формулы (2) и (3) принимают вид

$$p_{\xi}(x) = |a| p_{\eta}(ax + b), \quad (5)$$

$$p_{\eta}(x) = \frac{1}{|a|} p_{\xi}\left(\frac{x - b}{a}\right). \quad (6)$$

Пример. Случайная величина ω со значениями из \mathbb{R} равномерно распределена в интервале $(0; 1)$, т. е. имеет плотность $p_{\omega}(x) = 1$, $0 < x < 1$; $p_{\omega}(x) = 0$ вне интервала $(0; 1)$. Положим $\alpha = a\omega + b$, где $a > 0$. По формуле (6)

$$p_{\eta}(x) = \frac{1}{a} p_{\xi}\left(\frac{x - b}{a}\right).$$

Величина $(x - b)/a$ находится в интервале $(0; 1)$ при $b < x < b + a$. Таким образом, $p_{\eta}(x) = 1/a$ при $b < x < b + a$, $p_{\eta}(x) = 0$ вне этого интервала.

При $m = 2$ преобразование (1) принимает вид

$$\eta_1 = a_{11}\xi_1 + a_{12}\xi_2 + b_1, \quad (7)$$

$$\eta_2 = a_{21}\xi_1 + a_{22}\xi_2 + b_2$$

$$\det A = a_{11}a_{22} - a_{12}a_{21} = \Delta \neq 0, \quad (8)$$

$$A^{-1} = \frac{1}{\Delta} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix},$$

откуда

$$p_{\xi}(x, y) = |\Delta| p_{\eta}(a_{11}x + a_{12}y + b_1, a_{21}x + a_{22}y + b_2), \quad (9)$$

$$p_{\eta}(x, y) = \frac{1}{|\Delta|} p_{\xi}\left(\frac{1}{\Delta}(a_{22}(x - b_1) - a_{12}(y - b_2)), \frac{1}{\Delta}(-a_{21}(x - b_1) + a_{11}(y - b_2))\right). \quad (10)$$

Пример. Случайная величина $\xi = (\xi_1, \xi_2)$ равномерно распределена в круге $x^2 + y^2 < 1$, т. е. имеет в нем плотность $1/\pi$.

Тогда в результате применения преобразования (7) получается случайная величина η , равномерно распределенная в эллипсе

$$(a_{22}x - a_{12}y)^2 + (a_{11}y - a_{12}x)^2 \leq (a_{11}a_{22} - a_{12}a_{21})^2. \quad (11)$$

Это условие получается из (10) как условие положительности p_{ξ} .

Пусть $M\xi = b_0$, K_ξ — корреляционная матрица ξ . Тогда

$$M\eta = M(A\xi + b) = Ab_0 + b, \quad (12)$$

$$K_\eta = \text{cov}(\eta^0, (\eta^0)') = \text{cov}(A\xi^0, (\xi^0)' A')^1 = A \text{cov}(\xi^0, (\xi^0)') A',$$

или

$$K_\eta = AK_\xi A'. \quad (13)$$

Решенная задача преобразования плотности при линейном преобразовании случайной величины позволяет выделить в каждом типе распределения некоторое стандартное, а все остальные получать из него по формулам (5), (6).

Стандартное одномерное нормальное (гауссово) распределение задается плотностью

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

(14)

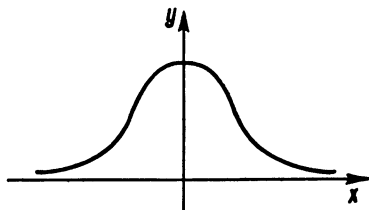


Рис. 15

Общий вид кривой $y = \varphi(x)$, называемой *кривой Гаусса*, показан на рис. 16.

Быстрое убывание $\varphi(x)$ при возрастании $|x|$ иллюстрируется табл. 2.

Таблица 2

x	$\varphi(x)$	x	$\varphi(x)$	x	$\varphi(x)$
0	0,399	3,5	$8,73 \cdot 10^{-4}$	7	$9,13 \cdot 10^{-12}$
0,5	0,352	4	$1,34 \cdot 10^{-4}$	7,5	$2,43 \cdot 10^{-13}$
1	0,242	4,5	$1,60 \cdot 10^{-5}$	8	$5,05 \cdot 10^{-15}$
1,5	0,130	5	$1,49 \cdot 10^{-6}$	8,5	$8,17 \cdot 10^{-17}$
2	$5,40 \cdot 10^{-2}$	5,5	$2,70 \cdot 10^{-7}$	9	$1,03 \cdot 10^{-18}$
2,5	$1,75 \cdot 10^{-2}$	6	$6,08 \cdot 10^{-9}$	9,5	$1,01 \cdot 10^{-20}$
3	$4,43 \cdot 10^{-3}$	6,5	$2,27 \cdot 10^{-10}$	10	$7,69 \cdot 10^{-23}$

Докажем, что $\int_{-\infty}^{\infty} \varphi(x) dx = 1$.

Пусть $I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$. Тогда

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-(x^2 + y^2)/2\} dx dy.$$

¹ При транспонировании вектора $A\xi^0$ получаем $(\xi^0)'A'$, где A' — транспонированная к A матрица.

Двойной интеграл I^2 вычисляем переходом к полярным координатам: $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, $dx dy = \rho d\rho d\varphi$. Имеем

$$I^2 = \int_0^{2\pi} d\varphi \int_0^{\infty} \rho e^{-\rho^2/2} d\rho = 2\pi \int_0^{\infty} d(-e^{-\rho^2/2}) = 2\pi;$$

$$M\xi = \int_{-\infty}^{\infty} x \varphi(x) dx = 0$$

как интеграл от нечетной функции;

$$D\xi = M\xi^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = 1$$

(интеграл сводится к вычисленному нами интегрированием по частям с заменой $u = x$, $dv = xe^{-x^2/2} dx$).

Одномерная случайная величина η называется *нормальной (гауссовской)* с параметрами $a \in \mathbb{R}$, $\sigma^2 > 0$, если она имеет плотность

$$p_{\eta}(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-a)^2/2\sigma^2}, \quad x \in \mathbb{R}. \quad (15)$$

Имеем

$$p_{\eta}(x) = \frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right), \quad (16)$$

так что по формуле (6) $p_{\eta}(x)$ совпадает с плотностью случайной величины $a + \sigma\xi$, где ξ — стандартная нормальная случайная величина. По формулам (12), (13)

$$M\eta = a, \quad D\eta = \sigma^2, \quad \sigma(\eta) = \sigma. \quad (17)$$

Итак, параметр a в выражении (15) означает математическое ожидание нормальной случайной величины, параметр σ^2 — ее дисперсию. Символическое обозначение $\eta = \mathcal{N}(a, \sigma^2)$ — нормальная случайная величина с параметрами a, σ^2 .

Функция распределения стандартной нормальной случайной величины

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz, \quad z \in \mathbb{R}, \quad (18)$$

называется *интегральной кривой Гаусса*. Эта функция не выражается через элементарные; она вычисляется приближенными приемами либо с помощью таблиц.

Функция $\Phi(x)$ уже встречалась нам в связи с интегральной предельной теоремой Лапласа (§ 2.7). Вероятность неравенства $|\eta - M\eta| < \sigma x$ равна $\Phi(x) - \Phi(-x)$. Ввиду того что чаще всего в практических расчетах требуется именно такая вероятность,

составлены таблицы функции

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz.$$

Имеем

$$\Phi(x) = \frac{1}{2} + \psi(x), \quad x \geq 0; \quad \Phi(x) = \frac{1}{2} - \psi(-x), \quad x \leq 0;$$

$$\Phi(x) - \Phi(-x) = 2\psi(x), \quad x \geq 0; \quad P(|\eta - M\eta| \geq \sigma x) = 1 - 2\psi(x), \\ x \geq 0.$$

В практических расчетах обычно пренебрегают выходом случайной величины за так называемые трехсигмовые пределы, т. е. событием $|\eta - M\eta| > 3\sigma$. По таблице находим, что вероятность такого события приблизительно равна 0,0027, т. е.

$$P(|\eta - M\eta| \leq 3\sigma) \approx 0,9973.$$

Неравенство Чебышева (§ 4) дает оценку

$$P(|\eta - M\eta| \leq 3\sigma) \geq 1 - (\sigma/3\sigma)^2 = 8/9 \approx 0,8889.$$

Если известно, что величина η нормальна, то, как видим, результат можно сильно уточнить. Можно ли использовать трехсигмовые пределы, зависит от практической ситуации.

Назовем *вероятным отклонением* значение E , при котором $P(-E \leq \eta - M\eta \leq E) = 1/2$. В ряде расчетов пренебрегают возможностью события $|\eta - M\eta| \geq 4E$, так как его вероятность меньше 0,01. Между E и σ существует приближенное соотношение $E \approx 0,674\sigma$. При приближенных расчетах (например, в теории стрельбы) применяется правило «25 — 16 — 7 — 2»: $\eta - M\eta$ попадает в интервал $(0; E)$ с вероятностью 0,25, в интервал $(E; 2E)$ с вероятностью $\approx 0,16$, в $(2E; 3E)$ с вероятностью $\approx 0,07$ и в $(3E; 4E)$ с вероятностью $\approx 0,02$.

В определенных случаях, когда требуется оценка вероятности весьма редкого события, приходится иметь дело с «хвостом» $1 - \Phi(x)$ нормального распределения при довольно больших значениях x . В этом случае полезна приводимая ниже оценка.

В интеграле

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2} dz, \quad x > 0,$$

сделаем замену переменных $z = x + u/x$, что приведет его к виду

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \lambda(x),$$

где

$$\lambda(x) = \int_0^\infty e^{-u - u^2/2x^2} du.$$

¹ Обозначение $\Phi(x)$ в таблицах используется именно для этой функции.

Подынтегральная функция лежит между $e^{-u} (1 - (u^2/2x^2))$ и e^{-u} . Отсюда

$$\int_0^{\infty} e^{-u} (1 - (u^2/2x^2)) du < \lambda(x) < \int_0^{\infty} e^{-u} du$$

или после вычисления этих интегралов

$$1 - x^{-2} < \lambda(x) < 1.$$

Отсюда получаем двустороннюю оценку «хвоста» нормального распределения

$$\left(\frac{1}{x} - \frac{1}{x^3}\right) \varphi(x) \leq 1 - \Phi(x) \leq \frac{1}{x} \varphi(x), \quad x > 0.$$

Относительная погрешность при замене $1 - \Phi(x)$ верхней и нижней оценками не больше x^{-2} . Так, при $x \geq 3,2$ относительная погрешность менее 0,1, при $x \geq 5$ — менее 0,04. Аппроксимация средней точкой интервала

$$1 - \Phi(x) \approx \left(\frac{1}{x} - \frac{1}{2x^3}\right) \varphi(x)$$

дает относительную погрешность, меньшую $1/2 x^2$.

Далее,

$$F_{\eta}(x) = P(\sigma\xi + a < x) = P\left(\xi < \frac{1}{\sigma}(x - a)\right) = \Phi((x - a)/\sigma). \quad (19)$$

Пусть ξ_i , $1 \leq i \leq m$, — независимые нормальные случайные величины с параметрами 0, σ_i^2 . Тогда плотность m -мерной случайной величины $\xi = (\xi_1, \dots, \xi_m)$

$$\begin{aligned} p_{\xi}(x) &= p_{\xi}(x_1, \dots, x_m) = \prod_{i=1}^m \left(\frac{1}{\sqrt{2\pi}\sigma_i} e^{-x_i^2/2\sigma_i^2} \right) = \\ &= (2\pi)^{-m/2} (\sigma_1 \dots \sigma_m)^{-1} \exp \left\{ -\frac{1}{2} \sum_{i=1}^m x_i^2/\sigma_i^2 \right\}. \end{aligned} \quad (20)$$

Введем диагональную матрицу

$$B = \begin{pmatrix} \sigma_1^{-2} & & 0 \\ & \sigma_2^{-2} & \\ & & \ddots \\ 0 & & & \sigma_m^{-2} \end{pmatrix}. \quad (21)$$

Легко проверить, что

$$x' B x = (x_1 \dots x_m) \begin{pmatrix} \sigma_1^{-2} & & 0 \\ & \ddots & \\ 0 & & \sigma_m^{-2} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \sum_{i=1}^m x_i^2/\sigma_i^2. \quad (22)$$

С учетом (22) можно записать

$$p_{\xi}(x) = (2\pi)^{-m/2} (\sigma_1 \dots \sigma_m)^{-1} \exp \left\{ -\frac{1}{2} x' B x \right\}. \quad (23)$$

Плотность (23) постоянна на концентрических эллипсоидах

$$\sum x_i^2 / \sigma_i^2 = c^2. \quad (24)$$

Этой плотностью (рис. 17) описывается, например, случайное отклонение (ξ_1, ξ_2) точки падения снаряда от цели по дальности и направлению при независимости ошибок по этим координатам.

Введем новую случайную величину η по формуле $\xi = C\eta$, где C — ортогональное преобразование пространства \mathbb{R}^m . Можно также записать $\eta = C'\xi$, где C' — матрица, транспонированная к C . Поскольку при ортогональном преобразовании элемент объема не изменяется, то формула (3) упрощается:

$$p_\eta(x) = p_\xi(Cx). \quad (25)$$

Известно, что C и C' — взаимно обратные матрицы. Подстановкой равенства (23) в (25) находим

$$p_\eta(x) = (2\pi)^{-m/2} (\sigma_1 \dots \sigma_m)^{-1} \times \\ \times \exp \left\{ -\frac{1}{2} x' C' B C x \right\}. \quad (26)$$

Обозначим

$$A = C' B C.$$

Поскольку C — ортогональное преобразование, то A — матрица, подобная B . Отсюда следует, что $A = (a_{ij})$ — положительно определенная матрица, т. е.

$$x' A x = \sum a_{ij} x_i x_j > 0$$

для всех $x = (x_1, \dots, x_m)$, кроме $x = (0, \dots, 0)$. Поскольку $\det C = \det C' = \pm 1$, то $\det A = \det B = (\sigma_1 \dots \sigma_m)^{-2}$. Следовательно, из (26) получаем

$$p_\eta(x) = (2\pi)^{-m/2} \sqrt{\det A} \exp \left\{ -\frac{1}{2} x' A x \right\}. \quad (27)$$

Наконец, возьмем вектор-столбец $a = (a_1, \dots, a_m)$ и образуем случайную величину $\zeta = \eta + a$. Тогда случайная величина ξ будет иметь плотность

$$p(x) = (2\pi)^{-m/2} \sqrt{\det A} \exp \left\{ -\frac{1}{2} (x - a)' A (x - a) \right\}. \quad (28)$$

Случайная величина ζ называется невырожденной m -мерной нормальной случайной величиной, если ее плотность задается формулой (28), где a — m -мерный случайный вектор, A — положительно определенная матрица. В скалярном виде (28) записывается так:

$$p_\zeta(x_1, \dots, x_m) = (2\pi)^{-m/2} \left(\begin{vmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{1m} & \dots & a_{mm} \end{vmatrix} \right)^{1/2} \times \\ \times \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^m a_{ij} (x_i - a_i) (x_j - a_j) \right\}. \quad (29)$$

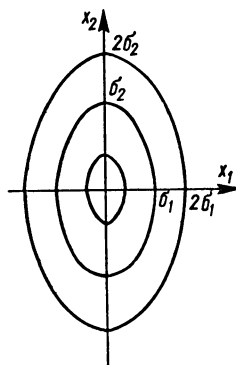


Рис. 17

Поскольку любую положительно определенную матрицу можно привести к диагональному виду, то путем линейного преобразования из нормальной случайной величины с плотностью (28) можно получить случайную величину $\xi = (\xi_1, \dots, \xi_m)$ с независимыми ξ_i и плотностью (20). Для этого полагаем

$$\xi = C(\zeta - a),$$

где C выбирается из того условия, что $CAC' = B$ — диагональная матрица. Таким образом, произвольная плотность вида (28) получается из плотности вида (20) случайной величины $\xi = (\xi_1, \dots, \xi_m)$ путем сдвига на постоянный вектор и поворота (ортogonalного преобразования) (рис. 18).

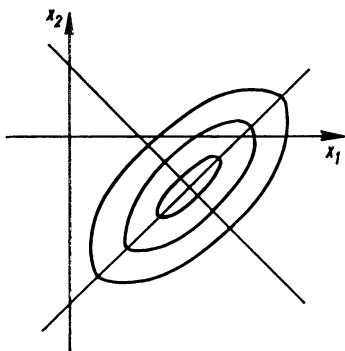


Рис. 18

Раскроем смысл матрицы A . По формуле (13)

$$K_{\eta} = C' K_{\xi} C. \quad (30)$$

Имеем $\text{cov}(\xi_i, \xi_j) = 0$ при $i \neq j$;
 $\text{cov}(\xi_i, \xi_i) = D\xi_i = \sigma_i^2$, т. е.

$$K_{\xi} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_m^2 \end{pmatrix} = B^{-1}.$$

Вспомнив, что $A = C'BC$, получаем

$$AK_{\eta} = (C'BC)(C'B^{-1}C) = C'B(CC')B^{-1}C = C'(BB^{-1})C = C'C = I,$$

т. е.

$$A = K_{\eta}^{-1}.$$

При смещении случайной величины на постоянную корреляционная матрица не изменяется, следовательно,

$$A = K_{\zeta}^{-1}. \quad (31)$$

Теперь формулу для нормальной плотности (28) можно записать следующим образом:

$$p_{\zeta}(x) = (2\pi)^{-m/2} (\det K_{\zeta})^{-1/2} \exp \left\{ -\frac{1}{2} (x - a)' K_{\zeta}^{-1} (x - a) \right\}. \quad (32)$$

Рассмотрим случай, когда $m = 2$. Обозначив $D\xi_1 = \sigma_1^2$, $D\xi_2 = \sigma_2^2$ (не следует смешивать с предыдущим обозначением $\sigma_i^2 = D\xi_i$ для независимых ξ_i), $\rho = \text{cov}(\xi_1, \xi_2)/\sigma_1\sigma_2$, получим

$$K_{\zeta} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}, \quad (33)$$

$$\det K_{\zeta} = \sigma_1^2\sigma_2^2(1 - \rho^2), \quad (34)$$

$$K_{\zeta}^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2(1 - \rho^2)} & -\frac{\rho}{\sigma_1\sigma_2(1 - \rho^2)} \\ -\frac{\rho}{\sigma_1\sigma_2(1 - \rho^2)} & \frac{1}{\sigma_2^2(1 - \rho^2)} \end{pmatrix}. \quad (35)$$

Отсюда

$$\rho_c(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-a_1}{\sigma_1}\right)^2 - 2\rho\frac{x_1-a_1}{\sigma_1}\frac{x_2-a_2}{\sigma_2} + \left(\frac{x_2-a_2}{\sigma_2}\right)^2\right)\right\}. \quad (36)$$

Из формулы (36) следует, что при $\rho = 0$ случайные величины ξ_1 и ξ_2 независимы. (Этот результат в § 10 обобщается и на многомерные нормальные случайные величины ξ_1, ξ_2 : из того, что $\text{cov}(\xi_1, \xi_2) = 0$, следует, что ξ_1 и ξ_2 независимы.)

Пусть теперь $\rho \neq 0$. Каким образом привести случайные величины ξ_1, ξ_2 к независимым ξ_1, ξ_2 ? При $m = 2$ ортогональное преобразование есть поворот осей координат на некоторый угол φ (рис. 19):

$$\begin{aligned} \xi_1 &= \eta_1 \cos \varphi + \eta_2 \sin \varphi, \\ \xi_2 &= -\eta_1 \sin \varphi + \eta_2 \cos \varphi, \end{aligned}$$

так что

$$\begin{aligned} C &= \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}, \\ C' &= \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \end{aligned}$$

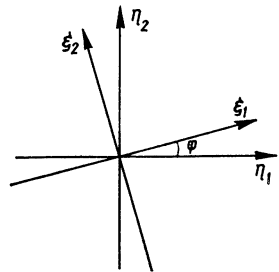


Рис. 19

Для того чтобы $CAC' = B$ была диагональной матрицей, необходимо и достаточно, чтобы $B^{-1} = CA^{-1}C' = CK_{\xi}C'$ была диагональной.

Непосредственное вычисление показывает, что оба недиагональных элемента матрицы $CK_{\xi}C'$ равны $\frac{1}{2}(\sigma_2^2 - \sigma_1^2) \sin 2\varphi + \sigma_1\sigma_2 \cos 2\varphi$. Для равенства этих элементов нулю достаточно выбрать такое φ , чтобы

$$\text{tg } 2\varphi = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}.$$

В таком случае $\xi_1 = (\xi_1 - a_1) \cos \varphi + (\xi_2 - a_2) \sin \varphi$ и $\xi_2 = -(\xi_1 - a_1) \sin \varphi + (\xi_2 - a_2) \cos \varphi$ будут независимы.

Нами было определено невырожденное нормальное распределение; кроме того, существует еще вырожденное. Пусть в пространстве \mathbb{R}^m задана гиперплоскость размерности k , $1 \leq k < m$. Любую точку этой гиперплоскости можно задать координатами t_1, \dots, t_k . Допустим, что в гиперплоскости задана невырожденная нормальная случайная величина $\alpha = (\alpha_1, \dots, \alpha_k)$. Любая точка гиперплоскости (t_1, \dots, t_k) имеет определенные координаты

$$x_j = \sum_{i=1}^k t_i d_{ij} + b_j, \quad 1 \leq j \leq m,$$

в базисе пространства \mathbb{R}^m . Таким образом, по α определяется m -мерная случайная величина $\xi = (\xi_1, \dots, \xi_m)$, где $\xi_j = \sum_{i=1}^k t_i d_{ij} + b_j$.

Эта величина и называется *вырожденной нормальной*, а ее распределение — *вырожденным нормальным распределением* ранга k . При этом говорят, что данное распределение сосредоточено на гиперплоскости, в то время как невырожденное распределение сосредоточено на всем \mathbb{R}^m .

Пусть ξ — невырожденная или вырожденная нормальная случайная величина со значениями из \mathbb{R}^m , A — любая матрица порядка $n \times m$, не переводящая пространство или гиперплоскость, где сосредоточено распределение ξ , в точку пространства \mathbb{R}^n . Тогда $A\xi + b$, где $b \in \mathbb{R}^n$, есть невырожденная или вырожденная нормальная случайная величина со значениями в \mathbb{R}^n . Ранг $A\xi + b$ не может быть больше ранга ξ .

Пример 1. Пусть η_1 и η_2 — отклонения точек падения двух выпущенных подряд снарядов от цели по дальности. Величины η_1, η_2 складываются из индивидуальных факторов ξ_1, ξ_2 и общего фактора ξ_3 . Найти плотность совместного распределения $\eta_1 = \xi_1 + \xi_3$ и $\eta_2 = \xi_2 + \xi_3$, если ξ_i — независимые нормальные случайные величины, $M\xi_i = 0$, $D\xi_1 = D\xi_2 = b^2$, $D\xi_3 = b_1^2$.

Заранее известно, что (η_1, η_2) — нормальная случайная величина. Имеем

$$\sigma_1^2 = D\eta_1 = D(\xi_1 + \xi_3) = b^2 + b_1^2,$$

$$\sigma_2^2 = D\eta_2 = D(\xi_2 + \xi_3) = b^2 + b_1^2,$$

$$\text{cov}(\eta_1, \eta_2) = M(\xi_1 + \xi_3)(\xi_2 + \xi_3) = M\xi_1\xi_2 + M\xi_1\xi_3 +$$

$$+ M\xi_2\xi_3 + M\xi_3^2 = 0 + 0 + 0 + b_1^2 = b_1^2;$$

$$\rho = \rho(\eta_1, \eta_2) = b_1^2/(b^2 + b_1^2).$$

Искомая плотность задается формулой (36), в которую подставлены данные $\sigma_1^2, \sigma_2^2, \rho$.

Пример 2. В распознающих технических системах по наблюдению m -мерного сигнала от объекта этот объект относится к тому или иному классу. Пусть, например, имеется два объекта: 1 и 2. Наблюдается m -мерная нормальная случайная величина с корреляционной матрицей K и математическими ожиданиями $a = (a_1, \dots, a_m)$ в случае объекта 1 и $b = (b_1, \dots, b_m)$ в случае объекта 2. Если получено наблюдение $x = (x_1, \dots, x_m)$, то принимают решение, что это объект 1, если $p(x - a) > p(x - b)$, где $p(x)$ — плотность нормальной величины с нулевым математическим ожиданием и корреляционной матрицей K . При противоположном неравенстве решают, что присутствует объект 2.

Из формулы (32) находим условие решения 1:

$$(x - a)' K (x - a) < (x - b)' K (x - b).$$

Допустим, что присутствует объект 1. Тогда вероятность δ ошибки распознавания (т. е. решения 2) равна вероятности события $(x - a)' K (x - a) > (x - b)' K (x - b)$, если подставить вместо x наблюдаемую случайную величину.

Имеем $x = a + \xi$, где ξ — нормальная случайная величина с математическим ожиданием 0 и корреляционной матрицей K . Следовательно,

$$\delta = P((\xi + a - a)' K (\xi + a - a) > (\xi + a - b)' K (\xi + a - b))$$

или после раскрытия скобок и сведения подобных членов

$$\delta = P((b-a)' K \xi + \xi' K (b-a) > (b-a)' K (b-a)).$$

(Таким же образом проверяется, что данное δ есть в то же время вероятностью решения 1 при существовании объекта 2.)

Видим, что δ — вероятность превышения одномерной случайной величиной ξ с нулевым математическим ожиданием постоянного числа λ . Имеем

$$\xi = (b-a)' K \xi + \xi' K (b-a) = \sum_{k=1}^m \mu_k \xi_k,$$

где μ_k — постоянные коэффициенты; $\lambda = (b-a)' K (b-a)$. Отсюда

$$D\xi = \sum_{k,j=1}^m \mu_k \mu_j \text{cov}(\xi_k, \xi_j) = \sigma^2.$$

Окончательно получаем

$$\delta = 1 - \Phi(\lambda/\sigma).$$

В случае $m = 1$, например, при $a < b$ условные ошибки имеет вид $\xi > (b-a)/2$, откуда

$$\delta = 1 - \Phi((b-a)/2).$$

В общем случае можно найти выражения $D\xi$ и λ в матричном виде. Обозначим $\Delta = b - a$, $K = (s_{ij})$. Ввиду того что, очевидно,

$$s_{it} = s_{tj}, \quad \xi = (b-a)' K \xi + \xi' K (b-a) = 2 \sum s_{ij} \Delta_i \xi_j, \\ \xi^2 = 4 \sum \Delta_i s_{ij} \xi_j \xi_{j'} s_{j'i'} \Delta_{i'}.$$

Поскольку $M\xi = 0$, $M\xi_j \xi_{j'} = s_{jj'}$, то

$$D\xi = M\xi^2 = 4 \sum \Delta_i s_{ij} s_{jj'} s_{j'i'} \Delta_{i'} = 4 \Delta' K^2 \Delta.$$

В то же время

$$\lambda = \Delta' K \Delta.$$

Итак,

$$\delta = 1 - \Phi(\Delta' K \Delta / 2 (\Delta' K^2 \Delta)^{1/2}).$$

Нормальный закон служит основой практических расчетов в совершенно необозримом множестве областей. Приведем примеры, взятые из различных книг.

Пример 3 (Е. С. Вентцель, Л. А. Овчаров, 1973). Завод изготавливает шарики, номинальный диаметр которых равен d_0 , а фактический диаметр L — случайная величина, распределенная по нормальному закону с математическим ожиданием d_0 и средним квадратическим отклонением σ . После изготовления каждый шарик проходит контроль, причем бракуются все шарики, проходящие сквозь отверстие диаметром $d_0 - 2\sigma$, и все шарики, не проходящие сквозь отверстие диаметром $d_0 + 2\sigma$. Найти закон распределения диаметра шариков, прошедших контроль (не забракованных).

Пример 4 (И. В. Дунин-Барковский, Н. В. Смирнов, 1955). Математическое ожидание отклонения нормально распределенного диаметра втулки равно 30 мк, $\sigma = 10$ мк. Требуется определить вероятность того, что диаметр втулки будет находиться в пределах от 10 до 50 мк.

§ 7. Распределение суммы случайных величин

Пусть $\xi = (\xi_1, \dots, \xi_m)$, $\eta = (\eta_1, \dots, \eta_m)$ — случайные величины. Рассматривать их сумму $\xi + \eta$, как и любую другую функцию от них обеих, правомерно лишь в том случае, если они заданы на

одном вероятностном пространстве, т. е. существует случайная величина $\xi = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_m)$. Обозначим

$$F(x, y) = P(\xi < x, \eta < y).$$

Тогда

$$F_{\xi+\eta}(z) = P(\xi + \eta < z) = \iint_{x+y < z} dF(x, y). \quad (1)$$

Если ξ имеет плотность $p(x, y)$, то (1) преобразуется к виду

$$F_{\xi+\eta}(z) = \iint_{x+y < z} p(x, y) dx dy. \quad (2)$$

Пример. Двумерная случайная величина (ξ, η) равномерно распределена в круге $x^2 + y^2 \leq R^2$. Найти функцию распределения случайной величины $\xi + \eta$.
Имеем

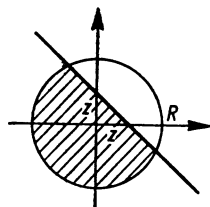


Рис. 20

$$F_{\xi+\eta}(z) = (\pi R^2)^{-1} \iint_A dx dy,$$

где A — сегмент, показанный на рис. 20 штриховкой. Двойной интеграл равен площади этого сегмента. Вычислив полученную площадь, находим

$$F_{\xi+\eta}(z) = 1 - \pi^{-1} (\arccos(z/R\sqrt{2}) + (z^2/2R^2)\sqrt{2R^2 - z^2}), \quad |z| \leq R\sqrt{2}.$$

При $z < -R\sqrt{2}$ имеем $F_{\xi+\eta}(z) = 0$, если $z > R\sqrt{2}$, то $F_{\xi+\eta}(z) = 1$.

Рассмотрим наиболее важный случай независимых случайных величин ξ, η . Имеем

$$\begin{aligned} F_{\xi+\eta}(z) &= \iint_{x+y < z} dF_{\xi}(x) dF_{\eta}(y) = \\ &= \int \left\{ \int_{y < z-x} dF_{\eta}(y) \right\} dF_{\xi}(x) = \int F_{\eta}(z-x) dF_{\xi}(x). \end{aligned}$$

Функцию распределения вида $\int F(z-x) dG(x)$, где F и G — функции распределения, называют *сверткой* F и G и обозначают $(F * G)(z)$. Таким образом, получено следующее важное утверждение: *функция распределения суммы двух независимых случайных величин равна свертке функций распределения этих величин:*

$$F_{\xi+\eta}(x) = \int F_{\xi}(x-y) dF_{\eta}(y) = \int F_{\eta}(x-y) dF_{\xi}(y). \quad (3)$$

(В предыдущей формуле заменено x на y и z на x , также замечено, что формула симметрична относительно F_{ξ} и F_{η} .)

Если распределения непрерывны, то равенство (3) перепишем так:

$$F_{\xi+\eta}(x) = \int_{-\infty}^{\infty} F_{\xi}(x-y) p_{\eta}(y) dy = \int_{-\infty}^{\infty} F_{\eta}(x-y) p_{\xi}(y) dy. \quad (4)$$

Если $p(x)$, $q(x)$ — плотности вероятности, то их сверткой назовем функцию

$$(p * q)(x) = \int_{-\infty}^{\infty} p(x-y) q(y) dy. \quad (5)$$

Имеем

$$\begin{aligned} \int_{-\infty}^{\infty} F_{\xi}(x-y) p_{\eta}(y) dy &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^x p_{\xi}(t-y) dt \right\} p_{\eta}(y) dy = \\ &= \int_{-\infty}^x \left\{ \int_{-\infty}^{\infty} p_{\xi}(t-y) p_{\eta}(y) dy \right\} dt = \int_{-\infty}^x (p_{\xi} * p_{\eta})(t) dt. \end{aligned}$$

Полученное равенство означает, что $\xi + \eta$ — непрерывная случайная величина и ее плотность есть свертка плотностей ξ и η :

$$p_{\xi+\eta}(x) = \int_{-\infty}^{\infty} p_{\xi}(x-y) p_{\eta}(y) dy = \int_{-\infty}^{\infty} p_{\eta}(x-y) p_{\xi}(y) dy. \quad (6)$$

Часто встречаются положительные случайные величины, т. е. такие, для которых плотность равна нулю при отрицательных значениях аргумента. Для независимых положительных случайных величин ξ , η формула (6) преобразуется к виду

$$p_{\xi+\eta}(x) = \int_0^x p_{\xi}(x-y) p_{\eta}(y) dy = \int_0^x p_{\eta}(x-y) p_{\xi}(y) dy. \quad (7)$$

Пример. Пусть ξ , η — независимые стандартные нормальные случайные величины. В этом случае

$$p_{\xi}(x) = p_{\eta}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

и тогда по формуле (6)

$$p_{\xi+\eta}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} [(x-y)^2 + y^2] \right\} dy.$$

Выражение в квадратных скобках преобразуем к сумме квадратов:

$$(x-y)^2 + y^2 = 2y^2 - 2xy + x^2 = (y\sqrt{2} - x/\sqrt{2})^2 + x^2/2.$$

Заменой $y\sqrt{2} - x/\sqrt{2} = t$ интеграл приводится к виду

$$p_{\xi+\eta}(x) = \frac{1}{2\pi \cdot \sqrt{2}} e^{-x^2/4} \int_{-\infty}^{\infty} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-x^2/(2 \cdot 2)}.$$

Записанное в таком виде выражение есть $\frac{2}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$ при $\sigma = \sqrt{2}$. Итак, $\xi + \eta$ — нормальная случайная величина с математическим ожиданием 0 и среднеквадратическим отклонением $\sqrt{2}$. Этот вывод согласуется с теоремой, изложенной в предыдущем параграфе.

§ 8. Комплексные случайные величины

Пусть ξ, η — одномерные случайные величины, заданные на одном вероятностном пространстве. Функцию элементарного события $\zeta = \xi + i\eta$, где i — мнимая единица, назовем *комплексной одномерной случайной величиной*. Аналогично, если на вероятностном пространстве заданы m -мерные величины $\xi = (\xi_1, \dots, \xi_m)$ и $\eta = (\eta_1, \dots, \eta_m)$, то функцию элементарного события $\zeta = (\zeta_1, \dots, \zeta_m)$, где $\zeta_k = \xi_k + i\eta_k$, $1 \leq k \leq m$, назовем *комплексной m -мерной случайной величиной*. В отличие от комплексных ранее рассматривавшиеся случайные величины называют *действительными*. Основные правила действий с комплексными случайными величинами следуют из правил действий с комплексными числами.

Если $\zeta = \xi + i\eta$, $\zeta' = \xi' + i\eta'$, то

$$\zeta + \zeta' = (\xi + \xi') + i(\eta + \eta'),$$

т. е. действительные и мнимые части суммы равны суммам действительных и соответственно мнимых частей слагаемых. Произведение комплексного числа $c = a + ib$ на комплексную случайную величину $\zeta = \xi + i\eta$ определяется формулой

$$c\zeta = (a\xi - b\eta) + i(a\eta + b\xi).$$

Эта формула сохраняется и в том случае, если c — одномерная комплексная случайная величина.

Скалярное произведение (t, ξ) m -мерного действительного вектора $t = (t_1, \dots, t_m)$ на $\xi = \xi + i\eta = (\xi_1 + i\eta_1, \dots, \xi_m + i\eta_m)$ определяется формулой

$$(t, \zeta) = \sum_{k=1}^m t_k \zeta_k = \sum_{k=1}^m t_k \xi_k + i \sum_{k=1}^m t_k \eta_k. \quad (1)$$

Компоненты m -мерной случайной комплексной величины можно представить в *показательной форме*

$$\zeta_k = r_k e^{i\theta_k}$$

или в *тригонометрической форме*

$$\zeta_k = r_k (\cos \theta_k + i \sin \theta_k),$$

где

$$r_k = (\xi_k^2 + \eta_k^2)^{1/2}, \quad \cos \theta_k = \xi_k / r_k, \quad \sin \theta_k = \eta_k / r_k.$$

Если θ_k определить произвольным образом при $r_k = 0$, то $(r_1, \dots, r_m, \theta_1, \dots, \theta_m)$ будет $2m$ -мерной действительной случайной величиной. Случайная величина $\bar{\zeta}_k = \xi_k - i\eta_k = r_k e^{-i\theta_k}$ называется *комплексно сопряженной* с величиной $\zeta_k = \xi_k + i\eta_k = r_k e^{i\theta_k}$. Имеем

$$\zeta_k \bar{\zeta}_k = |\zeta_k|^2 = r_k^2.$$

Математическое ожидание комплексной случайной величины $\zeta = \xi + i\eta$ определяется формулой

$$M\zeta = M\xi + iM\eta. \quad (2)$$

Задать m -мерную комплексную случайную величину ξ — все равно, что задать $2m$ -мерную действительную величину $(\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_m)$.

Назовем комплексные случайные величины ξ_j *независимыми*, если соответствующие $2m$ -мерные действительные величины независимы. Например, $\alpha + i\beta$ и $\gamma + i\delta$ независимы, если $(\alpha, \beta), (\gamma, \delta)$ независимы. В частности, если комплексные случайные величины — функции от различных действительных случайных величин, то они независимы.

Если ξ_1, \dots, ξ_n — независимые одномерные комплексные случайные величины, то

$$M(\xi_1, \dots, \xi_n) = M\xi_1 \dots M\xi_n, \quad (3)$$

т. е. математическое ожидание произведения равно произведению математических ожиданий.

Формулу (3) достаточно установить для $n = 2$, а затем применить индукцию. При $m = 2$ положим $\xi_1 = \alpha + i\beta$, $\xi_2 = \gamma + i\delta$, где $\alpha, \beta, \gamma, \delta$ — действительные величины. Тогда

$$\xi_1 \xi_2 = (\alpha\gamma - \beta\delta) + i(\beta\gamma + \alpha\delta), \quad M\xi_1 \xi_2 = M(\alpha\gamma - \beta\delta) + iM(\beta\gamma + \alpha\delta).$$

Поскольку по условию $(\alpha, \beta), (\gamma, \delta)$ независимы, то $M\alpha\gamma = M\alpha M\gamma$ и т. д. В результате получаем

$$\begin{aligned} M\xi_1 \xi_2 &= (M\alpha M\gamma - M\beta M\delta) + i(M\beta M\gamma - M\alpha M\delta) = \\ &= (M\alpha + iM\beta)(M\gamma + iM\delta) = M\xi_1 M\xi_2, \end{aligned}$$

что и требовалось доказать.

Если ξ — одномерная комплексная случайная величина, для которой $|\xi| \leq a$ с вероятностью 1, то

$$|M\xi| \leq a. \quad (4)$$

Для комплексных чисел $\omega_1, \dots, \omega_n$ имеем неравенство

$$|\omega_1 + \dots + \omega_n| \leq |\omega_1| + \dots + |\omega_n|.$$

Для дискретной величины ξ , принимающей значения z_k , $|z_k| \leq a$, с вероятностями p_k , имеем

$$|M\xi| = \left| \sum_{k=1}^n z_k p_k \right| \leq \sum_{k=1}^n |z_k| p_k \leq a \sum_{k=1}^n p_k = a,$$

а следовательно, неравенство (4) выполняется. На все остальные величины оно переносится с помощью монотонного предельного перехода.

В частности, если γ — действительная случайная величина, то, поскольку $|e^{i\gamma}| = 1$,

$$|Me^{i\gamma}| \leq 1. \quad (5)$$

§ 9. Характеристические функции

Пусть ξ — действительная случайная величина с функцией распределения $F_\xi(x)$. *Характеристической функцией* ξ (ее распределения) называется функция переменного $t \in \mathbb{R}$:

$$\varphi_\xi(t) = \mathbf{M} e^{it\xi} = \int e^{itx} dF_\xi(x). \quad (1)$$

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — m -мерная действительная случайная величина с функцией распределения $F_\xi(x) = F_\xi(x_1, \dots, x_m)$. *Характеристической функцией* ξ называется функция переменного $t = (t_1, \dots, t_m) \in \mathbb{R}^m$:

$$\varphi_\xi(t) = \mathbf{M} e^{i(t, \xi)} = \int e^{i(t, x)} dF_\xi(x) \quad (2)$$

или в развернутом виде

$$\begin{aligned} \varphi_\xi(t_1, \dots, t_m) &= \mathbf{M} \exp \left\{ i \sum_{k=1}^m t_k \xi_k \right\} = \\ &= \int \dots \int \exp \left\{ i \sum_{k=1}^m t_k x_k \right\} dF_\xi(x_1, \dots, x_m). \end{aligned}$$

Если ξ — непрерывная случайная величина, то в одномерном случае

$$\varphi_\xi(t) = \int_{-\infty}^{\infty} e^{itx} p_\xi(x) dx, \quad (3)$$

в общем случае

$$\begin{aligned} \varphi_\xi(t) &= \int_{-\infty}^{\infty} e^{i(t, x)} p_\xi(x) dx = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ i \sum_{k=1}^m t_k x_k \right\} \times \\ &\times p_\xi(x_1, \dots, x_m) dx_1 \dots dx_m. \end{aligned} \quad (4)$$

Установим вначале простейшие свойства характеристической функции.

1. Характеристическая функция существует для любой случайной величины ξ . Действительно, под интегралом в равенстве (2) — ограниченная непрерывная функция, а для них интеграл Стильеса существует (см. «Дополнение»).

2. Характеристическая функция ограничена по модулю единицей:

$$|\varphi_\xi(t)| \leq 1. \quad (5)$$

Для доказательства достаточно использовать свойство, доказанное в § 8.

3. Значение характеристической функции в точке нуль равно 1:

$$\varphi_\xi(0) = 1.$$

Действительно, тогда $e^{it\xi}$ — величина, с вероятностью 1 равная 1, так что и ее математическое ожидание равно 1.

4. Пусть $\varphi_{\xi}(t_1, \dots, t_m)$ — характеристическая функция случайной величины ξ . Положив некоторые t_k , равными нулю, получим характеристическую функцию случайной величины меньшей размерности, образованной в результате исключения ξ_k , соответствующих t_k , замененных нулями. В частности,

$$Me^{i(t_1\xi_1 + \dots + t_r\xi_r)} = \varphi_{\xi}(t_1, \dots, t_r, 0, \dots, 0). \quad (6)$$

Это свойство также следует из определения характеристической функции. Обращаем внимание на простоту получения характеристической функции случайной величины меньшей размерности по сравнению с плотностью: там требуется интегрирование по исключаемым x_k .

5. Пусть ξ_1, \dots, ξ_n — независимые случайные величины, $s_n = \xi_1 + \dots + \xi_n$. Тогда

$$\varphi_{s_n}(t) = \varphi_{\xi_1}(t) \dots \varphi_{\xi_n}(t), \quad (7)$$

т. е. характеристическая функция суммы независимых случайных величин есть произведение их характеристических функций.

Это свойство в более общем виде доказано в § 8: математическое ожидание произведения функций независимых случайных величин равно произведению их математических ожиданий. (Для равенства (7) достаточно требовать лишь независимости $\xi_1 + \dots + \xi_k$ и ξ_{k+1} при каждом k , $1 \leq k \leq n-1$.)

6. Пусть A — произвольная матрица порядка $n \times n$, b — произвольный вектор-столбец. Тогда

$$\varphi_{A\xi+b}(t) = \varphi_{\xi}(A't) e^{i(t,b)}, \quad (8)$$

где A' — матрица, транспонированная к матрице A .

Для доказательства заметим, что $(t, x) = t'x$ (произведение вектор-строки на вектор-столбец). Отсюда

$$\begin{aligned} \varphi_{\xi}(t) &= Me^{it'\xi}, \\ \varphi_{A\xi+b}(t) &= Me^{it'(A\xi+b)} = Me^{i(t'A)\xi + it'b} = \\ &= Me^{i(A't)\xi} e^{it'b} = \varphi_{\xi}(A't) e^{it'b} = \varphi_{\xi}(A't) e^{i(t,b)}, \end{aligned} \quad (9)$$

что и требовалось доказать. Единственный момент в этой цепочке равенств, требующий пояснения, — это замена $t'A$ на $(A't)'$. Эта замена законна, поскольку для произведения матриц A и B в транспонированной матрице меняется порядок произведения: $(AB)' = B'A'$; кроме того, $(A')' = A$.

Из равенства (8) следует, что для одномерной величины ξ

$$\varphi_{a\xi+b}(t) = \varphi_{\xi}(at) e^{ibt}. \quad (10)$$

7. Пусть $\xi = (\xi_1, \dots, \xi_m)$ и $\eta = (\eta_1, \dots, \eta_n)$ — независимые случайные величины с характеристическими функциями $\varphi_{\xi}(t_1, \dots, t_m)$ и $\varphi_{\eta}(t_{m+1}, \dots, t_{m+n})$. Тогда характеристическая функция случайной величины $\zeta = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n)$ есть произведение характеристических функций ξ и η :

$$\varphi_{\zeta}(t_1, \dots, t_{m+n}) = \varphi_{\xi}(t_1, \dots, t_m) \varphi_{\eta}(t_{m+1}, \dots, t_{m+n}). \quad (11)$$

Свойство (11) очевидно, так как в принятых предположениях случайная величина $e^{it\xi}$ представляется в виде произведения двух независимых случайных величин.

Равенство (11) имеет место также для произвольного числа независимых величин.

Пусть ξ — стандартная нормальная случайная величина. Ее характеристическая функция $\varphi(t)$, по определению, представляется интегралом

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx - x^2/2} dx.$$

Преобразуем показатель экспоненты:

$$itx - x^2/2 = -(x - it)^2/2 - t^2/2,$$

откуда

$$\varphi(t) = e^{-t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x - it)^2/2} dt.$$

Интеграл, стоящий после множителя $e^{-t^2/2}$, можно записать как $\frac{1}{\sqrt{2\pi}} \int_C e^{-z^2/2} dz$, где C — прямая комплексной плоскости, параллельная действительной оси. Поскольку $e^{-z^2/2}$ — аналитическая функция, убывающая к нулю при $z \rightarrow \infty$, если при этом z остается между прямой C и действительной осью, то по интегральной формуле Коши значение интеграла то же, что и интеграла по действительной оси, т. е.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1.$$

Окончательно получаем

$$\varphi(t) = e^{-t^2/2}. \quad (12)$$

По формуле (10) из (12) находим, что характеристическая функция нормальной случайной величины с математическим ожиданием a и дисперсией σ^2 имеет вид

$$\varphi_{\sigma\xi+a}(t) = e^{ita - \sigma^2 t^2/2}. \quad (13)$$

Пусть $\xi = (\xi_1, \dots, \xi_m)$, где ξ_i — независимые нормальные случайные величины, $M\xi_i = 0$, $D\xi_i = \sigma_i^2$. В силу формулы (11)

$$\varphi_{\xi}(t) = \exp \left\{ -\frac{1}{2} \sum_{k=1}^m \sigma_k^2 t_k^2 \right\}. \quad (14)$$

Заметим, что корреляционная матрица вектора ξ

$$K_{\xi} = M\xi\xi' = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_m^2 \end{pmatrix} \quad (15)$$

и в то же время выражение $\sum_{i=1}^m \sigma_i^2 t_i^2$ можно записать как $t' K_{\xi} t$. Следовательно, из (14) находим

$$\varphi_{\xi}(t) = \exp \left\{ -\frac{1}{2} t' K t \right\}. \quad (16)$$

Рассмотрим ортогональное преобразование $\xi = C\eta$, т. е. $\eta = C'\xi$. По формулам (8), (16) имеем

$$\varphi_{C'\xi}(t) = \varphi_{\xi}(Ct) = \exp \left\{ -\frac{1}{2} t' C' K C t \right\} = \exp \left\{ -\frac{1}{2} t' K_{\eta} t \right\}. \quad (17)$$

Пусть теперь $\zeta = \eta + a$, где a — произвольный вектор-столбец. Объединив (17) и (8), получим общую формулу

$$\varphi_{\zeta}(t) = \exp \left\{ -\frac{1}{2} t' K_{\zeta} t + i t' a \right\} \quad (18)$$

для характеристической функции невырожденной m -мерной нормальной случайной величины ζ . В скалярном виде эта формула записывается так:

$$\begin{aligned} \varphi_{\zeta}(t_1, \dots, t_m) = \\ = \exp \left\{ -\frac{1}{2} \sum_{k,j=1}^m \text{cov}(\zeta_k, \zeta_j) t_k t_j + i \sum_{k=1}^m t_k M \zeta_k \right\}. \end{aligned} \quad (19)$$

Формулу (18) или, что то же самое, (19) принимают в качестве определения характеристической функции произвольной (в том числе вырожденной) нормальной случайной величины.

Пусть $\zeta = (\zeta_1, \dots, \zeta_m)$ — m -мерная действительная случайная величина, $D\zeta_i < \infty$, $1 \leq i \leq m$. Если корреляционная матрица K_{ζ} имеет хотя бы один ненулевой элемент и характеристическая функция ζ имеет вид (18), то ζ называется m -мерной нормальной случайной величиной. Ранг распределения этой величины (размерность пространства, на котором оно сосредоточено) равен рангу матрицы K_{ζ} . Нормальное распределение невырождено, если K_{ζ} имеет ранг m .

Для установления дальнейших свойств характеристических функций нам понадобятся некоторые оценки, связанные с экспоненциальной функцией мнимого аргумента. Рассмотрим модуль разности этой функции и r -го отрезка ее ряда Тейлора:

$$R_r(x) = |e^{ix} - 1 - ix - \dots - (ix)^r/r!|.$$

Поскольку

$$\begin{aligned} i \int_0^x (e^{iu} - 1 - iu - \dots - (iu)^{r-1}/(r-1)!) du = \\ = (e^{ix} - ix - (ix)^2/2 - \dots - (ix)^r/r!) \Big|_0^x = \\ = e^{ix} - 1 - ix - \dots - (ix)^r/r!, \end{aligned}$$

то

$$R_r(x) \leq \left| \int_0^x R_{r-1}(u) du \right|, \quad r \geq 1. \quad (20)$$

Для $r = 0$

$$R_0(x) = |e^{ix} - 1| = \left| \int_0^x e^{iu} du \right| \leq |x|. \quad (21)$$

Из (20) и (21) по индукции получаем

$$R_r(x) \leq |x|^{r+1}/(r+1)!, \quad r \geq 0. \quad (22)$$

В то же время из выражения для $R_r(x)$ можно записать

$$R_r(x) \leq R_{r-1}(x) + |x|^r/r!, \quad r \geq 1.$$

Тогда из (22) и очевидного неравенства $R_0(x) \leq 2$ имеем

$$R_r(x) \leq 2|x|^r/r!, \quad r \geq 0. \quad (23)$$

Получены две оценки для $R_r(x)$, а именно: (22) и (23). В дальнейшем они будут комбинироваться: первая будет использоваться при малых $|x|$, вторая при больших.

8. Характеристическая функция $\varphi_\xi(t)$ случайной величины ξ равномерно непрерывна при $t \in \mathbb{R}^m$. При $h = (h_1, \dots, h_m)$

$$\begin{aligned} \Delta\varphi &= \varphi_\xi(t+h) - \varphi_\xi(t) = \int e^{i(t+h,x)} dF(x) - \int e^{itx} dF(x) = \\ &= \int e^{itx} (e^{i(h,x)} - 1) dF(x), \end{aligned}$$

откуда

$$|\Delta\varphi| \leq \int R_0(h, x) dF(x), \quad (24)$$

так как модуль интеграла не превосходит интеграл модуля. Разобьем пространство \mathbb{R}^m на две области: U_N , в которой $\sum |x_i| \leq N$, и V_N , в которой выполняется противоположное неравенство. Поскольку $V_{N+1} \subset V_N$ и $\bigcap_N V_N = \emptyset$, то по аксиоме непрерывности

$$\int_{V_N} dF(x) = P(V_N) \xrightarrow{N \rightarrow \infty} 0.$$

Поскольку, к тому же, $R_0 \leq 2$, то для некоторого N часть интеграла (24) по области V_N меньше $\varepsilon/2$. В области U_N вследствие (22) $R_0 \leq \sum |h_k x_k| < \varepsilon/2$ при $|h_k| < 1/2N$, так как $\sum |x_k| \leq N$. Отсюда находим, что правая часть неравенства (24) меньше ε при достаточно малых h_k . Требуемое доказано.

Большое практическое применение имеет вычисление моментов распределения с помощью дифференцирования характеристической функции. Справедливо следующее утверждение.

9. Пусть одномерная случайная величина ξ имеет конечный момент α_k порядка k . Тогда $\varphi_\xi(t)$ имеет непрерывную производную

порядка k , причем

$$\varphi_{\xi}^{(k)}(t) = i^k \int x^k e^{itx} dF(x), \quad (25)$$

где $F(x) = F_{\xi}(x)$ и, в частности,

$$\varphi_{\xi}^k(0) = i^k \alpha_k. \quad (26)$$

При $k = 0$ равенство (25) выполняется, если понимать под нулевой производной функции φ_{ξ} саму эту функцию. Докажем это равенство для значения $k = 1$. Рассмотрим разность отношения $\Delta\varphi/h$ и правой части (25) (наша цель доказать ее бесконечную малость). Имеем

$$\begin{aligned} & \left| \Delta\varphi/h - i \int x e^{itx} dF(x) \right| = \\ & = \left| \frac{1}{h} \int e^{itx} (e^{thx} - 1 - ihx) dF(x) \right| \leq \left| \frac{1}{h} \int R_1(hx) dF(x) \right|. \end{aligned}$$

Дальнейшие выкладки вполне аналогичны использованным при доказательстве непрерывности $\varphi_{\xi}(t)$: в силу (23)

$$\left| \frac{1}{h} \int_{|x|>N} R_1(hx) dF(x) \right| < \int_{|x|>N} |x| dF(x) < \varepsilon/2$$

при достаточно большом N , так как по предположению $\mathbf{M}|\xi| = \int |x| dF(x) < \infty$; в силу (22)

$$\left| \frac{1}{h} \int_{|x| \leq N} R_1(hx) dF(x) \right| \leq |h| \int_{|x| \leq N} x^2 dF(x) \leq |h| \cdot N < \varepsilon/2$$

при $|h| < 1/2N$. Итак, равенство (25), а вместе с ним и (26), установлено при $k = 1$. Для произвольного k это равенство доказывается по индукции. Переход от $k - 1$ к k ничем не отличается от произведенного нами перехода от 0 к 1: при конечном $\alpha_k = \int x^k dF(x)$ с подынтегральным выражением $x^k dF(x)$ можно обращаться так же, как и с $xdF(x)$ при $k = 1$.

В гл. 5 будет установлено, что распределение случайной величины однозначно определяется ее характеристической функцией. Это, однако, требует привлечения довольно сложного математического аппарата. Ситуация упрощается, когда ξ — непрерывная случайная величина. В этом случае формула (4) определяет характеристическую функцию как преобразование Фурье плотности. Из теории интеграла Фурье имеем следующую формулу обращения.

10. В любой точке x , в которой $p_{\xi}(x)$ непрерывна,

$$p_{\xi}(x) = \frac{1}{(2\pi)^m} \lim_{A_1, \dots, A_m \rightarrow \infty} \int_{-A_1}^{A_1} dt_1 \dots \int_{-A_m}^{A_m} e^{-i(t, x)} \varphi_{\xi}(t) dt_m. \quad (27)$$

Если $\int_{-\infty}^{\infty} |\varphi_{\xi}(t)| dt < \infty$, то для любого $x \in \mathbb{R}^m$

$$\begin{aligned} p_{\xi}(x) &= \frac{1}{(2\pi)^m} \int_{-\infty}^{\infty} dt_1 \dots \int_{-\infty}^{\infty} e^{-i(t,x)} \varphi_{\xi}(t) dt_m = \\ &= \frac{1}{(2\pi)^m} \int_{-\infty}^{\infty} e^{-i(t,x)} \varphi_{\xi}(t) dt; \end{aligned} \quad (28)$$

в этом случае $p_{\xi}(x)$ непрерывна при всех x .

§ 10. Распределение некоторых сумм

Характеристические функции позволяют находить распределение сумм независимых случайных величин.

1. Одномерная случайная величина с плотностью $p(x)$, равной 0 при $x < 0$ и равной $e^{-\lambda x}$ при $x > 0$, называется *экспоненциальной (экспоненциально распределенной)* случайной величиной с параметром $\lambda > 0$. Альтернативное название: *показательная случайная величина*. Ее характеристическая функция

$$\varphi(t) = \lambda \int_0^{\infty} e^{itx - \lambda x} dx = \lambda/(\lambda - it) = 1/(1 - it/\lambda). \quad (1)$$

Тогда

$$\varphi(t) = \sum_{n=0}^{\infty} (it)^n / \lambda^n.$$

Поскольку

$$\varphi(t) = \sum_{n=0}^{\infty} \varphi^{(n)}(0) t^n / n! = \sum_{n=0}^{\infty} i^n \alpha_n t^n / n!,$$

то при сравнении коэффициентов при t^n имеем

$$\alpha_n = n! / \lambda^n. \quad (2)$$

В частности, $\alpha_1 = 1/\lambda$.

2. Пусть $S_n = \xi_1 + \dots + \xi_n$, где ξ_k независимы и экспоненциальны с параметром 1. Тогда

$$\varphi_{S_n}(t) = 1/(1 - it)^n.$$

Как было выяснено в § 9, если $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx$, то

$$\varphi^{(k)}(t) = i^k \int_{-\infty}^{\infty} x^k e^{itx} p(x) dx.$$

Продифференцировав $1/(1 - it)$ $n - 1$ раз, получим

$$(n - 1)! i^{n-1} / (1 - it)^n = i^{n-1} \int_0^{\infty} x^{n-1} e^{itx - x} dx;$$

следовательно,

$$1/(1 - it)^n = \int_0^{\infty} x^{n-1} e^{itx-x} dx / (n-1)!$$

т. е. при $\lambda = 1$ плотность S_n равна $x^{n-1}e^{-x}/(n-1)!$ при $x > 0$ и 0 при $x < 0$. По формуле (9.10) находим, что при любом $\lambda > 0$

$$p_{S_n}(x) = \lambda^n x^{n-1} e^{-\lambda x} / (n-1)!, \quad x > 0; \quad p_{S_n}(x) = 0, \quad x < 0. \quad (2)$$

Данное распределение называется *распределением Эрланга* порядка n . Оно — частный случай *гамма-распределения* с параметрами $\lambda > 0$, $\alpha > 0$, имеющего плотность

$$p(x; \lambda, \alpha) = \lambda^\alpha x^{\alpha-1} e^{-\lambda x} / \Gamma(\alpha), \quad x > 0. \quad (3)$$

Методом контурного интегрирования показывается, что

$$\int_0^{\infty} e^{itx} p(x; \lambda, \alpha) dx = (1 - it/\lambda)^{-\alpha}. \quad (4)$$

Отсюда находим, что сумма независимых случайных величин, имеющих гамма-распределение с параметрами λ, α_1 и λ, α_2 , есть величина с гамма-распределением, параметрами которого будут $\lambda, \alpha_1 + \alpha_2$.

3. Пусть $\chi_n^2 = \xi_1^2 + \dots + \xi_n^2$, где ξ_k — независимые стандартные нормальные величины. Эта величина называется «*хи-квадрат*», а ее распределение — *распределением χ_n^2 («хи-квадрат» с n степенями свободы)*. Обозначим $\varphi_n(t) = M \exp \{it\chi_n^2\}$; через $p_n(x)$ обозначим плотность χ_n^2 . Тогда $\varphi_n(t) = \varphi_1^n(t)$, $n \geq 1$. Проще всего исследуется $p_2(x)$. Имеем

$$F_{\chi_2^2}(x) = P(\xi_1^2 + \xi_2^2 < x) = \frac{1}{2\pi} \iint_{u^2+v^2 < x} \exp \left\{ -\frac{1}{2}(u^2 + v^2) \right\} du dv.$$

Переходом к полярным координатам $u = \rho \cos \varphi$, $v = \rho \sin \varphi$ данный интеграл преобразуется к виду

$$\frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\sqrt{x}} e^{-\rho^2/2} \rho d\rho = 1 - e^{-x/2}, \quad (5)$$

откуда дифференцированием находим $p_2(x) = (1/2)e^{-x/2}$, $x > 0$, т. е. χ_2^2 — экспоненциальная случайная величина с параметром $\lambda = 1/2$. По формуле (1) $\varphi_2(t) = (1 - 2it)^{-1}$, откуда $\varphi_1(t) = (1 - 2it)^{-1/2}$, т. е.

$$\varphi_n(t) = (1 - 2it)^{-n/2}, \quad n = 1, 2, \dots \quad (6)$$

Двузначность квадратного корня устраняется тем, что $\varphi_1(t)$ непрерывна по t и $\varphi_1(0) = 1$. Эта формула — частный случай (4) при $\lambda = 1/2$, $\alpha = n/2$. Затем можно воспользоваться выражением (3):

$$p_n(x) = 2^{-n/2} x^{(n/2)-1} e^{-x/2} / \Gamma(n/2), \quad x > 0. \quad (7)$$

4. Рассмотрим случайную величину $\chi_2 = (\chi_2^2)^{1/2}$. Ее функция распределения выражается интегралом (5) с верхним пределом x вместо \sqrt{x} , т. е. $P(\chi_2 < x) = 1 - e^{-x^2/2}$, $x \geq 0$. Величина $\sigma\chi_2$ называется *распределенной по закону Релея* с параметром σ . Имеем

$$P(\sigma\chi_2 < x) = P(\chi_2 < x/\sigma) = 1 - e^{-x^2/2\sigma^2}, \quad x \geq 0, \quad (8)$$

$$p_{\sigma\chi_2}(x) = \frac{x}{\sigma} e^{-x^2/2\sigma^2}, \quad x \geq 0. \quad (9)$$

Непосредственное вычисление приводит к формулам

$$M(\sigma\chi_2) = \sigma \sqrt{\pi/2}, \quad D(\sigma\chi_2) = \sigma^2 (2 - \pi/2). \quad (10)$$

Заметим, что распределение Релея имеет применение в радиотехнике и радиолокации.

Пример. Амплитуда A радиолокационного сигнала, отраженного от цели, распределена по закону Релея; значения амплитуды в последовательных k обзорах независимы. Найти вероятность $p_k(z)$ того, что хотя бы в одном из этих обзоров значение амплитуды превзойдет пороговый уровень z .

Имеем

$$p_k(z) = 1 - (1 - \exp\{-z^2/2\sigma^2\})^k.$$

5. Пусть ξ_1, \dots, ξ_n — независимые равномерные в интервале $(0; 1)$ случайные величины. Требуется найти функцию распределения $F_n(x)$ суммы $S_n = \xi_1 + \dots + \xi_n$.

Удобно записать плотность ξ_k в виде

$$p_1(x) = E(x) - E(x-1),$$

где $E(x) = 0$ при $x \leq 0$, $E(x) = 1$ при $x > 0$. Тогда

$$F_n(x) = \int_{x_1+\dots+x_n < x} \dots \int \prod_{k=1}^n [E(x_k) - E(x_k-1)] dx_1 \dots dx_n.$$

Раскрыв квадратные скобки, получим под интегралом сумму слагаемых вида

$$(-1)^{\varepsilon_1+\dots+\varepsilon_n} E(x_1 - \varepsilon_1) \dots E(x_n - \varepsilon_n),$$

где ε_i равны 0 или 1. Число слагаемых, для которых $\varepsilon_1 + \dots + \varepsilon_n = m$, равно C_n^m . Сделаем замену переменных $x_k = \varepsilon_k + t_k$, $1 \leq k \leq m$. Тогда при $\varepsilon_1 + \dots + \varepsilon_n = m$

$$\begin{aligned} & \int_{x_1+\dots+x_n < x} \dots \int E(x_1 - \varepsilon_1) \dots E(x_n - \varepsilon_n) dx_1 \dots dx_n = \\ & = \int_{t_1+\dots+t_n < x-m} \dots \int E(t_1) \dots E(t_n) dt_1 \dots dt_n = \\ & = \int_{t_k > 0, t_1+\dots+t_m < x-m} \dots \int dt_1 \dots dt_m = \begin{cases} 0 & \text{при } x \leq m, \\ (x-m)^n/n! & \text{при } x > m \end{cases} \end{aligned}$$

(объем n -мерного симплекса). Итак,

$$F_n(x) = \sum_{0 \leq m \leq x} \frac{1}{m! (n-m)!} (-1)^m (x-m)^n \quad (11)$$

(формула Лобачевского).

6. Пусть ξ_1, \dots, ξ_n — независимые нормальные m -мерные случайные величины с математическими ожиданиями a_1, \dots, a_n и корреляционными матрицами K_1, \dots, K_n ; $S_n = \xi_1 + \dots + \xi_n$. Тогда

$$\begin{aligned} \Phi_{S_n}(t) &= \Phi_{\xi_1}(t) \dots \Phi_{\xi_n}(t) = \\ &= \exp \left\{ -\frac{1}{2} t' (K_1 + \dots + K_n) t + i t' (a_1 + \dots + a_n) \right\}. \quad (12) \end{aligned}$$

Следовательно, S_n есть нормальная случайная величина с математическим ожиданием $a_1 + \dots + a_n$ и корреляционной матрицей $K_1 + \dots + K_n$.

7. Пусть $\alpha = (\alpha_1, \dots, \alpha_l)$, $\beta = (\beta_1, \dots, \beta_l)$ — некоррелированные нормальные случайные величины, т. е. $\text{cov}(\alpha_i, \beta_j) = 0$, или в векторном виде $\mathbf{M} \alpha^0 (\beta^0)' = 0$. Докажем, что α и β независимы.

Можно выбрать пространства размерностей $m \leq j$ и $n \leq l$, на которых сосредоточены распределения α и β . Пусть, например, $(\alpha_1, \dots, \alpha_m)$ и $(\beta_1, \dots, \beta_n)$ — независимые нормальные величины, а все остальные α_i, β_i линейно выражаются через них. Тогда достаточно доказать, что $(\alpha_1, \dots, \alpha_m)$ и $(\beta_1, \dots, \beta_n)$ независимы. Если $\xi = (\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n)$, то матрица K_ξ имеет блочный вид, так что если $t = (t_1, \dots, t_{m+n})$, то $t' K_\xi t$ есть сумма функции от t_1, \dots, t_m и функции от t_{m+1}, \dots, t_{m+n} . Следовательно, $\Phi_\xi(t)$ есть произведение функции от t_1, \dots, t_m и функции от t_{m+1}, \dots, t_{m+n} . Применив формулу обращения (9.28), получим плотность ξ в виде произведения плотностей $(\alpha_1, \dots, \alpha_m)$ и $(\beta_1, \dots, \beta_n)$, что и требовалось доказать. Так же доказывается и более общее утверждение: если ξ_i — нормальные случайные величины произвольных размерностей, причем $\text{cov}(\xi_i, \xi_j) = 0$, $i \neq j$, то ξ_i независимы.

Докажем важное следствие. Пусть ξ_k — независимые одномерные нормальные случайные величины с параметрами a, σ^2 , $\bar{x} = \frac{1}{n} (\xi_1 + \dots + \xi_n)$, $\alpha_k = \xi_k - \bar{x}$. Тогда n -мерная случайная величина $(\alpha_1, \dots, \alpha_n)$ и случайная величина \bar{x} независимы. Для доказательства достаточно установить, что $\text{cov}(\alpha_k, \bar{x}) = 0$. Имеем

$$\begin{aligned} \text{cov}(\alpha_k, \bar{x}) &= \mathbf{M} \alpha_k \bar{x}^0 = \mathbf{M} (\xi_k^0 - \bar{x}^0) \bar{x}^0 = \\ &= \mathbf{M} \xi_k^0 \bar{x}^0 - \mathbf{D} \bar{x} = \frac{1}{n} \mathbf{M} \xi_k^0 \sum_{j=1}^n \xi_j^0 - \mathbf{D} \bar{x} = \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0, \end{aligned}$$

что и требовалось доказать.

8. Характеристическая функция биномиального распределения с параметрами n, p имеет вид

$$\Phi_n(t) = (q + p e^{it})^n; \quad (13)$$

при $n = 1$ величина $e^{it\xi_k}$, где ξ_k — индикатор успеха в k -м испытании, принимает значение e^{it} с вероятностью p и значение $1 = e^{it0}$ с вероятностью q ; степень получается ввиду того, что ξ_k независимы. Из формулы (13) видим, что $\varphi_{n+m}(t) = \varphi_n(t) \varphi_m(t)$, что очевидно и из вероятностных соображений.

9. Характеристическая функция $\psi_a(t)$ закона Пуассона с параметром a имеет вид

$$\psi_a(t) = e^{-a} \sum_{n=0}^{\infty} e^{int} a^n / n! = \exp \{a(e^{it} - 1)\}. \quad (14)$$

Отсюда

$$\psi_a(t) \psi_b(t) = \psi_{a+b}(t), \quad (15)$$

т. е. сумма независимых пуассоновских величин есть пуассоновская величина; параметр ее распределения равен сумме параметров слагаемых.

§ 11. Преобразования случайных величин

Пусть $\xi = (\xi_1, \dots, \xi_m)$ — случайная величина с функцией распределения $F(x)$, $x \in \mathbb{R}^m$. Если случайная величина η со значениями в \mathbb{R}^n имеет вид $\eta = f(\xi)$, то для функции распределения $G(y)$ величины η имеем универсальную формулу

$$G(y) = \int_{x: f(x) < y} dF(x). \quad (1)$$

Применим ее для нахождения распределения величины $t = \xi_1/\xi_2$, где ξ_1, ξ_2 — независимые стандартные нормальные случайные величины. Имеем

$$F_t(y) = \iint_{(x/z) < y} \varphi(x) \varphi(z) dx dz = \int_{-\infty}^{\infty} \varphi(z) dz \int_{(x/z) < y} \varphi(x) dx.$$

При $z > 0$ неравенство $(x/z) < y$ означает $x < yz$, откуда $\int \varphi(x) dx = \Phi(yz)$, при $z < 0$ имеем $1 - \Phi(yz) = \Phi(-yz)$.

Отсюда

$$\begin{aligned} F_t(y) &= 2 \int_0^{\infty} \varphi(z) \Phi(yz) dz, \quad p_t(y) = F'_t(y) = \\ &= 2 \int_0^{\infty} z \varphi(z) \varphi(yz) dz = \frac{1}{\pi} \int_0^{\infty} z \exp \left\{ -\frac{1}{2} z^2 (1 + y^2) \right\} dz; \end{aligned}$$

подстановкой $z^2 (1 + y^2) = u^2$ этот интеграл приводится к виду

$$p_t(y) = 1/(\pi (1 + y^2)). \quad (2)$$

Случайная величина $\sigma t + a$, где $\sigma \neq 0$ и a — постоянные, имеет плотность

$$p_{\sigma t + a}(x) = |\sigma| / (\pi (\sigma^2 + (x - a)^2)); \quad (3)$$

ее распределение называется *распределением Коши*. Математическое ожидание данной величины не существует.

Этим же приемом выводится формула для плотности случайной величины $t_n = \xi/\eta$, где ξ — стандартная нормальная случайная величина, $\eta = \chi_n/\sqrt{n}$, $\chi_n = (\chi_n^2)^{1/2}$, ξ и η независимы:

$$p_{t_n}(x) = \Gamma\left(\frac{n+1}{2}\right) \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} / \sqrt{\pi n} \Gamma\left(\frac{n}{2}\right). \quad (4)$$

Данное распределение называется *распределением Стьюдента* с n степенями свободы. Выражение (2) — частный случай (4) при $n = 1$. Поскольку при $|x| \rightarrow \infty$ плотность вида (4) имеет порядок малости $|x|^{-(n+1)}$, то момент α_k существует при $k \leq n-1$ и не существует при $k \geq n$.

Пусть ξ имеет плотность $p_\xi(x)$, $x \in \mathbb{R}^m$, $\eta = f(\xi)$, где f — функция со значениями в \mathbb{R}^m . Предположим, что в окрестности точки y существует конечное число обратных функций $x = \psi_k(y)$. Для окрестности V точки y имеем

$$P(\eta \in V) = \sum_k P(\xi \in V_k) = \sum_k \int_{V_k} p_\xi(x) dx, \quad (5)$$

где V_k — образ V при отображении $x = \psi_k(y)$.

Пусть $\psi_k(y)$ дифференцируемы в данной точке y . Тогда при V , стягивающейся к точке y , объемы V_k эквивалентны $|I_k| |V|$, где $|V|$ — объем V , I_k — якобиан преобразования ψ_k . Отсюда находим, что в данной точке плотность η существует и имеет вид

$$p_\eta(y) = \sum_k p_\xi \psi_k(y) |I_k|, \quad (6)$$

где

$$I_k = \det \left(\frac{\partial \psi_i(y)}{\partial y_j} \right). \quad (7)$$

Пример 1. Амплитуда радиосигнала изменяется по закону $A(t) = a(1 + \cos t)$. Найти плотность значения амплитуды в случайный момент времени τ , равномерно распределенный в интервале $(-\pi; \pi)$.

Согласно условию, $A(\tau) = a(1 + \cos \tau)$. При $0 < y < 2a$ имеем две функции $\psi_1(y)$ и $\psi_2(y)$, обратные к функции $y = a(1 + \cos x)$, а именно: $x = \pm \arccos((y/a) - 1)$ (рис. 21).

В одномерном случае якобиан преобразования $I_k = \psi_k(y)$. В то же время $p_\tau(x) = 1/2\pi$. В результате получаем

$$p_\eta(y) = 1/(\pi \sqrt{y(2a-y)}), \quad 0 < y < 2a.$$

Пример 2. Амплитуда импульса равна Ae^{-at^2} в момент t . Отсчет амплитуды снимается в момент ξ , распределенный по нормальному закону с математическим ожиданием 0 и дисперсией σ^2 . Найти плотность вероятности значения η амплитуды при отсчете,

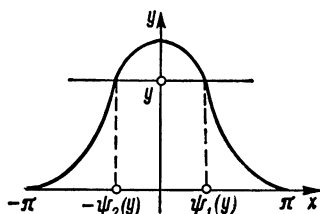


Рис. 21

Согласно условию, $\eta = Ae^{-a\xi^2}$. При $y < A$ уравнение $Ae^{-ax^2} = y$ имеет два корня:

$$x_1 = \psi_1(y) = \left(\frac{1}{a} \ln(A/y) \right)^{1/2} \quad \text{и} \quad x_2 = \psi_2(y) = -x_1.$$

Далее, $|\psi'_k(y)| = (2y(a \ln(A/y))^{1/2})^{-1}$, в то же время

$$p_{\xi}(x_k) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2 a} \ln(A/y) \right\} = \frac{1}{\sqrt{2\pi}} \left(\frac{y}{A} \right)^{1/2\sigma^2 a}.$$

Окончательно получаем

$$p_{\eta}(y) = \frac{1}{\sqrt{2\pi}} \left(\frac{y}{A} \right)^{1/2\sigma^2 a} / (y(a \ln(A/y))^{1/2}), \quad 0 < y < A.$$

§ 12. Условные распределения

Пусть ξ_1, ξ_2 — дискретные случайные величины. Их возможные значения будем обозначать символами i, j и т. д. По определению условной вероятности (§ 1.5)

$$\mathbf{P}(\xi_1 = i, \xi_2 = j) = \mathbf{P}(\xi_1 = i) \mathbf{P}(\xi_2 = j | \xi_1 = i). \quad (1)$$

Введя обозначения

$$p_{\xi_1}(i) = \mathbf{P}(\xi_1 = i), \quad p_{\xi_2|\xi_1}(j|i) = \mathbf{P}(\xi_2 = j | \xi_1 = i), \\ p_{\xi_1, \xi_2}(i, j) = \mathbf{P}(\xi_1 = i, \xi_2 = j),$$

можем переписать равенство (1) так:

$$p_{\xi_1, \xi_2}(i, j) = p_{\xi_1}(i) p_{\xi_2|\xi_1}(j|i). \quad (2)$$

Пусть A, B — любые множества возможных значений величин ξ_1, ξ_2 . Тогда, просуммировав (2) по $i \in A, j \in B$, получим

$$\mathbf{P}(\xi_1 \in A, \xi_2 \in B) = \sum_{i \in A} p_{\xi_1}(i) \sum_{j \in B} p_{\xi_2|\xi_1}(j|i). \quad (3)$$

Распределение $p_{\xi_2|\xi_1}(j|i)$ называется *условным распределением* дискретной случайной величины ξ_2 относительно дискретной случайной величины ξ_1 .

Пусть теперь ξ_1, ξ_2 — произвольные случайные величины. Обозначим через $F_{\xi_1}(x), F_{\xi_1, \xi_2}(x, y)$ функции распределения величин ξ_1 и (ξ_1, ξ_2) , т. е.

$$F_{\xi_1}(x) = \mathbf{P}(\xi_1 < x), \quad F_{\xi_1, \xi_2}(x, y) = \mathbf{P}(\xi_1 < x, \xi_2 < y).$$

Функцию $F_{\xi_2|\xi_1}(y|x)$ называют *условной функцией распределения* случайной величины ξ_2 относительно случайной величины ξ_1 , если для любых борелевских множеств A, B выполняется равенство

$$\mathbf{P}(\xi_1 \in A, \xi_2 \in B) = \iint_{x \in A, y \in B} dF_{\xi_1, \xi_2}(x, y) = \\ = \int_{x \in A} dF_{\xi_1}(x) \int_{y \in B} dF_{\xi_2|\xi_1}(y|x). \quad (4)$$

Если (ξ_1, ξ_2) — непрерывная случайная величина с плотностью $p_{\xi_1, \xi_2}(x, y)$, то равенству (4) удовлетворяет, т. е. может быть принята в качестве условной функции распределения, функция

$$F_{\xi_2|\xi_1}(y|x) = \int_{-\infty}^y (p_{\xi_1, \xi_2}(x, t)/p_{\xi_1}(x)) dt, \quad (5)$$

или проще

$$dF_{\xi_2|\xi_1}(y|x) = (p_{\xi_1, \xi_2}(x, y)/p_{\xi_1}(x)) dy. \quad (6)$$

Действительно,

$$\begin{aligned} \int_A p_{\xi_1}(x) dx \int_B dF_{\xi_2|\xi_1}(y|x) &= \int_A p_{\xi_1}(x) dx \int_B (p_{\xi_1, \xi_2}(x, y)/p_{\xi_1}(x)) dy = \\ &= \int_A dx \int_B p_{\xi_1, \xi_2}(x, y) dy = \mathbf{P}(\xi_1 \in A, \xi_2 \in B), \end{aligned} \quad (7)$$

т. е. пришли к равенству (4).

Если

$$F_{\xi_2|\xi_1}(y|x) = \int_{-\infty}^y p_{\xi_2|\xi_1}(t|x) dt, \quad (8)$$

то функция $p_{\xi_2|\xi_1}(y|x)$ называется *условной плотностью* случайной величины ξ_2 относительно случайной величины ξ_1 . Из равенства (7) заключаем, что данному определению удовлетворяет функция

$$p_{\xi_2|\xi_1}(y|x) = p_{\xi_1, \xi_2}(x, y)/p_{\xi_1}(x). \quad (9)$$

Пусть (ξ_1, ξ_2) — двумерная нормальная случайная величина; $\mathbf{M}\xi_1 = \mathbf{M}\xi_2 = 0$, $\mathbf{D}\xi_1 = \sigma_1^2$, $\mathbf{D}\xi_2 = \sigma_2^2$, $\text{cov}(\xi_1, \xi_2) = \rho\sigma_1\sigma_2$. Тогда, согласно § 3.6,

$$\begin{aligned} p_{\xi_1}(x) &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-x^2/2\sigma_1^2}, \\ p_{\xi_1, \xi_2}(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ &\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\}. \end{aligned}$$

После преобразования квадратичной формы в показателе экспоненты получаем

$$\begin{aligned} p_{\xi_2|\xi_1}(y|x) &= p_{\xi_1, \xi_2}(x, y)/p_{\xi_1}(x) = \\ &= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\rho^2\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\}. \end{aligned} \quad (10)$$

Формулу (10) можно истолковать так. При условии, что величина ξ_1 приняла значение x , величина ξ_2 нормальная с математическим ожиданием $\sigma_2\rho x/\sigma_1$ и дисперсией $(1-\rho^2)\sigma_2^2$. Можно также записать

$$\xi_2 = \sigma_2\rho\xi_1/\sigma_1 + \xi_3, \quad (11)$$

где первое и второе слагаемые независимы.

Обобщая этот прием, рассмотрим нормальную случайную величину $(\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n)$ со значениями в \mathbb{R}^{m+n} и найдем условную плотность $\eta = (\eta_1, \dots, \eta_n)$ относительно $\xi = (\xi_1, \dots, \xi_m)$. Предположим, что ξ — невырожденная m -мерная нормальная случайная величина (в противном случае некоторые ξ_j можно было бы исключить из рассмотрения). Найдем такое ортогональное преобразование $\xi^0 = C\gamma^0$, что $\gamma^0 = (\gamma_1^0, \dots, \gamma_m^0)$, где γ_k^0 — независимые нормальные величины с математическим ожиданием 0 и дисперсиями σ_k^2 . Будем искать выражение η^0 в виде

$$(\eta^0)' = (\gamma^0)' A + (\xi^0)', \quad (12)$$

где A — матрица порядка $m \times n$, а ξ^0 не зависит от γ^0 (следовательно, и от ξ). Умножим равенство (12) слева на γ^0 (вектор-столбец) и приравняем математические ожидания левой и правой частей:

$$M\gamma^0 (\eta^0)' = M\gamma^0 (\gamma^0)' A + M\gamma^0 (\xi^0)'.$$

В силу независимости γ и ξ последнее слагаемое равно нулю. Далее, $M\gamma^0 (\gamma^0)' = K_\gamma = C' K_\xi C$; $M\gamma^0 (\eta^0)' = C' M\xi^0 (\eta^0)'$.

Итак, переходим к равенству

$$C' M\xi^0 (\eta^0)' = K_\gamma A. \quad (13)$$

Матрицей, обратной к K_γ , является матрица

$$K_\gamma^{-1} = \begin{pmatrix} \sigma_1^{-2} & & 0 \\ & \ddots & \\ 0 & & \ddots & \\ & & & \sigma_m^{-2} \end{pmatrix}.$$

Умножив равенство (13) слева на K_γ^{-1} , получим выражение для матрицы A :

$$A = K_\gamma^{-1} C' M\xi^0 (\eta^0)'. \quad (14)$$

Транспонированием (12) находим

$$\eta^0 = A' \gamma^0 + \xi^0.$$

Умножив это равенство на (12) и взяв математическое ожидание, получим

$$K_\xi = K_\eta - A' K_\gamma A, \quad (15)$$

так как $A' M\gamma^0 (\xi^0)' = M\xi^0 (\gamma^0)' A = 0$ в силу некоррелированности ξ и γ . Возвращаясь к исходным переменным, имеем

$$\eta^0 = A' \gamma^0 + \xi^0 = A' C' \xi^0 + \xi^0. \quad (16)$$

Отсюда находим

$$p_{\eta|\xi}(y|x) = p_\xi(y - A' C' (x - M\xi) - M\eta). \quad (17)$$

Условным математическим ожиданием случайной величины η относительно случайной величины ξ называется такая функция $f(\eta)$,

что для любого борелевского множества A

$$\int_A f(x) dF_{\xi}(x) = \iint_{x \in A} y dF(x, y) = M(\eta; \xi \in A). \quad (18)$$

Как и в случае условной плотности, можно показать, что условию (18) удовлетворяет функция

$$f(x) = M(\eta | \xi = x) = \int_{-\infty}^{\infty} y p_{\eta|\xi}(y | x) dy. \quad (19)$$

Так, из формулы (17) получаем

$$M(\eta | \xi = x) = M\eta + A'C'(x - M\xi). \quad (20)$$

Пусть η — одномерная случайная величина, $(\xi_1, \dots, \xi_m, \eta)$ — $(m+1)$ -мерная нормальная величина. Поставим задачу аппроксимации η функций $g(\xi) = g(\xi_1, \dots, \xi_m)$ так, чтобы $M(\eta - g(\xi))^2$ было минимальным. Функция $g(\xi)$, удовлетворяющая этому условию, называется *регрессией* η на ξ . Имеем

$$M(\eta - g(\xi))^2 = \int dF_{\xi}(x) \int (y - g(x))^2 dF_{\eta|\xi}(y | x). \quad (21)$$

Чтобы минимизировать двойной интеграл, достаточно минимизировать внутренний интеграл при любом x . Обозначив $a(x) = M(\eta | \xi = x)$, получим

$$\begin{aligned} & \int (y - g(x))^2 dF_{\eta|\xi}(y | x) = \\ & = \int (y - a(x) + a(x) - g(x))^2 dF_{\eta|\xi} = \int (y - a(x))^2 dF_{\eta|\xi} + \\ & + 2(Q(x) - g(x)) \int (y - a(x)) dF_{\eta|\xi} + \int (a(x) - g(x))^2 dF_{\eta|\xi}. \end{aligned}$$

Первое слагаемое не зависит от $g(x)$, второе равно 0, а третье минимально при $g(x) = a(x)$. Итак, регрессия η на ξ равна

$$g(x) = M(\eta | \xi = x). \quad (22)$$

В случае нормальных величин $g(x)$ — линейная функция x , задаваемая формулой (20).

Задачи

1. Случайная величина ω распределена равномерно в интервале $(0; 1)$. Обозначим через ω' дробную часть величины $2\omega'$. (Например, при $\omega = 0,4$ имеем $\omega' = 0,8$; при $\omega = 0,6$ имеем $\omega' = 0,2$.) Найти функцию распределения двумерной случайной величины (ω, ω') . Будет ли эта величина непрерывной?

2. Функция распределения $F(x)$ одномерной случайной величины равна x при $0 < x < \frac{1}{2}$ и равна 1 при $x > \frac{1}{2}$. Вывести формулу для $\int x^n dF(x)$ ($n = 1, 2, \dots$).

3. Случайные величины ξ_k независимы и имеют плотность e^{-x} при $x > 0$. Найти математическое ожидание величины $\eta = \xi_1^2 \xi_2^2 \xi_3^3$.

4. Двумерная случайная величина (ξ_1, ξ_2) равномерно распределена в шестиугольнике с вершинами $(\cos \frac{k\pi}{3}; \sin \frac{k\pi}{3})$, $0 \leq k \leq 5$. Вычислить математическое ожидание случайной величины $\xi_1^2 + \xi_2^2$.

5. Случайная величина (ξ, η) имеет плотность вида $c/(1+x^2+y^2)$. Найти плотность случайной величины ξ .

6. Пусть ξ, η — независимые случайные величины, причем $p_\eta(x) \leq c$ для всех x . Доказать, что $p_{\xi+\eta}(x)$ существует и во всех точках ее непрерывности $p_{\xi+\eta}(x) \leq c$.

7. Случайные величины α, β, γ независимы и равномерно распределены в интервале $(0; 1)$. Найти плотность величины ζ , равной средней из α, β, γ по величине (например, при $\alpha = 0,1, \beta = 0,9, \gamma = 0,7$ имеем $\zeta = 0,7$).

8. Случайная величина ξ имеет плотность, равную $cx^2(1-x)^2$ при $0 < x < 1$ и равную 0 вне интервала $(0; 1)$. Найти постоянную c и коэффициент корреляции величин ξ и ξ^2 .

9. Случайные величины ξ, η принимают значения ± 1 с вероятностью $1/2$. Выразить $P(\xi = \eta)$ через коэффициент корреляции величин ξ и η .

10. Какому условию удовлетворяет распределение одномерной случайной величины ξ , если $M \sin^2 \xi = 0$?

11. Обозначим через s_k число успехов в первых k испытаниях последовательности независимых испытаний с вероятностью успеха p . Вывести формулу для $\text{cov}(s_k, s_m)$ для произвольных $k \geq 0, m \geq 0$.

12. Случайная величина (ξ, η, ζ) имеет корреляционную матрицу вида

$$\begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}.$$

Найти коэффициент корреляции случайных величин $\xi + \eta$ и $\eta - \zeta$.

13. Записать формулу для распределения суммы независимых случайных величин ξ, η , где ξ — биномиальная величина с параметрами $n = 4, p = 1/2$, η имеет плотность $1 - |x|$ при $|x| < 1$.

14. Найти Me^{ξ} , где ξ — равномерная величина в интервале $(0; \frac{\pi}{3})$.

15. Пусть x — одномерная случайная величина с характеристической функцией $\varphi(t)$ и конечным $M\xi^4$; $\psi(t) = \ln \varphi(t)$ (имеется в виду ветвь логарифма, для которой $\psi(0) = 0$). Найти формулы для первых четырех производных функции $\psi(t)$ в точке $t = 0$. Какой вид примут эти формулы при $M\xi = 0$?

16. Пусть $\omega_1, \omega_2, \omega_3$ — независимые случайные величины с плотностью $2e^{-2x}$, $x > 0$. Определим ζ формулой

$$\zeta = \min \{\omega_1 + \omega_2, \omega_3\}.$$

Найти плотность ζ .

17. Трехмерная случайная величина (α, β, γ) имеет плотность $e^{-(x+y+z)}$ при $x > 0, y > 0, z > 0$. Найти плотность двумерной величины (ξ, η) , где $\xi = \alpha/(\alpha + \beta + \gamma), \eta = \beta/(\alpha + \beta + \gamma)$.

18. Пусть ξ_1, \dots, ξ_n — независимые случайные величины с плотностью $\lambda e^{-\lambda x}$, $x > 0$. Методом характеристических функций найти плотность величины

$$\zeta = \xi_1 + \frac{1}{2} \xi_2 + \dots + \frac{1}{n} \xi_n.$$

19. Случайные величины ξ, η, ζ нормальны; $M\xi = M\eta = M\zeta = 0$; $D\xi = D\eta = D\zeta = \sigma^2$; $r(\xi, \eta) = r(\xi, \zeta) = r(\eta, \zeta) = \rho$. Найти $g(x, y) = M(\zeta | \xi = x, \eta = y)$.

20. Обозначим $\varphi(t) = 1 - |t|$ при $|t| \leq 1, \varphi(t) = 0$ при $|t| > 1$. Методом преобразования Фурье доказать, что $\varphi(t)$ — характеристическая функция, и найти соответствующую ей плотность вероятности.

21. На основании результата предыдущей задачи доказать, что если $\varphi(t)$ — четная, непрерывная, вогнутая при $t > 0$ кусочно-линейная функция, $\varphi(0) = 1, \varphi(t) = 0$ при $t \geq T$, то $\varphi(t)$ — характеристическая функция.

§ 1. Цепь Маркова с дискретным временем

Рассмотрим последовательность испытаний, занумерованных числами $0, 1, 2, \dots$, с возможными исходами E_1, E_2, \dots . Пусть e_n — исход n -го испытания. Назовем последовательность $(e_n, n \geq 0)$ *цепью Маркова*, если при известном исходе n -го испытания исход $(n+1)$ -го испытания не зависит от исходов 0 -го, \dots , $(n-1)$ -го испытаний:

$$\begin{aligned} P(e_{n+1} = E_j | e_n = E_i, e_0 = E_{i_0}, \dots, e_{n-1} = E_{i_{n-1}}) = \\ = P(e_{n+1} = E_j | e_n = E_i). \end{aligned} \quad (1)$$

Свойство, выраженное равенством (1), называют *свойством марковости* и кратко формулируют так: *при фиксированном настоящем будущее не зависит от прошлого*. Для любой последовательности испытаний с возможными исходами E_i можно записать

$$\begin{aligned} P(e_0 = E_{i_0}, \dots, e_n = E_{i_n}) = \\ = P(e_0 = E_{i_0}) P(e_1 = E_{i_1} | e_0 = E_{i_0}) P(e_2 = E_{i_2} | e_0 = \\ = E_{i_0}, e_1 = E_{i_1}) \dots P(e_n = E_{i_n} | e_0 = E_{i_0}, \dots, e_{n-1} = E_{i_{n-1}}); \end{aligned}$$

для цепи Маркова это соотношение на основании (1) значительно упрощается:

$$\begin{aligned} P(e_0 = E_{i_0}, \dots, e_n = E_{i_n}) = \\ = P(e_0 = E_{i_0}) P(e_1 = E_{i_1} | e_0 = E_{i_0}) P(e_2 = E_{i_2} | e_1 = E_{i_1}) \dots \\ \dots P(e_n = E_{i_n} | e_{n-1} = E_{i_{n-1}}). \end{aligned} \quad (2)$$

Цепь Маркова называется *однородной*, если $P(e_n = E_j | e_{n-1} = E_i)$ не зависит от n .

С помощью цепей Маркова описывают движение материальных частиц, а также поведение различных систем. Исход n -го испытания часто называют *положением частицы* на n -м шаге либо *состоянием системы* на n -м шаге (в момент n). Мы будем называть e_n *состоянием цепи Маркова* на n -м шаге. Если E_i — числа, то e_n — случайные величины, обычно обозначаемые греческими буквами, например ξ_n . В таком случае говорят, что случайные величины ξ_n образуют *цепь Маркова*. Основная часть теории цепей Маркова не зависит от того, какова природа состояний. Поэтому последние можно закодировать числами: $1, 2, \dots, r$, если множество состояний конечно, и $0, 1, 2, \dots$, если это множество счетно. В дальнейшем будем рассматриваться только однородные цепи Маркова, в связи с чем слово «однородные» будет опускаться. Если $\xi_n = i$, $\xi_{n+1} = j$, то говорят, что произошел *переход* цепи Маркова из состояния i в состояние j .

Для характеристики цепи Маркова $(\xi_n, n \geq 0)$ задаются:

начальные вероятности $p_i^{(0)} = P(\xi_0 = i)$;
 вероятности перехода $p_{ij} = P(\xi_{n+1} = j | \xi_n = i)$ (из состояния i в состояние j за один шаг).

Матрица $P = (p_{ij})$ называется *матрицей перехода* цепи Маркова. Для цепей Маркова со счетным множеством состояний P есть матрица с бесконечным числом строк и столбцов. Вектор $p^{(0)} = (p_i^{(0)})$ называется *начальным распределением* цепи Маркова.

Имеем $\sum_j p_{ij} = 1$, так как переход из i в одно из состояний j — достоверное событие. Матрица с неотрицательными элементами, суммы которых по строкам равны 1, называется *стохастической матрицей*. Таким образом, матрица перехода цепи Маркова является стохастической матрицей. Равенство (2) можно записать так:

$$P(\xi_0 = i_0, \dots, \xi_n = i_n) = p_{i_0}^{(0)} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n}. \quad (3)$$

Равенство (3) также может быть взято в качестве определения однородной цепи Маркова, так как из него в силу определения условной вероятности последовательно получаем (1): например,

$$\begin{aligned} P(\xi_2 = j | \xi_0 = i_0, \xi_1 = i) &= \\ &= P(\xi_0 = i_0, \xi_1 = i, \xi_2 = j) / P(\xi_0 = i_0, \xi_1 = i) = \\ &= p_{i_0}^{(0)} p_{i_0 i} p_{ij} / (p_{i_0}^{(0)} p_{i_0 i}) = p_{ij}. \end{aligned}$$

Пусть $\omega_0, \omega_1, \omega_2, \dots$ — исходы независимых испытаний. Положим

$$\xi_0 = f_0(\omega_0), \xi_n = f(\xi_{n-1}, \omega_n), \quad n \geq 1. \quad (4)$$

Тогда, очевидно, $(\xi_n, n \geq 0)$ — цепь Маркова. Если вероятностные меры, определяющие исходы $\omega_n, n \geq 1$, совпадают то $(\xi_n, n \geq 0)$ — однородная цепь Маркова. Обратно, любую цепь Маркова можно представить в виде (4). Пусть, например, $(\xi_n, n \geq 0)$ — цепь Маркова с двумя состояниями 1 и 2 и матрицей перехода

$$P = \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix}, \quad (5)$$

причем $a \leq b$. Для любого $n \geq 1$ рассмотрим испытание с тремя исходами: $\omega_n = 0$ с вероятностью a , $\omega_n = 1$ с вероятностью $b - a$, $\omega_n = 2$ с вероятностью $1 - b$. При $\xi_{n-1} = 1$ положим $\xi_n = 1$ при $\omega_n = 0$, $\xi_n = 2$ при $\omega_n \geq 1$. При $\xi_{n-1} = 2$ положим $\xi_n = 1$ при $\omega_n \leq 1$, $\xi_n = 2$ при $\omega_n = 2$. При $n = 0$ зададим испытание с исходами $\omega_0 = 1, 2$, имеющими вероятности $p_1^{(0)}, p_2^{(0)}$, и положим $\xi_0 = \omega_0$. Этими условиями последовательность (ξ_n) задается как цепь Маркова. Если $\xi_{n-1} = i$, то, подставив вероятности исходов, благоприятствующих событию $\{\xi_n = j\}$, легко убедиться, что это те же вероятности, которые являются элементами матрицы (4). Например,

$$P(\xi_n = 1 | \xi_{n-1} = 2) = P(\omega_n = 0) + P(\omega_n = 1) = a + (b - a) = b.$$

Приведем некоторые примеры цепей Маркова.

Пример 1. Пусть S_n — число успехов в n независимых испытаниях с вероятностью успеха p в каждом; $S_0 = 0$. Тогда (S_n) — цепь Маркова, так как $S_n = S_{n-1} + I_n$, где I_n — индикатор успеха в n -м испытании, так что имеем соотношение вида (4): I_n не зависит от $(I_j, j < n)$. Матрица перехода имеет вид

$$P = \begin{pmatrix} q & p & & & \\ & q & p & & \\ & & q & p & \\ & & & q & p \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Поскольку состояния цепи Маркова имеют вид $0, 1, 2, \dots$, то строки и столбцы P нумеруются теми же числами.

Здесь и далее пустые поля в матрице означают, что они заполнены нулями: так, нулевая строка имеет вид $(q \ p \ 0 \ 0 \ \dots)$.

Вообще, если S_n — положение частицы на n -м шаге при неограниченном или ограниченном случайном блуждании, то $((S_n), n \geq 0)$ — цепь Маркова.

Пример 2 (классический, принадлежит А. А. Маркову — создателю теории цепей Маркова). Имеется бесконечная последовательность урн. В каждой урне вначале находится два шара: черный и белый. На n -м шаге ($n \geq 0$) из n -й урны наудачу вынимается шар и бросается в $(n+1)$ -ю урну; следовательно, при $n \geq 1$ один шар выбирается из трех. Закодируем черный шар единицей, белый — нулем. Этим самым определена случайная величина ξ_n — результат выбора шара из n -й урны. Последовательность $(\xi_n, n \geq 0)$ есть цепь Маркова. Действительно, пусть при n -м испытании выпала единица. Тогда в $(n+1)$ -й урне оказалось две единицы и один нуль; следовательно, выбор единицы имеет вероятность $2/3$. Аналогичная ситуация и при выборе нуля на n -м шаге. Короче говоря, матрицу перехода можно задать формулой

$$P(\xi_{n+1} = i | \xi_n = i) = 2/3, \quad P(\xi_{n+1} = 1 - i | \xi_n = i) = 1/3$$

или записью

$$P = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}.$$

Начальное распределение задается формулой $p_0^{(0)} = p_1^{(0)} = 1/2$, или в матричном виде $p^{(0)} = (1/2, 1/2)$.

Пример 3. Пусть $(v_i, i \geq 0)$ — бернуллиева последовательность; $P(v_i = 0) = P(v_i = 1) = 1/2$. Определим $\xi_0 = v_0$; $\xi_n = v_{n-1} \oplus v_n$ (здесь \oplus — символ сложения по модулю два: $0 \oplus 0 = 1 \oplus 1 = 0, 0 \oplus 1 = 1 \oplus 0 = 1$). К формуле, определяющей ξ_n , можно применить формулу (4); следовательно, (ξ_n) — цепь Маркова. Имеем $P(\xi_n = j | \xi_{n-1} = i) = P(i \oplus v_n = j) = P(v_n = i + j) = 1/2$, т. е.

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

По формуле (3) получаем

$$P(\xi_0 = i_0, \xi_1 = i_1, \dots, \xi_n = i_n) = \frac{1}{2} \frac{1}{2} \dots \frac{1}{2} = \frac{1}{2^{n+1}},$$

т. е. всевозможные значения $\xi_0, \xi_1, \dots, \xi_n$ независимы. Приходим к выводу, что $(\xi_n, n \geq 0)$ — бернуллиева последовательность с $p = 1/2$, как и исходная последовательность (v_n) , из которой она получена.

Следующие примеры дают представление о разнообразии применений цепей Маркова.

Пример 4. При доводке технического изделия устраняются технологические дефекты, но могут вноситься новые. Каждый дефект на n -м шаге в результате корректировки превращается в $0, 1$ или 2 дефекта на $(n+1)$ -м шаге с вероят-

ностью соответственно p_0 , p_1 и p_2 . Обозначим через ξ_n число дефектов после n -й корректировки. Тогда, если $\xi_{n-1} = i$, то $\xi_n = i + m_i(2) - m_i(0)$, где $m_i(0)$ — число устраненных, $m_i(2)$ — число удвоившихся дефектов. Если исходы корректировок различных дефектов независимы, то имеем i независимых испытаний с исходами 0, 1, 2, и тогда $m_i(0)$, $m_i(2)$ — соответственно число исходов 0, 2 в этих испытаниях. Из § 2.8

$$P_i(a, b, c) = \frac{i!}{a! b! c!} p_0^a p_1^b p_2^c,$$

где a, b, c — неотрицательные целые числа, в сумме равные i . Отсюда

$$p_{ij} = \sum_{i+c=a=j} P_i(a, b, c).$$

Пример 5. При расчете систем хранения, передачи и обработки текстовой информации желательнее иметь вероятностную модель реальной речи. При решении определенных задач хорошо оправдывает себя модель источника языковой информации, в которой слова выбираются из реального словаря, а вероятности следования одного слова за другим вычислены на основании обработки реальных текстов очень большого объема. Например, если в реальных текстах всегда после слова «бразды» идет «правления», то вероятность перехода первого во второе равна 1. Вот образец текста, полученного в результате реализации такого случайного источника информации¹:

ОБЩЕСТВО ИМЕЛО ВЫРАЖЕНИЕ МГНОВЕННОГО ОРУДИЯ К ДОСТИЖЕНИЮ ДОЛЖНОСТЕЙ ОДИН В РАСЧЕТЫ НА БЕЗНРАВСТВЕННОСТИ В ПОЭЗИИ РЕЗВИТЬСЯ ВСЕ ГРЫЗЕТ СВОИ БРАЗДЫ ПРАВЛЕНИЯ НАЧАЛА ЕГО ПОШЛОЙ

Пример 6. В одной из моделей исследования операций рассматривается противоборство сторон А и Б, располагающих вначале соответственно i_0 и j_0 огневыми точками. Если на n -м шаге сторона А имеет $i > 0$ уцелевших огневых точек, а сторона Б имеет $j > 0$, то вероятность того, что следующей будет подавлена огневая точка стороны А, равна

$$C_2 / (C_1 i + C_2 j) = f(i, j),$$

где отношение C_1/C_2 характеризует сравнительную эффективность средств стороны А по отношению к Б. С вероятностью 1 — $f(i, j)$ подавляется огневая точка стороны Б.

Состояниями цепи Маркова являются точки (i, j) , где $i, j \in \mathbb{Z}^+$. Если $i > 0$, $j > 0$, то вероятность перехода из (i, j) в $(i-1, j)$ равна $f(i, j)$. Переход из (i, j) в $(i, j-1)$ имеет вероятность $1 - f(i, j)$. Остальные переходы имеют вероятность 0. При $j = 0$ либо $i = 0$ переход из (i, j) в (i, j) имеет вероятность 1 (уцелевшие огневые точки сохраняются).

§ 2. Рекуррентные формулы

Обозначим $p_i^{(n)} = P(\xi_n = i)$, $p^{(n)} = (p_i^{(n)})$, т. е. $p^{(n)}$ — вектор, компонентами которого являются $p_i^{(n)}$. Тогда, просуммировав (1.1) по всем возможным значениям i_0, \dots, i_{n-1} и положив $i_0 = i$, $i_n = j$, получим

$$p_j^{(n)} = \sum_{i_1, \dots, i_{n-1}} p_i^{(0)} p_{i i_1} \dots p_{i_{n-1} j}, \quad (1)$$

или

$$p^{(n)} = p^{(0)} P \dots P = p^{(0)} P^n,$$

¹ См.: К. Шеннон. Работы по теории информации и кибернетике / Пер. с англ. под. ред. Р. Л. Добрушина, О. Б. Лупанова. — М., 1963. — С. 225.

так как суммированию по промежуточным индексам соответствует произведение матриц.

По формуле полной вероятности

$$p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}, \quad (2)$$

где $p_{ij}^{(n)} = P(\xi_n = j \mid \xi_0 = i)$ — вероятность перехода цепи Маркова за n шагов из состояния i в состояние j . Формулу (2) можно записать так:

$$p^{(n)} = p^{(0)} P^n, \quad (3)$$

где

$$P^{(n)} = \| p_{ij}^{(n)} \|$$

— матрица, называемая *матрицей перехода цепи Маркова* за n шагов. Сравнив (3) с (1), видим, что

$$p^{(0)} P^{(n)} = p^{(0)} P^n, \quad (4)$$

откуда следует важная формула

$$P^{(n)} = P^n. \quad (5)$$

Взяв $p_i^{(0)} = 1$, $p_k^{(0)} = 0$ при $k \neq i$, в формуле (4) слева получим i -ю строку матрицы $P^{(n)}$, справа — i -ю строку матрицы P^n . Отсюда при любом m , $1 \leq m \leq n$, получаем

$$p^{(m)} P^{(n-m)} = p^{(n)}, \quad (6)$$

$$P^n = P^m P^{n-m}, \quad (7)$$

или в скалярном виде

$$p_j^{(n)} = \sum_i p_i^{(m)} p_{ij}^{(n-m)}, \quad (8)$$

$$p_{ij}^{(n)} = \sum_k p_{ik}^{(m)} p_{kj}^{(n-m)}. \quad (9)$$

Соотношения (6) — (9) называются *уравнениями Маркова*. Наиболее часто они используются при $m = 1$ и $m = n - 1$, что позволяет рекуррентным образом вычислять $p_j^{(1)}$, $p_j^{(2)}$, ... (соответственно $p_{ij}^{(1)} = p_{ij}$, $p_{ij}^{(2)}$, ...). Приведем формулы для $m = n - 1$ как наиболее используемые:

$$P^{(n)} = P^{(n-1)} P, \quad p^{(n)} = p^{(n-1)} P, \quad (10)$$

или в скалярном виде

$$p_{ij}^{(n)} = \sum_k p_{ik}^{(n-1)} p_{kj}, \quad p_j^{(n)} = \sum_k p_k^{(n-1)} p_{kj}. \quad (11)$$

Следует заметить, что, например, для вычисления P^{57} вместо 56 умножений матриц достаточно сделать 8 умножений:

$$P^2 = P P, \quad P^4 = P^2 P^2, \quad P^8 = P^4 P^4, \quad P^{16} = P^8 P^8, \\ P^{32} = P^{16} P^{16}, \quad P^{48} = P^{32} P^{16}, \quad P^{56} = P^{48} P^8, \quad P^{57} = P^{56} P.$$

Пример. В некоторых ЭВМ арифметические операции реализуются на базе сложения и умножения двоичных чисел по модулю два. Предполагая, что при n суммированиях двоичных чисел каждое суммирование, независимо от остальных, приводит к ошибке (0 вместо 1 или наоборот) с вероятностью q , вычислить вероятность p_n безошибочности результата.

Заметим, что результат будет безошибочным в том и только в том случае, если допущено четное число ошибок. Если S_n — индикатор четности числа ошибок при n суммированиях, то $p_n = P(S_n = 0)$. Как и в примере 3 § 1, $S_n = S_{n-1} \oplus \oplus v_n$, однако здесь $v_n = 0$ с вероятностью $p = 1 - q$, $v_n = 1$ с вероятностью q .

Матрица перехода

$$P = \begin{pmatrix} p & q \\ q & p \end{pmatrix}.$$

По формуле (11)

$$p_0^{(n)} = p_0^{(n-1)}p + p_1^{(n-1)}q,$$

или, так как $p_0^{(n)} = p_n$, $p_1^{(n)} = 1 - p_n$,

$$p_n = p_{n-1}p + (1 - p_{n-1})q. \quad (12)$$

Удобно сделать замену

$$p = \frac{1}{2}(1 + \Delta); \quad q = \frac{1}{2}(1 - \Delta), \quad p_n = \frac{1}{2}(1 + \delta_n), \quad 1 - p_n = \frac{1}{2}(1 - \delta_n). \quad (13)$$

Тогда (12) приводится к виду

$$\frac{1}{2}(1 + \delta_n) = \frac{1}{4}(1 + \delta_{n-1})(1 + \Delta) + \frac{1}{4}(1 - \delta_{n-1})(1 - \Delta),$$

или после сведения подобных членов

$$\delta_n = \delta_{n-1}\Delta, \quad n \geq 1. \quad (14)$$

Поскольку $\xi_0 = 0$, то $P(\xi_n = 0) = 1 = (1 + \delta_0)/2$, откуда $\delta_0 = 1$. Тогда из (14) имеем $\delta_n = \Delta^n$, и окончательно

$$p_n = (1 + \Delta^n)/2. \quad (15)$$

Так, при $q = 10^{-7}$ имеем $\Delta = 1 - 2q = 1 - 2 \cdot 10^{-7}$. При $n = 2,5 \cdot 10^8$ имеем

$$p_n = \frac{1}{2}(1 + (1 - 2 \cdot 10^{-7})^{2,5 \cdot 10^8}) \approx 0,5(1 + e^{-0,5}) \approx 0,8033.$$

Заметим, что вероятность полного отсутствия ошибок при n суммированиях $(1 - q)^n \approx e^{-0,25} \approx 0,7788$; таким образом, с вероятностью $\approx 0,0245$ ошибки присутствуют, но взаимно уничтожаются.

Рассмотрим *цепи Маркова с запретами*. Пусть A — некоторое подмножество множества состояний цепи Маркова $(\xi_n, n \geq 0)$. Обозначим через $p_j^{(n)}(A)$ вероятность события $\{\xi_n = j; \xi_m \in A, 0 \leq m \leq n - 1\}$. Эта вероятность понимается как вероятность нахождения цепи Маркова в состоянии j в момент n при запрете выхода из множества состояний A .

Пример. Параметр технического устройства в n -й момент времени равен ξ_n , где (ξ_n) — цепь Маркова. При выходе ξ_n из допустимой области A наступает либо отказ, либо аварийное отключение устройства. Тогда $p_j^{(n)}(A)$ — вероятность того, что в n -й момент времени параметр равен j и при этом до $(n - 1)$ -го момента включительно указанные нежелательные события не наступили,

Каким же образом рассчитываются вероятности $p_j^n(A)$? Для них также существуют рекуррентные формулы. Имеем

$$p_j^{(n)}(A) = \sum_{k \in A} p_k^{(n-1)}(A) p_{kj}, \quad n \geq 1, \quad (16)$$

$$p_j^{(0)}(A) = p_j^{(0)}. \quad (17)$$

При расчете для данного n можно промежуточные $p_k^{(m)}$, $m < n$, вычислять только для $k \in A$; остальные в формулу (16) не входят.

Другой способ решения той же задачи — в рассмотрении цепи Маркова (η_n) с вероятностями перехода p_{ij}^* , равными p_{ij} при $i \in A$, 1 при $i = j \notin A$ и 0 при $i \notin A$, $j \neq i$. Попав в какое-либо состояние вне множества A , цепь Маркова (η_n) навсегда в нем остается¹. При $n \geq 1$ имеем

$$p_j^{(n)}(A) = P(\eta_n = j), \quad (18)$$

где правая часть (18) вычислена при том же начальном распределении, что и левая часть.

Полезна также следующая интерпретация. Рассмотрим цепь Маркова (ζ_n) , принудительно удерживаемую в множестве состояний A . Ее вероятности перехода

$$f_{ij} = p_{ij} / \sum_{k \in A} p_{ik}, \quad j \in A, \quad (19)$$

начальное распределение

$$p_j^{(0)} / \sum_k p_k^{(0)}, \quad j \in A. \quad (20)$$

Если $\zeta_m = i$, то вычисляется вероятность θ_m , с которой ζ_{m+1} удержалось бы в множестве A в течение одного шага при отсутствии принуждения. Имеем

$$\theta_m = \sum_{k \in A} p_{ik}.$$

В таком случае при $j \in A$

$$p_j^{(n)}(A) = M I_{nj} \theta_0 \dots \theta_{n-1}, \quad (21)$$

где I_{nj} — индикатор события $\{\zeta_n = j\}$.

Пример. При обучении оператора управлению каким-либо объектом находящийся рядом специалист вовремя исключает опасные ситуации, осуществляя принуждение; в конце сеанса обучения по степени самостоятельности действий учащегося (это и есть произведение $\theta_0 \dots \theta_{n-1}$) можно судить о надежности его подготовки.

§ 3. Финальные вероятности

Для цепей Маркова типичен случай, когда

$$p_{ij}^{(n)} \xrightarrow{n \rightarrow \infty} \pi_i \quad \text{либо} \quad p_j^{(n)} \xrightarrow{n \rightarrow \infty} \pi_j.$$

¹ Можно также «склеить» все состояния, не входящие в A , в одно, заменяя индивидуальные p_{ij} , $j \notin A$, их суммой.

В обоих случаях назовем π_j *финальной вероятностью* состояния j . Финальная вероятность интерпретируется как вероятность того, что система находится в состоянии j при условии, что она функционирует неограниченно долго («начиная с момента $t = -\infty$ »). Если $\sum_j \pi_j = 1$, то (π_j) называется *предельным распределением* цепи Маркова (ξ_n) . Видим, что π_j зависят от начального состояния $\xi_0 = i$ либо начального распределения $p^{(0)} = (p_i^{(0)})$. Если для любого $p^{(0)}$ существует одно и то же предельное распределение, т. е.

$$\forall p^{(0)} \lim_{n \rightarrow \infty} p_j^{(n)} = \pi_j, \quad \sum_j \pi_j = 1, \quad (1)$$

то (π_j) называется *эргодическим распределением* цепи Маркова (ξ_n) , а сама (ξ_n) называется *эргодической*.

Определение эргодичности цепи Маркова и эргодического распределения (π_j) можно сформулировать еще и так:

$$\forall i \lim_{n \rightarrow \infty} p_{ij}^{(n)} \rightarrow \pi_j, \quad \sum_j \pi_j = 1. \quad (2)$$

Оба определения эквивалентны. В самом деле, (2) — частный случай (1), так как $p_{ij}^{(n)} = p_j^{(n)}$ при $p_k^{(0)}$, равном 1 при $k = i$ и равном 0 для всех остальных k . Далее, так как $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$, то, взяв такое N , чтобы $\sum_{i=0}^N p_i^{(0)} > 1 - \varepsilon$ и соответственно $\sum_{i=N+1}^{\infty} p_i^{(0)} < \varepsilon$, при условии (2) получим

$$p_j^{(n)} - \pi_j = \sum_{i=0}^N p_i^{(0)} (p_{ij}^{(n)} - \pi_j) + \sum_{i=N+1}^{\infty} p_i^{(0)} (p_{ij}^{(n)} - \pi_j).$$

Первая сумма стремится к нулю при $n \rightarrow \infty$, так как выражение в скобках бесконечно мало вследствие (2). Вторая сумма меньше ε , так как $|p_{ij}^{(n)} - \pi_j| \leq 1$ и $\sum_{j=N+1}^{\infty} p_j^{(0)} < \varepsilon$. Итак, выполняется равенство (1).

Пример 1. Пусть (ξ_n) — цепь Маркова с двумя состояниями 0, 1 и матрицей перехода $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Тогда $\xi_n = \xi_0$ с вероятностью 1, т. е.

$$p_{00}^{(n)} = 1, \quad p_{01}^{(n)} = 0, \quad p_{11}^{(n)} = 1, \quad p_{10}^{(n)} = 0.$$

Видим, что финальные вероятности существуют, но зависят от начального состояния. Вывод: цепь Маркова — не эргодическая.

Пример 2. Пусть $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Тогда ξ_n попеременно равно 1 и 0. Например, при $\xi_0 = 0$ имеем $\xi_n = (1 - (-1)^n)/2$, откуда $p_{00}^{(n)} = (1 + (-1)^n)/2$. Финальные вероятности не существуют,

Пример 3. Пусть $P = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$, где $0 < p < 1$, $q = 1 - p$. Как было выяснено в примере 1 § 2.2,

$$p_{ij}^{(n)} = \frac{1}{2} (1 \pm \Delta^n), \quad |\Delta| < 1.$$

Следовательно, $p_{ij}^{(n)} \xrightarrow{n \rightarrow \infty} 1/2$. Итак, цепь Маркова эргодическая; $(1/2, 1/2)$ — ее эргодическое распределение.

Распределение (p_j) называется *стационарным*, если выполняется система уравнений

$$p_i = \sum_j p_i p_{ij}, \quad (3)$$

$$\sum_i p_i = 1. \quad (4)$$

Взяв $p_j^{(n-1)} = p_j$, из (2.2) получим

$$p_j^{(n)} = \sum_i p_i^{(n-1)} p_{ij} = \sum_i p_i p_{ij} = p_j.$$

По индукции заключаем, что $p_j^{(m)} = p_j$ для всех $m \geq n$. Это означает, что если на некотором шаге распределение (ξ_n) стационарно, то и на всех последующих шагах будет то же распределение, т. е. стационарность распределения означает его независимость от времени.

Если (π_i) — предельное распределение цепи Маркова, то оно стационарно. Докажем это свойство. Выберем такое N , что

$\sum_{i=0}^N \pi_i > 1 - \varepsilon$ и соответственно $\sum_{i=N+1}^{\infty} \pi_i < \varepsilon$. Поскольку $p_j^{(n)} \rightarrow \pi_j$,

то $\sum_{j=0}^N p_j^{(n)} > 1 - \varepsilon$ при $n > n_0$, т. е.

$$\sum_{j=N+1}^{\infty} p_j^{(n)} < \varepsilon, \quad n > n_0. \quad (5)$$

На основании (2.11)

$$p_j^{(n)} = \sum_{i=0}^N p_i^{(n-1)} p_{ij} + \sum_{i=N+1}^{\infty} p_i^{(n-1)} p_{ij}. \quad (6)$$

Вследствие (5) вторая сумма в (6) меньше ε . Перейдя к пределу при $n \rightarrow \infty$, получим

$$\left| \pi_j - \sum_{i=0}^N \pi_i p_{ij} \right| < \varepsilon.$$

Перейдя к пределу при $N \rightarrow \infty$, будем иметь неравенство

$$\left| \pi_j - \sum_{i=0}^{\infty} \pi_i p_{ij} \right| \leq \varepsilon,$$

а так как $\varepsilon > 0$ произвольно, то

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \quad (7)$$

что и требовалось доказать. Равенство

$$\sum_i \pi_i = 1 \quad (8)$$

вытекает из того, что (π_i) — распределение.

Пусть теперь (p_i) — стационарное распределение. Положив $p_i^{(0)} = p_i$, получим, что $p_i^{(n)} = p_i$, а следовательно, $p_i^{(n)} = \pi_i$ при данном $p_i^{(0)}$. Таким образом, стационарное распределение является предельным распределением цепи Маркова, если начальное распределение совпадает с ним.

Итак, между стационарными и предельными распределениями установлено двустороннее соответствие.

Предположим, что цепь Маркова (ξ_n) является эргодической, (π_i) — ее эргодическое распределение. Тогда (π_i) — единственное стационарное распределение: в противном случае существовало бы другое стационарное (оно же предельное) распределение вопреки условию (1).

Итак, для эргодической цепи Маркова система уравнений (7), (8) (или в других обозначениях (3), (4)) имеет единственное неотрицательное решение $(p_i \geq 0)$, представляющее собой эргодическое распределение (ξ_n) .

Если имеется предельное распределение (π_i) или стационарное распределение (p_i) , то числа π_i , p_i называются *предельными* (соответственно *стационарными*) *вероятностями*.

Пример. Пусть (ξ_n) — цепь Маркова с состояниями 1, 2, 3 и матрицей перехода

$$\begin{pmatrix} 0 & 0,5 & 0,5 \\ 0,7 & 0,15 & 0,15 \\ 0,5 & 0 & 0,5 \end{pmatrix}.$$

Имеем следующую систему уравнений для стационарных вероятностей:

$$\begin{aligned} p_1 &= 0,7p_2 + 0,5p_3, \\ p_2 &= 0,5p_1 + 0,15p_3, \\ p_3 &= 0,5p_1 + 0,15p_2 + 0,5p_3, \\ p_1 + p_2 + p_3 &= 1. \end{aligned}$$

Из второго уравнения $p_1 = 1,7 p_2$ и затем из третьего $p_3 = 2p_2$. Подстановка выражений p_1 и p_3 в четвертое уравнение приводит к равенству $4,7 p_2 = 1$, $p_2 \approx 0,213$. Затем определяются $p_1 \approx 0,362$, $p_3 \approx 0,425$. Итак, эргодическое распределение приблизительно равно вектору $(0,362; 0,213; 0,425)$. Заметим, что в процессе решения первое уравнение вообще не использовано. В общем случае из уравнений системы (3) одно — следствие остальных. В самом деле, суммирование обеих частей (3) по j приводит к тождеству $\sum_i p_i = \sum_i p_i$, так как $\sum_i p_{ij} = 1$.

Смысл эргодического распределения раскрывается в следующем утверждении, называемом *элементарной эргодической теоремой*.

Теорема. Пусть $T_j(n)$ — число тех k , $0 \leq k \leq n-1$, для которых $\xi_k = j$. Тогда для любого $\varepsilon > 0$

$$P(|n^{-1}T_j(n) - \pi_j| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0. \quad (9)$$

Таким образом, эргодическая цепь Маркова из n последовательных шагов находится в состоянии j примерно $n\pi_j$ шагов.

Д о к а з а т е л ь с т в о элементарной эргодической теоремы дадим в предположении, что число состояний цепи Маркова конечно. Пусть $\tau_i^{(1)}, \tau_i^{(2)}, \dots$ — расположенные в порядке возрастания номера шагов, на которых $v_n = i$. Тогда последовательность состояний на следующих за ними шагах есть последовательность независимых испытаний, в которых исход j имеет вероятность p_{ij} . Обозначим через $S_N(j|i)$ число исходов j среди N таких испытаний. Если $T_{ij}(n)$ — число переходов из состояния i в состояние j , т. е. число таких k , $0 \leq k \leq n-2$, что $\xi_k = i$, $\xi_{k+1} = j$, то

$$T_{ij}(n) = T_i(n) S_{T_i(n)}(j|i) + \theta_{ij}, \quad (10)$$

где θ_{ij} равно либо 0, либо -1 . Формулу (10) можно истолковать так. Из первых n шагов (включая нулевой) будет $N = T_i(n)$ таких, на которых $\xi_k = i$. Тогда число шагов $k+1$, на которых $\xi_{k+1} = j$, равно $S_N(j|i) = S_{T_i(n)}(j|i)$. Число $T_{ij}(n)$ может быть равно этому числу, а может быть меньше на единицу, так как в случае $\xi_{n-1} = i$ последний переход из i в j не учитывается. Разобьем отрезок $\{0, 1, \dots, n-1\}$ числами $0 = t_0 < t_1 < \dots < t_m = n-1$ на отрезки $\{0, 1, \dots, t_1\}, \{t_1 + 1, \dots, t_2\}, \dots, \{t_{m-1} + 1, \dots, t_m\}$, причем так, чтобы $n\epsilon - 1 \leq t_1 < n\delta, n\delta - 1 \leq t_2 - t_1 < n\delta, \dots, n\delta - 1 \leq t_m - t_{m-1} < n\delta$. (При достаточно большом n это всегда возможно). Обозначим через B_{ij} событие, состоящее в том, что

$$|T_{ij}(n) - T_i(n) p_{ij}| \geq 2n\delta + 1. \quad (11)$$

Возможен один из m случаев: либо $t_0 \leq T_i(n) \leq t_1$, либо $t_1 < T_i(n) \leq t_2$, ..., либо $t_{m-1} < T_i(n) \leq t_m$. В r -м из этих случаев

$$S_{t_{r-1}}(j|i) \leq T_{ij}(n) \leq S_{t_r}(j|i). \quad (12)$$

Из (11) получаем

$$\begin{aligned} (S_{t_{r-1}}(j|i) - t_{r-1}p_{ij}) - (T_i - t_{r-1}) p_{ij} &\leq T_{ij} - T_i p_{ij} \leq \\ &\leq (S_{t_r}(j|i) - t_r p_{ij}) + (t_r - T_i) p_{ij} + 1, \end{aligned}$$

где для упрощения символики записано $T_i = T_i(n)$, $T_{ij} = T_{ij}(n)$. Поскольку по условию $t_{r-1} < T_i \leq t_r$ и $t_r - t_{r-1} < n\delta$, то из последнего неравенства вытекает следующее:

$$\begin{aligned} (S_{t_{r-1}}(j|i) - t_{r-1}p_{ij}) - n\delta &< T_{ij} - T_i p_{ij} < \\ &< (S_{t_r}(j|i) - t_r p_{ij}) + n\delta + 1. \end{aligned} \quad (13)$$

Значит, чтобы произошло событие B_{ij} , необходимо, чтобы для некоторого r

$$|S_{t_r}(j|i) - t_r p_{ij}| \geq n\delta. \quad (14)$$

При $r = 0$ неравенство (14) при достаточно большом n невозможно, так как $t_0 = 0$, $S_0(j|i)$ равно 0 или 1. При $r \geq 1$ вероятность этого неравенства стремится к нулю в силу закона больших чисел в

форме Бернулли. Итак,

$$P(B_{ij}) \xrightarrow{n \rightarrow \infty} 0.$$

Поскольку пар (i, j) конечное число (только в этом пункте данный факт используется), то

$$P\left(\bigcup_{i,j} B_{ij}\right) \leq \sum_{i,j} P(B_{ij}) \xrightarrow{n \rightarrow \infty} 0, \quad (15)$$

т. е. вероятность противоположного события

$$A = \{|T_{ij} - T_i p_{ij}| \leq 2n\delta\} \quad (16)$$

стремится к 1 при $n \rightarrow \infty$.

При условии A , так как $T_j = \sum T_{ij} + \theta_j$, где $\theta_j = 0, 1$ (единица за счет возможного случая $\xi_0 = j$), то из (16) суммированием по i и делением на n находим

$$\left| \frac{T_j}{n} - \sum_i \frac{T_{ij}}{n} p_{ij} \right| \leq 2\delta + \frac{1}{n}. \quad (17)$$

Таким образом, система уравнений (7) выполняется при замене π_j на T_j/n с малой погрешностью. Уравнение (8) выполняется точно, так как $\sum_j T_j = n$.

Примем без доказательства, что малым значениям погрешностей соответствуют малые отклонения решения системы¹. Тогда $|T_j/n - \pi_j| < \varepsilon$ при достаточно малом δ и достаточно большом n . Поскольку, по доказанному, вероятность такого события стремится к 1, то тем самым элементарная эргодическая теорема доказана.

Пример. Пусть эргодическая цепь Маркова (ξ_n) описывает поведение технической системы. Множество состояний ξ_n разобьем на два множества A и B , считая $i \in A$ состояниями работоспособности, $i \in B$ состояниями отказа. Найдем эргодические вероятности π_i и определим

$$P(A) = \sum_{i \in A} \pi_i, \quad P(B) = \sum_{i \in B} \pi_i.$$

Тогда из n последовательных шагов система находится в исправном состоянии в течение примерно $nP(A)$ шагов, т. е. долю времени, приблизительно равную $P(A)$. Поэтому число $P(A)$ называют *коэффициентом готовности системы*.

§ 4. Предельное поведение вероятностей

Пусть ξ_n — цепь Маркова с конечным множеством состояний и матрицей перехода P . Введем *производящую функцию*

$$\varphi(z) = \sum_{n=0}^{\infty} p^{(n)} z^n, \quad |z| < 1. \quad (1)$$

¹ Докажем это свойство в случае, когда все $p_{ij} > 0$, а следовательно, $p_{ij} \geq \rho > 0$. Из (7) получаем $\pi_j \geq \rho \sum_i \pi_i = \rho$. Известно, что неотрицательное решение системы (7) — (8) единственно. Если (x_j) — любое другое решение, то при достаточно малом $a > 0$ решение $((1-a)\pi_j + ax_j)$ неотрицательно. Из этого следует, что такого (x_j) не может быть. Тогда заключаем, что решение (π_j) единственно, т. е. определитель системы отличен от нуля, а следовательно, упомянутые отклонения бесконечно малы.

В скалярном виде $\varphi(z) = (\varphi_i(z))$, где

$$\varphi_j(z) = \sum_{n=0}^{\infty} p_j^{(n)} z^n. \quad (2)$$

Поскольку $p^{(n)} = p^{(0)} P^n$, то

$$\varphi(z) = \sum_{n=0}^{\infty} p^{(0)} P^n z^n = p^{(0)} + p^{(0)} z P \sum_{n=0}^{\infty} P^n z^n = p^{(0)} + p^{(0)} P \varphi(z).$$

Введя единичную матрицу I , так что $\varphi(z) I = \varphi(z)$, получим

$$\varphi(z) (I - zP) = p^{(0)}. \quad (3)$$

Определитель матрицы $I - zP$ при $z \rightarrow 0$ стремится к 1, а следовательно, при достаточно малом z она невырождена, т. е. обладает обратной матрицей $(I - zP)^{-1}$. Умножив на последнюю матрицу левую и правую части (3), получим

$$\varphi(z) = p^{(0)} (I - zP)^{-1}. \quad (4)$$

Поскольку $|I - zP|$ — полином, то существует r его корней, где r — число состояний цепи Маркова. Эти корни обозначим $1/\rho_0, \dots, 1/\rho_{r-1}$. Их можно определить из уравнения

$$|P - \rho I| = 0. \quad (5)$$

Наиболее типичный случай — все корни различны.

Будем искать частные решения уравнения (4) в виде $b_k/(1 - \rho_k z)$, $0 \leq k \leq r-1$, где b_k — векторы. Это эквивалентно тому, что частное решение для $p^{(n)}$ есть $b_k \rho_k^n$, $n \geq 0$. Для определения b_k имеем матричное уравнение

$$b_k (P - I \rho) = 0. \quad (6)$$

Матрица этого уравнения имеет ранг, на единицу меньший ее размерности, т. е. уравнение имеет единственное, с точностью до постоянного множителя, решение b_k . Различные b_k линейно независимы. Общее решение имеет вид

$$p^{(n)} = \sum_{k=0}^{r-1} c_k b_k \rho_k^n, \quad (7)$$

где c_k — числовые коэффициенты. Для их определения достаточно решить матричное уравнение

$$\sum c_k b_k = p^{(0)}, \quad (8)$$

т. е. разложить $p^{(0)}$ по базису (b_k) . Одно из чисел ρ_k (обозначим его ρ_0) равно 1; для него

$$b_0 (P - I) = 0,$$

т. е. это уравнение для эргодического распределения; следовательно, можно положить $b_0 = \pi = (\pi_k)$.

Из теории стохастических матриц известно, что все ρ_k , $k \geq 1$, удовлетворяют неравенству

$$|\rho_k| < 1.$$

Поскольку $\rho^{(n)} \xrightarrow{n \rightarrow \infty} \pi$ и, кроме того, $\rho^{(n)} \rightarrow c_0 \pi$, то $c_0 = 1$, т. е.

$$\rho^{(n)} = \pi + \sum_{k=1}^{r-1} c_k b_k \rho_k^n. \quad (9)$$

Если среди чисел $\rho_1, \dots, \rho_{r-1}$ наибольшее по модулю ρ_1 , а все остальные меньше, то

$$\rho^{(n)} = \pi + c_1 b_1 \rho_1^n + O(\rho_2^n), \quad (10)$$

где предполагается, что ρ_2 — наибольшее по модулю из чисел ρ_2, \dots, ρ_r . Эта формула показывает, что скорость приближения $\rho^{(n)}$ к эргодическому распределению экспоненциальная, и дает асимптотически точную оценку погрешности $\rho^{(n)} - \pi$.

Пример. Пусть $(\xi_n, n \geq 0)$ — цепь Маркова с состояниями 1, 2, 3 и матрицей перехода

$$P = \begin{pmatrix} 0,1 & 0,7 & 0,2 \\ 0,4 & 0,4 & 0,2 \\ 0,4 & 0 & 0,6 \end{pmatrix}.$$

Для определения ρ_1, ρ_2, ρ_3 имеем уравнение (5), которое в данном случае принимает вид

$$\begin{vmatrix} 0,1 - \rho & 0,7 & 0,2 \\ 0,4 & 0,4 - \rho & 0,2 \\ 0,4 & 0 & 0,6 - \rho \end{vmatrix} = 0.$$

После раскрытия определителя получаем кубическое уравнение $\rho^3 - 1,1\rho^2 - 0,02\rho + 0,12 = 0$. Один корень известен: $\rho_0 = 1$. Разделив кубический многочлен на $\rho - 1$, получим квадратное уравнение $\rho^2 - 0,1\rho - 0,12 = 0$. Его корни $\rho_1 = 0,4$, $\rho_2 = -0,3$.

При $\rho = \rho_1 = 1$ уравнение (6) имеет вид

$$(x, y, z)(P - I) = 0,$$

где x, y, z — компоненты вектора b_0 . Подставив конкретное значение матрицы P , находим

$$P - I = \begin{pmatrix} 0,1 - 1 & 0,7 & 0,2 \\ 0,4 & 0,4 - 1 & 0,2 \\ 0,4 & 0 & 0,6 - 1 \end{pmatrix} = \begin{pmatrix} -0,9 & 0,7 & 0,2 \\ 0,4 & -0,6 & 0,2 \\ 0,4 & 0 & -0,4 \end{pmatrix},$$

откуда получаем систему уравнений

$$\begin{aligned} -0,9x + 0,4y + 0,4z &= 0, \\ 0,7x - 0,6y &= 0, \\ 0,2x + 0,2y - 0,4z &= 0. \end{aligned}$$

Как выше было отмечено для общего случая, эти уравнения зависимы, так что возьмем из них только первые два, но, чтобы (x, y, z) было распределением, вместо третьего возьмем уравнение $x + y + z = 1$. Таким образом, имеем систему

$$\begin{aligned} -0,9x + 0,4y + 0,4z &= 0, \\ 0,7x - 0,6y &= 0, \\ x + y + z &= 1. \end{aligned}$$

Из второго уравнения получаем $y = 1,166\ 6667x$, затем из третьего $z = 1 - 2,166\ 6667x$. Подстановкой этих выражений в первое уравнение находим $x = 0,307\ 6923$; далее имеем $y = 0,358\ 9743$; $z = 0,333\ 3334$.

Таким образом,

$$b_0 = \pi = (0,307\ 6923; 0,358\ 9743; 0,333\ 3334).$$

Переходим к корню $\rho_1 = 0,4$. В этом случае

$$P - \rho_1 I = \begin{pmatrix} 0,1 - 0,4 & 0,7 & 0,2 \\ 0,4 & 0,4 - 0,4 & 0,2 \\ 0,4 & 0 & 0,6 - 0,4 \end{pmatrix} = \begin{pmatrix} -0,3 & 0,7 & 0,2 \\ 0,4 & 0 & 0,2 \\ 0,4 & 0 & 0,2 \end{pmatrix},$$

и мы получаем систему уравнений

$$\begin{aligned} -0,3x + 0,4y + 0,4z &= 0, \\ 0,7x &= 0. \end{aligned}$$

(Причина, по которой взято лишь два уравнения, объяснена выше; условие $x + y + z = 1$ здесь не имеет смысла, так как b_k при $k \neq 0$ есть не обязательно распределение вероятностей.) Одну из переменных можно задать произвольно. Возьмем, например, $y = 1$. Тогда получим $x = 0$, $z = -1$, $b_2 = (0, 1, -1)$.

Для корня $\rho_2 = -0,3$ находим

$$P - \rho_2 I = \begin{pmatrix} 0,1 + 0,3 & 0,7 & 0,2 \\ 0,4 & 0,4 + 0,3 & 0,2 \\ 0,4 & 0 & 0,6 + 0,3 \end{pmatrix} = \begin{pmatrix} 0,4 & 0,7 & 0,2 \\ 0,4 & 0,7 & 0,2 \\ 0,4 & 0 & 0,9 \end{pmatrix}.$$

Для $b_2 = (x, y, z)$ имеем систему уравнений

$$\begin{aligned} 0,4x + 0,4y + 0,4z &= 0, \\ 0,7x + 0,7y &= 0. \end{aligned}$$

Отсюда $z = 0$. Можно положить $x = 1$, $y = -1$. Имеем $b_2 = (1, -1, 0)$.

Пусть задано начальное распределение $p^{(0)} = (0,15; 0,6; 0,25)$. Согласно (8), решаем систему уравнений

$$b_0 + c_1 b_1 + c_2 b_2 = p^{(0)}$$

(коэффициент при b_0 равен 1). Имеем

$$(0,307\ 6923; 0,358\ 9743; 0,333\ 3334) + c_1 (0, 1, -1) + c_2 (1, -1, 0) = (0,15; 0,6; 0,25),$$

или в скалярном виде

$$\begin{aligned} 0,307 \dots + 0 \cdot c_1 + 1 \cdot c_2 &= 0,15, \\ 0,358 \dots + 1 \cdot c_1 - 1 \cdot c_2 &= 0,6, \\ 0,333 \dots - 1 \cdot c_1 + 0 \cdot c_2 &= 0,25. \end{aligned}$$

Третье и первое из этих уравнений дают значения $c_1 = 0,083\ 3334$; $c_2 = -0,157\ 6923$. Используя второе уравнение для проверки, убеждаемся, что оно удовлетворяется. Результат запишем в скалярном виде:

$$\begin{aligned} p_1^{(n)} &= 0,307\ 6923 - 0,157\ 6923 (-0,3)^n, \\ p_2^{(n)} &= 0,358\ 9743 + 0,083\ 3334 \cdot 0,4^n + 0,157\ 6923 (-0,3)^n, \\ p_3^{(n)} &= 0,333\ 3334 - 0,083\ 3334 \cdot 0,4^n. \end{aligned}$$

Отклонение $p^{(n)}$ от эргодического распределения оценивается на основании *неравенства Бернштейна*: если существует такое

состояние i_0 , что $p_{ii_0} \geq \varepsilon > 0$ для всех i , то

$$\sum_j |p_j^{(n)} - q_j^{(n)}| \leq 2(1 - \varepsilon)^n, \quad (11)$$

где $p_j^{(n)}, q_j^{(n)}$ — распределения ξ_n при различных начальных распределениях $p^{(0)}, q^{(0)}$.

Для доказательства неравенства Бернштейна построим бернуллиеву последовательность $(\theta_n, n \geq 1)$, где $P(\theta_n = 1) = \varepsilon$, $P(\theta_n = 0) = 1 - \varepsilon$ и в случае $\theta_n = 1$ полагаем $\xi_n = i_0$. Пусть $\xi_{n-1} = i$. Если $\theta_n = 0$, событие $\{\xi_n = i_0\}$ наступает с вероятностью $(p_{ii_0} - \varepsilon)/(1 - \varepsilon)$; при $j \neq i_0$ событие $\{\xi_n = j\}$ наступает с вероятностью $p_{ij}/(1 - \varepsilon)$. Легко видеть, что определенная таким образом последовательность (ξ_n) есть цепь Маркова с матрицей перехода (p_{ij}) .

Если $\theta_k = 1$ хотя бы для одного k , $1 \leq k \leq n$, то ξ_n имеет распределение, не зависящее от начального распределения. Поэтому

$$p_j^{(n)} = (1 - \varepsilon)^n a_j^{(n)} + (1 - (1 - \varepsilon)^n) c_j^{(n)}, \quad (12)$$

$$q_j^{(n)} = (1 - \varepsilon)^n b_j^{(n)} + (1 - (1 - \varepsilon)^n) c_j^{(n)}, \quad (13)$$

где $(1 - \varepsilon)^n$ — вероятность события $\{\theta_k = 0, 1 \leq k \leq n\}$; $c_j^{(n)}$ — вероятность события $\{\xi_n = j\}$ при условии, что $\theta_k = 1$ хотя бы для одного k , $1 \leq k \leq n$; $a_j^{(n)}$ и $b_j^{(n)}$ — вероятности события $\{\xi_n = j\}$ при условии $\{\theta_k = 0, 1 \leq k \leq n\}$, соответствующие начальным вероятностям $p_j^{(0)}, q_j^{(0)}$. Из равенств (12), (13) находим

$$\sum_j |p_j^{(n)} - q_j^{(n)}| \leq (1 - \varepsilon)^n \left| \sum_j a_j^{(n)} + \sum_j b_j^{(n)} \right| = 2(1 - \varepsilon)^n,$$

что и требовалось доказать.

Введем события $A_0 = \{\xi_k = 0, 1 \leq k \leq n\}$, $A_i = \{\xi_i = 1, \xi_k = 0, i < k \leq n\}$, $1 \leq i \leq n$. Тогда A_i — несовместные события, причем $A_0 \cup A_1 \cup \dots \cup A_n = \Omega$. По формуле полной вероятности

$$p_j^{(n)} = \sum_{i=0}^n P(A_i) P(\xi_n = j | A_i).$$

Легко видеть, что при $i \geq 1$ вероятность $P(\xi_n = j | A_i)$ зависит от n, i только через их разность и поэтому данную вероятность можно обозначить $\psi_j(n - i)$ при $i \geq 1$, $\psi_j(n)$ при $i = 0$.

Заметив, что $P(A_0) = (1 - \varepsilon)^n$, $P(A_i) = \varepsilon(1 - \varepsilon)^{n-i}$, получим

$$\begin{aligned} p_j^{(n)} &= (1 - \varepsilon)^n \psi_j(n) + \varepsilon \sum_{i=1}^n (1 - \varepsilon)^{n-i} \psi_j(n - i) = \\ &= (1 - \varepsilon)^n \psi_j(n) + \varepsilon \sum_{k=0}^{n-1} (1 - \varepsilon)^k \psi_j(k). \end{aligned} \quad (14)$$

Поскольку $\psi_j(k)$ ограничены, то существует предел

$$\pi_j = \lim_{n \rightarrow \infty} p_j^{(n)} = \varepsilon \sum_{k=0}^{\infty} (1 - \varepsilon)^k \psi_j(k). \quad (15)$$

Но так как $\sum_{j=0}^N p_j^{(n)} \leq 1$, то и $\sum_{j=0}^N \pi_j \leq 1$, откуда $\sum_{j=0}^{\infty} \pi_j \leq 1$.

В то же время $\sum_{j=0}^{\infty} \psi_j(k) = 1$ для любого k , а следовательно, найдется такое N , что $\sum_{j=0}^{N-1} \psi_j(k) > 1 - \delta$, $0 \leq k \leq K$. На основании этого из (15) имеем

$$\sum_{j=0}^N \pi_j \geq \varepsilon \sum_{k=0}^{N-1} (1 - \varepsilon)^k (1 - \delta) > 1 - \delta - (1 - \varepsilon)^N.$$

В силу произвольности $\delta > 0$ и N получаем, что $\sum_{j=0}^{\infty} \pi_j = 1$. Переходя в неравенстве $p_{ij}^{(n)} \geq \sum_{n=0}^{N-1} p_i^{(n-1)} p_{ij}$ к пределу при $n \rightarrow \infty$, получим $\pi_j \geq \sum_{i=0}^{N-1} \pi_i p_{ij}$. Теперь можно перейти к пределу при $N \rightarrow \infty$. Имеем $\pi_j \geq \sum_{i=0}^{\infty} \pi_i p_{ij}$. Следовательно, $\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij} + \varepsilon_j$, где $\varepsilon_j \geq 0$. Просуммировав последнее равенство по j , получим $1 = 1 + \sum_j \varepsilon_j$, откуда $\varepsilon_j = 0$. Итак, (π_j) — стационарное распределение. Взяв $q^{(0)} = (\pi_j)$, получим $q^{(n)} = (\pi_j)$ при любом $n \geq 0$. На основании (11) получаем неравенство

$$\sum_j |p_j^{(n)} - \pi_j| \leq 2(1 - \varepsilon)^n, \quad (16)$$

так же называемое *неравенством Бернштейна*.

Пример. Имеется цепь Маркова с состояниями 1, 2, 3, 4 и матрицей перехода

$$P = \begin{pmatrix} 0,2 & 0,5 & 0,25 & 0,25 \\ 0,25 & 0,2 & 0,3 & 0,25 \\ 0,25 & 0,25 & 0,2 & 0,3 \\ 0,3 & 0,25 & 0,25 & 0,2 \end{pmatrix}.$$

Найти эргодическое распределение $\pi = (\pi_j)$ и оценить отклонение $p^{(n)}$ от π .

Уравнения для эргодического распределения имеют вид

$$\pi_j = \sum_{i=1}^4 \pi_i p_{ij}, \quad 1 \leq j \leq 4; \quad \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1.$$

Поскольку в данном случае сумма элементов каждого столбца матрицы P равна 1, то данным уравнениям удовлетворяет решение $\pi_j = 1/4$, $1 \leq j \leq 4$. Взяв $i_0 = 1$ (можно также взять равное 2, 3 или 4), найдем $p_{ii_0} \geq 0,2$. Итак, при любом начальном распределении

$$|p_1^{(n)} - 0,25| + |p_2^{(n)} - 0,25| + |p_3^{(n)} - 0,25| + |p_4^{(n)} - 0,25| \leq 2 \cdot 0,8^n.$$

Бывают случаи, когда во всех столбцах матрицы P имеются нули. В таком случае, чтобы применить неравенство Бернштейна,

рассматриваем цепь Маркова $(\eta_n = \xi_{mn})$, где m такое, что матрица P^m уже не содержит нулей. Тогда поскольку (η_n) — цепь Маркова с матрицей перехода P^m , то для нее применимо неравенство Бернштейна. Подробности покажем на примере цепи Маркова (ξ_n) , для которой

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Непосредственным вычислением получаем

$$P^3 = \begin{pmatrix} 1/8 & 3/8 & 3/8 & 1/8 \\ 1/8 & 1/8 & 3/8 & 3/8 \\ 3/8 & 1/8 & 1/8 & 3/8 \\ 3/8 & 3/8 & 1/8 & 1/8 \end{pmatrix}.$$

Как и в предыдущем примере, убеждаемся, что $\pi_j = (1/4, 1/4, 1/4, 1/4)$. Элементы матрицы P^3 больше или равны $1/8$. Отсюда

$$\sum_{j=1}^4 |p_j^{(3n)} - q_j^{(3n)}| \leq 2 (7/8)^n.$$

Если вместо $3n$ взять $3n + r$, где r равно 1 или 2, то заметим, что $p_j^{(3n+r)}, q_j^{(3n+r)}$ можно рассматривать как распределения на $3n$ -м шаге при начальных распределениях $p^{(r)}, q^{(r)}$. Итак,

$$\sum_{j=1}^4 |p_j^{(3n+r)} - q_j^{(3n+r)}| \leq 2 (7/8)^n, \quad r = 0, 1, 2.$$

В частности,

$$\sum_{j=1}^4 |p_j^{(3n+r)} - 1/4| \leq 2 (7/8)^n, \quad r = 0, 1, 2.$$

Пусть $(\xi_n, n \geq 0)$ — цепь Маркова с конечным или счетным множеством состояний X . Состояние j называется *достижимым из состояния i* , если $p_{ij}^{(n)} > 0$ хотя бы для одного $n > 0$. Достижимо ли одно состояние из другого, можно определить с помощью *графа переходов* цепи Маркова (рис. 22). Вершины графа изображают состояния. Если $p_{ij} > 0$, то от вершины i к вершине j проводится направленная дуга; направление указывается стрелкой. Тогда j достижимо из i в том и только том случае, если из вершины i можно попасть в вершину j , двигаясь в направлении стрелок. Так, на рис. 22 состояние 5 достижимо из 1 (путь $1 \rightarrow 4 \rightarrow 3 \rightarrow 5$), состояние 2 не достижимо ни из какого другого состояния. Состояния i, j называются *сообщающимися*, если каждое из них достижимо из другого. Например, на рис. 22 состояния 3 и 5 сообщающиеся.

Подмножество $Z \subset X$ называется *классом сообщающихся состояний*, если любые два состояния i, j из Z сообщаются. Класс Z сообщающихся состояний называется *замкнутым*, если из

него нет выхода, т. е. $p_{ij} = 0$ для любых $i \in Z, j \notin Z$ (тогда и $p_{ij}^{(n)} = 0$ для любого $n \geq 1$). Во множестве X выделяются замкнутые классы сообщающихся состояний X_1, X_2, \dots и множество X_0 состояний, не принадлежащих никаким замкнутым классам сообщающихся состояний. Состояния из X_0 называются *несущественными*. Можно показать, что для любого $i \in X$ и любого начального распределения $p^{(0)}$

$$p_{ij}^{(n)} \xrightarrow{n \rightarrow \infty} 0, \quad p_j^{(n)} \xrightarrow{n \rightarrow \infty} 0, \quad j \in X_0. \quad (17)$$

Таким образом, при большом n вероятностью пребывания ξ_n в несущественном состоянии j можно пренебречь. При конечном множестве состояний X

$$\sum_{j \in X_0} p_{ij}^{(n)} \rightarrow 0, \quad \sum_{j \in X_0} p_j^{(n)} \rightarrow 0. \quad (18)$$

Принцип доказательства соотношений 18 покажем, пользуясь рис. 23: имеется четыре несущественных состояния, остальные же

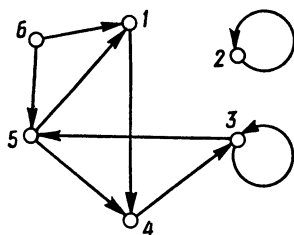


Рис. 22

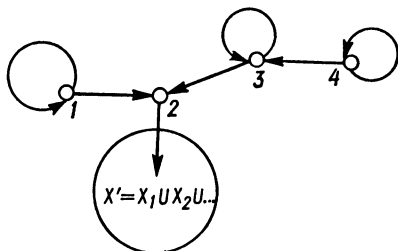


Рис. 23

принадлежат некоторым из множеств X_1, X_2, \dots . Из состояния 1 можно попасть в X' за 2 шага, из 2 — за 1 шаг, из 3 — за 2 шага, из 4 — за 3 шага. Поскольку из X' возврата нет, то

$$\sum_{j \in X'} p_{ij}^{(n)} \geq \sum_{j \in X'} p_{ij}^{(m)} \quad (19)$$

для любого $n \geq m$. Правая часть (19) положительна при (i, m) , равном $(1, 2); (2, 1); (3, 2); (4, 3)$, поэтому имеем $\sum_{j \in X'} p_{ij}^{(3)} \geq \alpha > 0$, т. е. при условии $\xi_k = i \in X_0$ событие $\{\xi_{k+3} \in X'\}$ имеет вероятность, не меньшую α . Отсюда

$$\sum_{j \in X_0} p_j^{(n)} \leq (1 - \alpha)^{[n/3]}, \quad (20)$$

где $[x]$ — целая часть числа x . Действительно, для того чтобы $\xi_n \in X_0$, необходимо, чтобы ξ_3, ξ_6, \dots принадлежали X_0 , т. е. чтобы произошло $[n/3]$ независимых событий вероятности, не большей $1 - \alpha$. Из неравенства (20) следуют соотношения (18).

Если $\xi_0 \in X_k, k \geq 1$, то $\xi_n \in X_k$ при любом $n \geq 0$. Следовательно, можно рассматривать поведение цепи Маркова в замкнутом классе сообщающихся состояний X_k независимо от всех остальных

состояний, т. е. отождествить X_k с множеством X всех состояний. Итак, рассмотрим цепь Маркова $(\xi_n, n \geq 0)$ с множеством сообщающихся состояний X . Обозначим $u_i^{(n)} = p_{ii}^{(n)}$ — вероятность возвращения в состояние i за n шагов, $f_i^{(n)}$ — вероятность первого возвращения в состояние i за n шагов: $f_i^{(n)} = P(\xi_n = i; \xi_k \neq i, 0 < k < n | \xi_0 = i)$, $f_i = \sum_{n=1}^{\infty} f_i^{(n)}$ — вероятность возвращения в состояние i .

Основное значение в изучении цепей Маркова имеет классификация состояний.

Состояние i называется *невозвратным*, если $f_i < 1$, и *возвратным*, если $f_i = 1$. В последнем случае время возвращения γ_i в состояние i есть случайная величина с законом распределения

$$P(\gamma_i = n) = f_i^{(n)}, \quad n = 1, 2, \dots \quad (21)$$

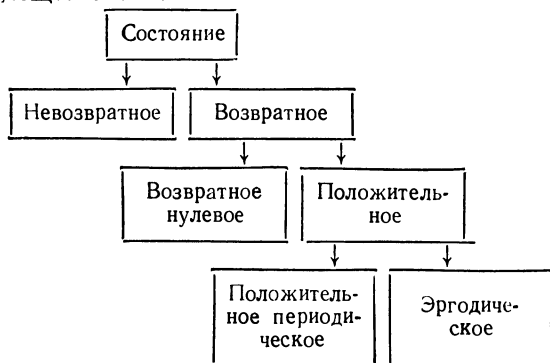
Возвратное состояние i называется *возвратным нулевым*, если

$$M\gamma_i = \sum_{n=1}^{\infty} n f_i^{(n)} = \infty, \quad (22)$$

и *положительным*, если

$$M\gamma_i = \sum_{n=1}^{\infty} n f_i^{(n)} < \infty. \quad (23)$$

Состояние i называется *периодическим* с периодом $r > 1$, если наибольший общий делитель тех n , для которых $f_i^{(n)} > 0$, равен r . Состояние называется *апериодическим*, если указанный наибольший общий делитель равен 1. Положительное апериодическое состояние называется *эргодическим*. В теории и практике наиболее важны именно эргодические состояния. Классификацию состояний покажем на следующей схеме:



Каким же образом распознать тип состояния цепи Маркова? Для этого существуют следующие критерии.

1. *Критерий возвратности*. Для возвратности состояния i необходимо и достаточно, чтобы

$$\sum_{n=1}^{\infty} u_i^{(n)} = \infty. \quad (24)$$

Таким образом, если этот ряд сходится, то состояние не возвратно, если расходится — возвратно.

Предположим, что состояние i не возвратно. Тогда вероятность k возвращений, после которых возвращений уже не будет, равна $f_i^k (1 - f_i)$. Если α — общее число возвращений, то

$$M\alpha = \sum_{k=1}^{\infty} k f_i^k (1 - f_i) = f_i / (1 - f_i). \quad (25)$$

В то же время, если I_n — индикатор возвращения в состояние на n -м шаге, то

$$\alpha = \sum_{n=1}^{\infty} I_n. \quad (26)$$

Таким образом,

$$f_i / (1 - f_i) = \sum_{n=1}^{\infty} M I_n = \sum_{n=1}^{\infty} u_i^{(n)} < \infty.$$

Пусть теперь состояние i возвратно, т. е. $f_i = 1$. Если $\gamma_i^{(1)}, \gamma_i^{(2)}, \dots$ — последовательные времена возвращения в состояние i , то для любого N событие $\{\gamma_i^{(1)} + \dots + \gamma_i^{(N)} < \infty\}$ достоверно и является объединением событий $\{\gamma_i^{(1)} + \dots + \gamma_i^{(N)} \leq n\}$. По аксиоме непрерывности

$$P(\gamma_i^{(1)} + \dots + \gamma_i^{(N)} \leq n) \xrightarrow{n \rightarrow \infty} 1. \quad (27)$$

Однако события $\{\gamma_i^{(1)} + \dots + \gamma_i^{(N)} \leq n\}$ и $\left\{ \sum_{k=1}^n I_k \geq N \right\}$ равны: сумма N времен возвращения не больше n в том и только том случае, если за n шагов произошло хотя бы N возвращений. Таким образом,

$$\begin{aligned} \sum_{k=1}^n u_i^{(k)} &= M \sum_{k=1}^n I_k \geq \sum_{j=N}^{\infty} j P\left(\sum_{k=1}^n I_k = j\right) \geq N P\left(\sum_{k=1}^n I_k > N\right) = \\ &= N P(\gamma_i^{(1)} + \dots + \gamma_i^{(N)} \leq n) \sim N \end{aligned} \quad (28)$$

вследствие (27). Поскольку N произошло, то из (28) следует, что $\sum_{k=1}^{\infty} u_i^{(k)} = \infty$. Критерий возвратности доказан.

2. *Критерий положительности.* Возвратное состояние является нулевым в том и только том случае, если

$$\lim_{n \rightarrow \infty} u_i^{(n)} = 0, \quad (29)$$

и положительным, если

$$\overline{\lim}_{n \rightarrow \infty} u_i^{(n)} > 0. \quad (30)$$

3. *Критерий эргодичности.* Возвратное состояние является эргодическим в том и только том случае, если существует

$$\lim_{n \rightarrow \infty} u_i^{(n)} > 0. \quad (31)$$

Критерии положительности и эргодичности доказаны В. Феллером методом рекуррентного события (§ 2.13). Именно, возвращение в состояние i является рекуррентным событием. Например, (31) следует из теоремы восстановления § 2.14: левая часть (31) есть величина, обратная к математическому ожиданию времени возвращения.

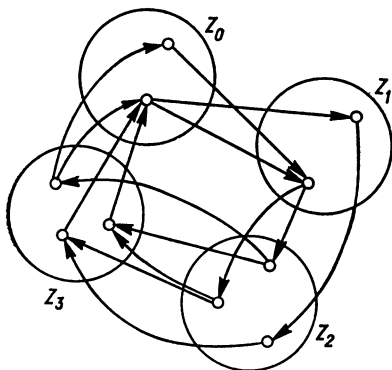


Рис. 24

Цепь Маркова, все состояния которой сообщаются, называется *неприводимой*. Оказывается, что все состояния такой цепи принадлежат одному и тому же типу: либо все они невозвратные, либо возвратные нулевые, либо эргодические, либо положительные периодические с одним и тем же периодом. Докажем некоторые из этих свойств.

Пусть i, j — различные состояния. Ввиду того что они сообщаются, найдутся такие m и k , что $p_{ii}^{(m)} > 0$, $p_{jj}^{(k)} > 0$. Следовательно,

$$p_{ii}^{(n+m+k)} \geq p_{ij}^{(m)} p_{jj}^{(n)} p_{ji}^{(k)}, \quad p_{jj}^{(n)} \geq p_{ji}^{(k)} p_{ii}^{(n-k-m)} p_{ij}^{(m)};$$

например, возвратиться в i за $n + m$ шагов можно путем перехода из i в j за m шагов, возвращения в j за n шагов и перехода из j в i за k шагов. Обозначив $p_{ij}^{(m)} p_{ji}^{(k)} = \alpha > 0$, получим

$$\alpha p_{ii}^{(n-k-m)} \leq p_{jj}^{(n)} \leq \frac{1}{\alpha} p_{ii}^{(n+k+m)},$$

или, что то же самое,

$$\alpha u_i^{(n-k-m)} \leq u_j^{(n)} \leq \frac{1}{\alpha} u_i^{(n+k+m)}. \quad (32)$$

Таким образом, если ряд $\sum_n u_i^{(n)}$ сходится, то сходится и ряд $\sum_n u_j^{(n)}$; если первый расходится, то расходится и второй. Подобным же образом, либо $u_i^{(n)} \rightarrow 0$, $u_j^{(n)} \rightarrow 0$ при $n \rightarrow \infty$, либо верхние пределы обеих последовательностей положительны. Итак, либо i, j невозвратные, либо i, j возвратные нулевые, либо i, j положительные состояния.

Цепь Маркова называется *невозвратной*, *возвратной нулевой*, *эргодической*, *положительной периодической* судя по тому, какому из этих свойств удовлетворяют ее состояния.

Неприводимая цепь Маркова называется *апериодической*, если наибольший делитель тех n , для которых $f_i^{(n)} > 0$, равен 1. Если

этот наибольший общий делитель равен $r > 1$, цепь называется *периодической* с периодом r .

Неприводимая аperiodическая цепь Маркова с конечным числом состояний является *эргодической*, т. е.

$$p_{ij}^{(n)} \xrightarrow{n \rightarrow \infty} \pi_j, \quad p_j^{(n)} \xrightarrow{n \rightarrow \infty} \pi_j, \quad (33)$$

где (π_j) — эргодическое распределение. Неприводимая периодическая цепь Маркова с периодом r обладает следующим свойством. Множество X ее состояний разбивается на классы Z_0, Z_1, \dots, Z_{r-1} так, что если $\xi_0 \in Z_m$, то $\xi_{rn} \in Z_m$ при любых $n = 1, 2, \dots$; $\xi_{rn+k} \in Z_{m+k}$, где $m+k$ заменяется на $m+k-r$ при $m+k \geq r$. Таким образом, граф переходов данной цепи Маркова характеризуется тем, что p_{ij} может быть положительным лишь в случае, если $i \in Z_0, j \in Z_1$, либо $i \in Z_1, j \in Z_2, \dots$, либо $i \in Z_{r-1}, j \in Z_0$ (рис. 24). Пусть $i \in Z_a, j \in Z_b$, где $b - a = m \pmod{r}$, т. е. $b - a - m$ делится на r . Тогда

$$p_{ij}^{(rn+m)} \xrightarrow{n \rightarrow \infty} \pi_j^{(m)}, \quad (34)$$

$$p_{ij}^{(rn+k)} = 0, \quad k - m \not\equiv 0 \pmod{r}, \quad (35)$$

где $(\pi_j^{(m)})$, $0 \leq m \leq r-1$, — распределения, однозначно определяемые системой уравнений

$$\pi_j^{(m)} = \sum_{i \in Z_{m-1}} \pi_i^{(m-1)} p_{ij}, \quad j \in Z_m, \quad 0 \leq m \leq r-1, \quad (36)$$

$$\sum_{j \in Z_0} \pi_j^{(0)} = 1 \quad (37)$$

(в равенстве (36) при $m = 0$ значение $m-1$ полагается равным r).

Пусть $p^{(0)}$ характеризуется тем, что

$$\sum_{i \in Z_k} p_j^{(0)} = \alpha_k, \quad 0 \leq k \leq r-1.$$

Тогда

$$p_j^{(n)} - \sum_{k=0}^{r-1} \alpha_k p_{kj}^{(n)} \xrightarrow{n \rightarrow \infty} 0, \quad (38)$$

где i_k — любое состояние из Z_k .

Свойства, выражаемые соотношениями (33) — (38), выполняются и для неприводимой цепи Маркова с бесконечным (счетным) множеством состояний, если эта цепь является положительной, т. е.

$\lim_{n \rightarrow \infty} u_i^{(n)} > 0$ для некоторого состояния i .

Доказательство всех этих утверждений основано на рассмотрении рекуррентного события — возвращения цепи Маркова в фиксированное состояние i . Метод доказательства совпадает с тем, который был использован при доказательстве существования эргодического распределения при условиях, достаточных для выполнения неравенства Бернштейна.

§ 5. Экспоненциальное распределение и процесс Пуассона

В гл. 3 нам уже встречалось *экспоненциальное (показательное) распределение* с параметром λ , задаваемое плотностью вероятности $\lambda e^{-\lambda x}$ при $x > 0$ и 0 при $x < 0$.

Функция распределения

$$F(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x \leq 0; \end{cases} \quad (1)$$

начальные моменты $\alpha_k = k!/\lambda^k$. Основное свойство экспоненциального распределения выражается равенством

$$P(\xi > t + x | \xi > x) = e^{-\lambda t}, \quad x \geq 0, \quad t \geq 0, \quad (2)$$

следующим из определения условной вероятности:

$$P(\xi > t + x, \xi > x) = P(\xi > t + x) = e^{-\lambda(t+x)},$$

$$P(\xi > x) = e^{-\lambda x},$$

так что при делении первого равенства на второе получаем (2). Данное равенство имеет следующую важную интерпретацию. Пусть ξ — случайная длительность некоторого процесса (например, длительность безотказной работы технического элемента). Событие $\{\xi > x\}$ означает, что за время x процесс не закончится, событие $\{\xi > x + t\}$ означает, что процесс не закончился еще в течение t единиц времени. Тогда равенство (2) выражает независимость остатка времени до окончания процесса от того времени, на протяжении которого он уже продолжался, если до этого процесс не закончился. Так, экспоненциальное распределение времени жизни (безотказной работы) элемента означает независимость остатка жизни от возраста. Для сравнения допустим, что время жизни элемента ξ распределено по закону Релея

$$P(\xi < x) = 1 - e^{-x^2/2}.$$

Если элемент работает уже в течение времени x , то вероятность того, что он проработает еще хотя бы единицу времени, равна $e^{-(x+1)^2/2} / e^{-x^2/2} = e^{-x-1/2}$, т. е. экспоненциально убывает в зависимости от возраста. Свойство независимости остатка жизни от возраста выполняется только для экспоненциальной величины среди любых положительных случайных величин. Для доказательства обозначим $P(t) = 1 - F(t)$. Тогда $P(nt) = (P(t))^n$: вероятность прожить время nt равно вероятности прожить n последовательных интервалов длины t , а эти события, по условию, независимы. Возможны три случая: либо $P(1) = 0$, либо $P(1) = 1$, либо $0 < P(1) < 1$. В первом случае, вследствие равенства $P(1) = (P(1/n))^n$, $P(1/n) = 0$, т. е. $F(1/n) = 1$ для любого $n = 1, 2, \dots$. По аксиоме непрерывности $F(+0) = 1$, т. е. $\xi = 0$ с вероятностью 1, что противоречит предположению о положительности ξ . Во вто-

ром случае $P(n) = (P(1))^n = 1$, т. е. $F(n) = 0$ для любого $n \geq 1$, вопреки свойству функции распределения $F(x) \rightarrow 1, x \rightarrow \infty$. Остается третий случай: $0 < P(1) < 1$. Обозначим $P(1) = e^{-\lambda}$; очевидно, $0 < \lambda < \infty$. Тогда $P(1/n) = e^{-\lambda/n}$; $P(k/n) = (P(1/n))^k = e^{-\lambda k/n}$. Таким образом, $P(t) = e^{-\lambda t}$ для всех рациональных чисел $t > 0$. Поскольку $P(t)$ — невозрастающая функция, то это равенство выполняется для всех $t \geq 0$.

Из равенства (2)

$$\begin{aligned} P(x < \xi < x + \Delta t | \xi > x) &= (e^{-\lambda x} - e^{-\lambda(x+\Delta t)})/e^{-\lambda x} = \\ &= 1 - e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t), \end{aligned} \quad (3)$$

т. е. вероятность окончания процесса за малое время Δt равна $\lambda \Delta t$ с точностью до бесконечно малой величины более высокого порядка малости независимо от времени x , на протяжении которого процесс длится до данного момента. Равенство (3) можно записать в дифференциалах:

$$P(x < \xi < x + dt | \xi > x) = \lambda dt. \quad (4)$$

Пусть $\xi_n, n \geq 1$, — последовательность независимых экспоненциальных случайных величин с параметром λ . Последовательно отложим их на оси времени, в результате чего образуется последовательность случайных моментов времени $S_n = \xi_1 + \dots + \xi_n, n \geq 1$. Последовательность $(S_n, n \geq 1)$ называется *простейшим потоком однородных событий*, а S_n — моментом n -го события потока. Построим функцию $X(t)$, при любом $t \geq 0$ равную числу событий потока в интервале $(0; t)$. Событие $\{X(t) = n\}$ равно событию $\{S_n < t \leq S_{n+1}\}$. Отсюда

$$P(X(t) = n) = P(S_n < t) - P(S_{n+1} < t).$$

В гл. 3 было установлено, что S_n имеет плотность вероятности $\lambda^n x^{n-1} e^{-\lambda x} / (n-1)!$. Следовательно,

$$P(X(t) = n) = \frac{1}{(n-1)!} \int_0^t \lambda^n x^{n-1} e^{-\lambda x} dx - \frac{1}{n!} \int_0^t \lambda^{n+1} x^n e^{-\lambda x} dx.$$

Второй из этих интегралов проинтегрируем по частям, сделав замену $u = \lambda^{n+1} x^n / n!$, $dv = e^{-\lambda x} dx$; тогда $\int v du$ сократится с первым интегралом, и останется лишь

$$uv|_0^t = \lambda^n t^n e^{-\lambda t} / n!.$$

Итак,

$$P(X(t) = n) = e^{-\lambda t} (\lambda t)^n / n!. \quad (5)$$

Видим, что $X(t)$ — случайная величина, имеющая распределение Пуассона с параметром λt .

Пусть $\Delta_1, \Delta_2, \dots, \Delta_m$ — непересекающиеся интервалы времени длины τ_1, \dots, τ_m . Тогда числа событий простейшего потока с параметром λ в интервалах $\Delta_1, \dots, \Delta_m$ независимы и имеют распределения Пуассона с параметрами $\lambda \tau_1, \dots, \lambda \tau_m$.

Принцип доказательства вполне ясен из рассмотрения следующего частного случая. Пусть $\lambda = 1$, $m = 2$, $\Delta_1 = (0, \tau)$, $\Delta_2 = (\tau, t)$, где $0 < \tau < t$. Найдем вероятность события $A = \{X(\tau) = 2, X(t) - X(\tau) = 1\}$. Событие A состоит в том, что в интервал $(0; \tau)$ попадает два события потока, а в интервал $(\tau; t)$ — одно. Событию A благоприятствуют значения $(\xi_1, \xi_2, \xi_3, \xi_4)$, при которых $0 < \xi < \tau$, $0 < \xi_2 < \tau < \xi_1$, $\tau - (\xi_1 + \xi_2) < \xi_3 < t - (\xi_1 + \xi_2)$, $\xi_4 > t - (\xi_1 + \xi_2 + \xi_3)$. Следовательно,

$$P(A) = \int_0^{\tau} e^{-x_1} dx_1 \int_0^{\tau-x_1} e^{-x_2} dx_2 \int_{\tau-x_1-x_2}^{t-x_1-x_2} e^{-x_3} dx_3 \int_{t-x_1-x_2-x_3}^{\infty} e^{-x_4} dx_4.$$

В данном выражении внутренний интеграл вычисляется непосредственно: он равен $\exp\{-t + x_1 + x_2 + x_3\}$. Следовательно, экспоненты под знаком интеграла сокращаются, и имеем

$$\begin{aligned} P(A) &= e^{-t} \int_0^{\tau} dx_1 \int_0^{\tau-x_1} dx_2 \int_{\tau-x_1-x_2}^{t-x_1-x_2} dx_3 = e^{-t} \int_0^{\tau} dx_1 \int_0^{\tau-x_1} (t - \tau) dx_2 = \\ &= e^{-t} (t - \tau) \int_0^{\tau} (\tau - x_1) dx_1 = e^{-t} (t - \tau) \tau^2/2. \end{aligned}$$

Полученное выражение равно произведению $(e^{-\lambda\tau} (\lambda\tau)^2/2!) \times \times (e^{-\lambda(t-\tau)} \lambda (t - \tau)/1!)$ при $\lambda = 1$, что соответствует сформулированному утверждению в данном частном случае. Таким образом, если v_1, \dots, v_m — числа событий простейшего потока в интервалах $\Delta_1, \dots, \Delta_m$, то

$$P(v_1 = k_1, \dots, v_m = k_m) = \prod_{i=1}^m (e^{-\lambda\tau_i} (\lambda\tau_i)^{k_i}/k_i!). \quad (6)$$

Предположим, что каждому t из некоторого числового интервала поставлена в соответствие случайная величина $\xi(t)$. Тогда совокупность этих величин $(\xi(t))$ называется *случайным процессом*. Обычно t интерпретируется как время и $\xi(t)$ называется *значением случайного процесса* в момент t . Для упрощения символики пишут просто: случайный процесс $\xi(t)$ — без внешних скобок. Видно, что $X(t)$ — число событий простейшего потока в интервале $(0; t)$ — случайный процесс. Он называется *процессом Пуассона* с параметром λ . Поскольку число v событий простейшего потока в интервале $\Delta = (a; b)$ есть приращение процесса Пуассона в этом интервале: $v = X(b) - X(a)$, то процесс Пуассона есть *процесс с независимыми приращениями*, что означает независимость приращений в непересекающихся интервалах времени. Если известно, что в данный интервал попало n событий потока, то их моменты независимы и равномерно распределены в данном интервале. Докажем это свойство в случае $n = 2$. Возьмем интервал $(t_0; t_0 + t)$ и в нем точки $t_1 < t_1 + \Delta_1 < t_2 < t_2 + \Delta_2$. Вероятность наступления по одному событию простейшего потока в интервалах $(t_1; t_1 + \Delta_1)$ и $(t_2; t_2 +$

$+ \Delta_2$) и 0 событий в интервалах $(t_0; t_1)$, $(t_1 + \Delta_1; t_2)$, $(t_2 + \Delta; t_0 + t)$ в силу формулы (6) равна

$$(e^{-\lambda \Delta_1} \lambda \Delta_1) (e^{-\lambda \Delta_2} \lambda \Delta_2) e^{-\lambda(t_1 - t_0)} e^{-\lambda(t_2 - t_1 - \Delta_1)} e^{-\lambda(t_0 + t - t_0 - \Delta)} = e^{-\lambda t} \lambda^2 \Delta_1 \Delta_2.$$

Для получения условной вероятности это выражение нужно разделить на вероятность n событий в интервале $(t_0; t_0 + t)$. В результате получим $2\Delta_1 \Delta_2 / t^2$. При $\Delta_1 \rightarrow 0$, $\Delta_2 \rightarrow 0$ находим, что условная плотность моментов событий потока равна $2/t^2$, т. е. не зависит от x_1, x_2 ; это и достаточно для равномерности распределения. Множитель 2 (при произвольном n вместо 2 будет $n!$) объясняется тем, что мы заранее подчинили моменты событий условию $x_1 < x_2$, вследствие чего плотность в области $\{x_1 < x_2\}$ удвоилась, а в области $\{x_1 > x_2\}$ обратилась в нуль.

Процесс Пуассона имеет колоссальное число приложений к реальным явлениям. Многие приложения укладываются в следующую схему. Производятся независимые испытания с очень малой вероятностью p успеха в каждом испытании. (Отдельным испытанием может быть выстрел по самолету из зенитного орудия, прием очень слабого радиолокационного сигнала, поиск архивного документа, не положенного на место, в одной папке, и пр.). Масштаб времени выбирается так, что в единицу времени производится λ/p испытаний. По теореме Пуассона (§ 2.3) вероятность k событий за время t (т. е. в $\lambda t/p$ испытаниях) при малом p приблизительно равна

$$e^{-\lambda t} (\lambda t)^k / k!,$$

что соответствует процессу Пуассона с параметром λ .

Пример 1. Время безотказной работы радиоэлемента имеет экспоненциальное распределение с параметром λ . После отказа элемента немедленно включается новый; длительности безотказной работы различных элементов независимы. Найти вероятность того, что за время t откажет не более 3 элементов.

Искомая вероятность

$$P(X(t) \leq 3) = e^{-\lambda t} (1 + \lambda t + (\lambda t)^2/2 + (\lambda t)^3/6).$$

Пример 2. Сбой вычислительной машины образуют простейший поток с параметром λ . Машина работает в периодическом режиме: в течение времени τ_0 записывает информацию, затем в течение времени τ_1 обрабатывает ее. Во время записи сбой недопустимы, во время обработки допустим один сбой. Найти вероятность того, что за время n ($\tau_0 + \tau_1$) ни разу не возникнет недопустимая ситуация.

Поскольку числа сбоев в непересекающихся интервалах независимы, то искомая вероятность имеет вид

$$(e^{-\lambda \tau_0})^n (e^{-\lambda \tau_1} + \lambda \tau_1 e^{-\lambda \tau_1})^n = e^{-\lambda(\tau_0 + \tau_1)n} (1 + \lambda \tau_1)^n.$$

Пример 3. В приемник поступают сигналы, подлежащие регистрации. Если некоторый сигнал поступил в момент t и зарегистрирован, то все сигналы, которые поступают в интервале $(t; t + \tau)$, не регистрируются. (Постоянная τ называется «мертвым временем» приемника.) Поток сигналов — простейший с параметром λ . Найти вероятность того, что все сигналы, поступающие в интервале длины t , зарегистрированы.

Вероятность поступления n сигналов равна $e^{-\lambda t} (\lambda t)^n / n!$. Моменты $t_1 < \dots < t_n$ этих сигналов должны удовлетворять условию $0 < t_1 < t_1 + \tau < t_2 < \dots < t_{n-1} + \tau < t_n < t$. Поскольку условное распределение (t_1, \dots, t_n) при известном n равномерно, то условная вероятность регистрации всех сигналов равна отношению объема области значений (t_1, \dots, t_n) , удовлетворяющих указанным

неравенствам, к объему области $\{0 < t_1 < \dots < t_n < t\}$, равному $t^n/n!$. Сделав замену переменных $t_1 = x_1$, $t_2 = x_2 + \tau$, ..., $t_n = x_n + (n-1)\tau$, приведем «благоприятную» область к виду $\{0 < x_1 < x_2 < \dots < x_n < t - (n-1)\tau\}$. Ее объем равен $(t - (n-1)\tau)^n/n!$. Окончательно получаем для искомой вероятности выражение

$$e^{-\lambda t} \sum_{0 \leq n < t/\tau + 1} \lambda^n (t - (n-1)\tau)^n/n!.$$

Пример 4. Цель находится в зоне обзора системы наведения в течение времени t . Процесс наведения подвержен срывам, моменты которых образуют простейший поток с параметром λ . Для успешного наведения нужно, чтобы за время пребывания цели в зоне обзора был хотя бы один интервал длины τ без срывов. Обозначим вероятность этого события через $P(t)$. Легко видеть, что $P(t) = 0$ при $t < \tau$. Ограничимся случаем $\tau \leq t \leq 2\tau$. Рассматриваемое событие в этом случае равно объединению несовместных событий $A_0 = \{\xi_1 > \tau\}$, $A_1 = \{\xi_1 < t - \tau, \xi_2 > \tau\}$, $A_2 = \{\xi_1 + \xi_2 < t - \tau, \xi_3 > \tau\}$ и т. д., где ξ_1 — время от появления цели до первого срыва, ξ_2 — время от первого до второго срыва и т. д. Поскольку $\xi_1 + \dots + \xi_n$ имеет распределение Эрланга порядка n с плотностью $\lambda^n x^{n-1} e^{-\lambda x}/(n-1)!$, то

$$P(t) = e^{-\lambda \tau} + \sum_{n=1}^{\infty} \int_0^{t-\tau} (\lambda^n x^{n-1} e^{-\lambda x}/(n-1)!) dx e^{-\lambda \tau} = e^{-\lambda \tau} (1 + \lambda(t-\tau)).$$

Пример 5. В радиоканале возникают шумовые импульсы, моменты начала которых образуют простейший поток с параметром λ . Канал начинает работать в момент 0, до которого шумовых импульсов не было. Длительности импульсов — независимые случайные величины с функцией распределения $B(x)$. Найти вероятность $p_k(t)$ наличия k импульсов в момент $t > 0$.

С вероятностью $e^{-\lambda t} (\lambda t)^n/n!$ в интервале $(0; t)$ появляется n импульсов. Если такое событие наступило, то ввиду независимости моментов импульсов события $\{i\text{-й импульс не закончился до момента } t\}$ независимы. Вероятность каждого из этих событий равна

$$\frac{1}{t} \int_0^t (1 - B(t-x)) dx = \frac{1}{t} \int_0^t (1 - B(x)) dx = f.$$

Следовательно, в указанной ситуации число импульсов в момент t — биномиальная случайная величина. Учитывая вероятность условия, получаем

$$\begin{aligned} p_k(t) &= \sum_{n=k}^{\infty} e^{-\lambda t} ((\lambda t)^n/n!) C_n^k f^k (1-f)^{n-k} = \\ &= (e^{-\lambda t} (\lambda t f)^k/k!) \sum_{n=k}^{\infty} (\lambda t (1-f))^{n-k}/(n-k)! = e^{-\lambda t f} (\lambda t f)^k/k!, \end{aligned}$$

или, если возвратиться к исходным обозначениям,

$$p_k(t) = \exp \left\{ -\lambda \int_0^t (1 - B(x)) dx \right\} \left(\lambda \int_0^t (1 - B(x)) dx \right)^k / k!.$$

Таким образом, число импульсов в момент t — случайная величина, распределенная по закону Пуассона с параметром

$$\lambda \int_0^t (1 - B(x)) dx.$$

Пример 6. Процесс дробового эффекта, изучаемый в радиотехнике, характеризуется простейшим с параметром λ потоком импульсов, k -й из которых имеет

начальную ординату ξ_k и ординату $\xi_k e^{-\alpha t}$ через время t после его возникновения (равномерная скорость разрядки), где ξ_k — независимые случайные величины с функцией распределения $A(x)$. Требуется исследовать распределение суммарной ординаты импульсов $\eta(T)$ в момент T .

Пусть поток импульсов начался в момент 0. Если до момента T было n импульсов, то их ординаты в момент T — независимые случайные величины $\gamma_1, \dots, \gamma_n$. Если $Me^{it\xi_k} = \psi_0(t)$, то, поскольку $\gamma_k = \xi_k e^{-\alpha \omega_k}$, где ω_k — равномерная в интервале $(0; T)$ случайная величина (время от начала импульса до момента T), то

$$Me^{it\gamma_k} = \frac{1}{T} \int_0^T \psi_0(te^{-\alpha x}) dx = \psi(t).$$

Далее, обозначив $\varphi(t) = Me^{it\eta(T)}$, получим

$$\begin{aligned} \varphi(t) &= \sum_{n=0}^{\infty} e^{-\lambda T} ((\lambda T)^n / n!) \psi^n(t) = \exp \{-\lambda T (1 - \psi(t))\} = \\ &= \exp \left\{ -\lambda \int_0^T (1 - \psi_0(te^{-\alpha x})) dx \right\}. \end{aligned}$$

Это выражение позволяет находить моменты распределения $\eta(T)$. Так, если $M\xi_k = \beta < \infty$, то

$$M\eta(T) = -i\varphi'(0) = \frac{\lambda\beta}{\alpha} (1 - e^{-\alpha T}).$$

При $T \rightarrow \infty$, что соответствует случаю, когда поток импульсов начался «в момент $-\infty$ », $\varphi(t)$ принимает предельное выражение

$$\varphi(t) = \exp \left\{ -\lambda \int_0^{\infty} (1 - \psi_0(te^{-\alpha x})) dx \right\}.$$

§ 6. Цепь Маркова с непрерывным временем

Пусть $X(t)$ — процесс Пуассона. Возьмем моменты времени $t_1 < t_2 < \dots < t_n < t_{n+1}$. Ввиду того, что приращения $X(t)$ в интервалах $(t_1; t_2)$, $(t_2; t_3)$, ..., $(t_n; t_{n+1})$ независимы, а

$$X(t_n) = X(t_1) + (X(t_2) - X(t_1)) + \dots + (X(t_n) - X(t_{n-1})),$$

то

$$\begin{aligned} P(X(t_1) = i_1, \dots, X(t_{n+1}) = i_{n+1}) &= \\ = P(X(t_1) = i_1, X(t_2) - X(t_1) = i_2 - i_1, \dots, X(t_{n+1}) - X(t_n) = \\ = i_{n+1} - i_n) &= P(X(t_1) = i_1) P(X(t_2) - X(t_1) = i_2 - i_1) \dots \\ \dots P(X(t_{n+1}) - X(t_n) = i_{n+1} - i_n). \end{aligned} \quad (7)$$

В то же время

$$\begin{aligned} P(X(t_1) = i_1, \dots, X(t_n) = i_n) &= \\ = P(X(t_1) = i_1) \dots P(X(t_n) - X(t_{n-1}) = i_n - i_{n-1}) \end{aligned} \quad (8)$$

(в равенстве (7) нужно заменить $n+1$ на n). Разделив равенство (7) на (8), по определению условной вероятности получаем

$$\begin{aligned} P(X(t_{n+1}) = i_{n+1} | X(t_1) = i_1, \dots, X(t_n) = i_n) &= \\ = P(X(t_{n+1}) - X(t_n) = i_{n+1} - i_n). \end{aligned} \quad (9)$$

Аналогично,

$$\begin{aligned} P(X(t_n) = i_n, X(t_{n+1}) = i_{n+1}) = \\ = P(X(t_n) = i_n) P(X(t_{n+1}) - X(t_n) = i_{n+1} - i_n), \end{aligned}$$

откуда заключаем, что правая часть равенства (9) есть условная вероятность события $\{X(t_{n+1}) = i_{n+1}\}$ при условии $\{X(t_n) = i_n\}$. Итак,

$$\begin{aligned} P(X(t_{n+1}) = i_{n+1} | X(t_1) = i_1, \dots, X(t_n) = i_n) = \\ = P(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n). \end{aligned} \quad (10)$$

Свойство (10) называется *свойством марковости* случайного процесса $X(t)$ и может быть сформулировано так: *если известно настоящее (значение процесса в момент t_n), то будущее (значение в момент t_{n+1}) не зависит от прошлого (значений в моменты t_1, \dots, t_{n-1})*.

Случайный процесс $\xi(t), t \geq 0$ с конечным или счетным множеством состояний называется *марковским*, если для любых $0 \leq t_1 < t_2 < \dots < t_{n+1}$ и любых i_1, \dots, i_{n+1} выполняется свойство марковости (10) для $\xi(t)$ вместо $X(t)$. Если правая часть (10) зависит от t_n, t_{n+1} только через их разность, марковский процесс называется *однородным*, или *цепью Маркова с непрерывным временем*. Последнее название и будет ниже использоваться.

Если $\xi(t)$ — цепь Маркова с непрерывным временем, то $(\xi(an + b), n \geq 0)$ при любых $a > 0$ и $b \geq 0$ представляет собой цепь Маркова (с дискретным временем).

Обозначим

$$p_{ij}(t) = P\{\xi(t_0 + t) = j | \xi(t_0) = i\} \quad (11)$$

и назовем $p_{ij}(t)$ *вероятностью перехода* цепи Маркова (с непрерывным временем) из состояния i в состояние j за время $t \geq 0$. Из определения цепи Маркова следует, что данная функция не зависит от $t_0 \geq 0$. Из свойства марковости и формулы полной вероятности следует *уравнение Чепмена — Колмогорова*:

$$p_{ij}(t + \tau) = \sum_k p_{ik}(t) p_{kj}(\tau) \quad (12)$$

при $t \geq 0, \tau \geq 0$. Очевидно, $p_{ij}(0) = 1$ при $j = i, p_{ij}(0) = 0$ в противном случае.

Обозначим $p_i(0) = P(\xi(0) = i)$. Последовательность $p(0) = (p_i(0))$ называется *начальным распределением* цепи Маркова $(\xi(t), t \geq 0)$. Последовательность $(p_i(t))$ называется *распределением (вероятностей состояний) цепи Маркова* в момент t . По формуле полной вероятности

$$p_i(t) = \sum_j p_i(0) p_{ij}(t), \quad (13)$$

а следовательно,

$$p_j(t + \tau) = \sum_k p_k(t) p_{kj}(\tau). \quad (13a)$$

Уравнение (13a) также называется *уравнением Чепмена — Колмогорова*.

Введем матрицу

$$P(t) = (p_{ij}(t)),$$

называемую *матрицей перехода* цепи Маркова за время t . Тогда уравнения (12) и (13) в матричном виде запишутся так:

$$P(t + \tau) = P(t) P(\tau), \quad (14)$$

$$p(t) = p(0) P(t) \quad (15)$$

(здесь $p(t)$, $p(0)$ — векторы-строки).

Примем следующие предположения.

1. Существование производной матрицы $P(t)$ при $t = 0$:

$$P(\Delta) = I + A\Delta + R(\Delta)\Delta, \quad (16)$$

где I — единичная матрица, $A = (a_{ij})$ — постоянная матрица, называемая *матрицей интенсивностей*, $R(\Delta) = (r_{ij}(\Delta))$ — матрица, для которой по любому $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех i одновременно

$$\sum_j |r_{ij}(\Delta)| < \varepsilon, \quad (17)$$

если $0 \leq \Delta \leq \delta$. При $j \neq i$ число a_{ij} называется *интенсивностью перехода* цепи Маркова из состояния i в состояние j . Число $a_{ii} = -a_{ii}$ называется *интенсивностью выхода* цепи Маркова из состояния i .

2. Ограниченность интенсивностей выхода из состояний:

$$a_i \leq \lambda. \quad (18)$$

3. Условие *ординарности*: по любому $\varepsilon > 0$ найдется такое $\delta > 0$, что одновременно для всех i

$$P\left(\bigcup_{0 < t < \Delta} \{i \neq \xi(t_0 + t) \neq \xi(\Delta) \mid \xi(t_0) = i\}\right) < \varepsilon \Delta, \quad (19)$$

если $0 < \Delta \leq \delta$. Словами свойство ординарности выражается так: вероятность хотя бы двух изменений состояния цепи Маркова в бесконечно малом интервале бесконечно мала по сравнению с длиной этого интервала, независимо от состояния цепи Маркова в начальный момент интервала.

Пусть i — некоторое состояние, событие A состоит в том, что $\xi(t) = i$ при всех t , $t_0 \leq t \leq t_0 + T_0$. Тогда

$$P(A \mid \xi(t_0) = i) \leq P(\xi(t_0 + kT/n) = i, 1 \leq k \leq n \mid \xi(t_0) = i) = \\ = (p_{ii}(T/n))^n.$$

Из (16) получаем

$$p_{ii}(\Delta) = 1 + a_{ii}\Delta + r_{ii}(\Delta)\Delta = 1 - a_i\Delta + r_{ii}(\Delta)\Delta, \quad (20)$$

откуда

$$P(A \mid \xi(t_0) = i) \leq (1 - a_i T/n + r_{ii}(T/n)(T/n))^n. \quad (21)$$

Поскольку $r_{ii}(T/n) \rightarrow 0$ при $n \rightarrow \infty$ вследствие условия (17), то правая часть (20) стремится к $e^{-a_i T}$ при $n \rightarrow \infty$. Итак,

$$P(A \mid \xi(t_0) = i) \leq e^{-a_i T}. \quad (22)$$

Пусть B — событие, состоящее в том, что $\xi(t_0 + kT/n) = i$, $1 \leq k \leq n$. Если $\xi(t_0) = i$ и наступило событие B , то событие A может не наступить лишь в том случае, если для некоторого k , $1 \leq k \leq n$, найдется такое t , $t_0 + (k-1)T/n < t < t_0 + kT/n$, что

$$i = \xi(t_0 + (k-1)T/n) \neq \xi(t) \neq \xi(t_0 + kT/n) = i.$$

Вероятность такого события для данного k , вследствие условия ординарности, меньше $\varepsilon T/n$, а хотя бы для одного k — меньше εT . Таким образом,

$$P(A | \xi(t_0) = i) \geq (1 - a_i T/n + r_{ii}(T/n)(T/n))^n - \varepsilon T.$$

Перейдя к пределу при $n \rightarrow \infty$, получим

$$P(A | \xi(t_0) = i) \geq e^{-a_i T} - \varepsilon T. \quad (23)$$

Ввиду произвольности $\varepsilon > 0$, из (22) и (23) имеем

$$P(A | \xi(t_0) = i) = e^{-a_i T}. \quad (24)$$

Равенство (24) означает следующее. Если $\xi(t_0) = i$, то время, на протяжении которого $\xi(t)$ постоянно находится в состоянии i , имеет экспоненциальное распределение с параметром a_i в случае $a_i > 0$ и бесконечно с вероятностью 1 при $a_i = 0$.

Из (16) имеем

$$p_{ij}(\Delta) = a_{ij}\Delta + r_{ij}(\Delta)\Delta, \quad j \neq i. \quad (25)$$

Ввиду (20) и (25)

$$1 = \sum_j p_{ij}(\Delta) = 1 + \sum_j a_{ij}\Delta + \sum_j r_{ij}(\Delta)\Delta.$$

После сокращения на Δ получаем

$$\sum_j a_{ij} = 0, \quad (26)$$

или, что то же самое,

$$a_i = \sum_{j \neq i} a_{ij}. \quad (27)$$

Итак, интенсивность выхода из состояния i равна сумме интенсивностей перехода из i в другие состояния.

Вследствие (19) вероятность $\bar{p}_{ij}(\Delta)$ того, что в интервале длины Δ произойдет переход из i в j ($j \neq i$) и других переходов не будет при условии, что в начальный момент этого интервала $\xi(t) = i$, выражается формулой

$$\bar{p}_{ij}(\Delta) = a_{ij}\Delta + o(\Delta). \quad (28)$$

Пусть $\xi(0) = i_0$ и $S = A_1 B_1 A_2 B_2 \dots A_k B_k$, где события A_1, \dots, B_k определяются следующим образом:

$A_1 = \{\text{на отрезке } [0; t_1] \quad \xi(t) = i_0\};$

$B_1 = \{\text{на отрезке } [t_1; t_1 + \Delta_1] \text{ произошел переход из } i_0 \text{ в } i_1 \text{ и других переходов не было}\};$

$A_2 = \{\text{на отрезке } [t_1 + \Delta_1; t_2] \quad \xi(t) = i_1\};$

$B_2 = \{\text{на отрезке } [t_2; t_2 + \Delta_2] \text{ произошел переход из } i_1 \text{ в } i_2 \text{ и других переходов не было}\};$

.....

$B_k = \{\text{на отрезке } [t_k; t_k + \Delta_k] \text{ произошел переход из } i_{k-1} \text{ в } i_k \text{ и других переходов не было}\}.$

Вследствие (24) и (28) имеем

$$P(A_1) = e^{-a_{i_0} t_1}, \quad P(A_2) = e^{-a_{i_1}(t_2 - t_1 - \Delta_1)}, \quad \dots$$

$$\dots, \quad P(A_k) = e^{-a_{i_{k-1}}(t_k - t_{k-1} - \Delta_{k-1})},$$

$$P(B_1) = a_{i_0 i_1} \Delta_1 + o(\Delta_1), \quad \dots, \quad P(B_k) = a_{i_{k-1} i_k} \Delta_k + o(\Delta_k).$$

Перемножив эти равенства, получаем

$$P(S) = e^{-a_{i_0} t_1} a_{i_0 i_1} \Delta_1^{-a_{i_1}(t_2 - t_1)} a_{i_1 i_2} \Delta_2 \dots e^{-a_{i_{k-1}}(t_k - t_{k-1})} a_{i_{k-1} i_k} \Delta_k + o(\Delta_1 \dots \Delta_k), \quad (29)$$

или более просто с помощью дифференциалов

$$P(S) = \prod_{m=1}^k \{e^{-a_{i_{m-1}}(t_m - t_{m-1})} a_{i_{m-1} i_m} dt_m\}, \quad (30)$$

где $dt_m = \Delta_m$, $1 \leq m \leq k$; $t_0 = 0$.

Равенство (30) раскрывает вероятностный смысл цепи Маркова с непрерывным временем, указывая способ последовательного построения $\xi(t)$ как функции времени.

В предположениях 1—3 функция $\xi(t)$ с вероятностью 1 ступенчата, т. е. изменяется скачками, причем в любом конечном интервале их лишь конечное число. Обозначим через t_n момент n -го скачка при $n \geq 1$; $t_0 = 0$. Обозначим также через i_n значение $\xi(t)$ в интервале между t_n и t_{n+1} ($n \geq 0$).

Для определения i_0 реализуем случайное испытание, исход i_0 которого равен i с вероятностью $p_i(0)$. Если $a_{i_0} = 0$, полагаем $\xi(t) = i_0$, $0 \leq t < \infty$. Если $a_{i_0} > 0$, берем случайную величину τ_0 , имеющую экспоненциальное распределение с параметром a_{i_0} , и полагаем $\xi(t) = i_0$ при $0 \leq t < \tau_0 = t_1$. Если функция $\xi(t)$ определена при $t < t_n$, причем $\xi(t_n - 0) = i_{n-1}$, то реализуем случайное испытание, исход i_n которого равен j с вероятностью $a_{i_{n-1}j}/a_{i_{n-1}}$, и при $a_{i_{n-1}} > 0$ случайную величину τ_n , имеющую экспоненциальное распределение с параметром a_{i_n} . Полагаем $t_{n+1} = t_n + \tau_n$, $\xi(t) = i_n$ при $t_n \leq t < t_{n+1}$. При $a_{i_{n-1}} = 0$ полагаем $\xi(t) = a_{i_n}$, $t_n \leq t < \infty$.

Испытания и величина, реализуемые в процессе построения $\xi(t)$, независимы от всего предыдущего.

Пример. Цепь Маркова $\xi(t)$ с непрерывным временем имеет состояния 1, 2, 3, матрицу интенсивностей вида

$$A = \begin{pmatrix} -1 & 0,25 & 0,75 \\ 3 & -4 & 1 \\ 2 & 0 & -2 \end{pmatrix}$$

и начальное распределение

$$p^{(0)} = (0,2; 0,5; 0,3).$$

Покажем, как строится функция $\xi(t)$.

Реализуем испытание с исходами 1, 2, 3, имеющими соответственно вероятности 0,2; 0,5; 0,3. Допустим, что наступил исход 2. Реализуем случайную величину τ_0 — экспоненциальную с параметром 4 — и полагаем $\xi(t) = 2$ при $0 \leq t < \tau_0 = t_1$. Затем реализуем испытание с исходами 1 и 3, имеющими вероятности $3/4$ и $1/4$ (из второй строки матрицы A). Пусть исход испытания равен 1. Тогда реализуем τ_1 — экспоненциальную с параметром 1 — и полагаем $\xi(t) = 1$ при $\tau_0 \leq t < \tau_0 + \tau_1 = t_2$. Теперь реализуем испытание с исходами 2 и 3, имеющими вероятности $0,25/1 = 0,25$ и $0,75/1 = 0,75$ (из первой строки A). При исходе 3 реализуем τ_2 — экспоненциальную с параметром 2 — и полагаем $\xi(t) = 3$ при $t_2 \leq t < t_2 + \tau_2 = t_3$. Этот процесс продолжается до бесконечности.

Положим в уравнении (14) $\tau = \Delta$ и подставим $P(\Delta)$ из (16). Тогда

$$P(t + \Delta) = P(t)(I + A\Delta + R(\Delta)\Delta),$$

или

$$\frac{1}{\Delta}(P(t + \Delta) - P(t)) = P(t)A + P(t)R(\Delta).$$

При $\Delta \rightarrow 0$ получаем уравнение

$$\frac{d}{dt}P(t) = P(t)A, \quad (31)$$

или в скалярном виде

$$p'_{ij}(t) = \sum_k p_{ik}(t) a_{kj}. \quad (32)$$

Для $p(t) = (p_j(t))$ имеет место аналогичное матричное уравнение

$$\frac{d}{dt}p(t) = p(t)A, \quad (33)$$

или в скалярном виде

$$p'_j(t) = \sum_k p_k(t) a_{kj}. \quad (34)$$

Каждое из уравнений (31), (33), или, что то же самое, (32), (34) называется *системой прямых дифференциальных уравнений Колмогорова*.

Теперь снова используем то же уравнение (14). Записав его в виде

$$P(t + \Delta) = P(\Delta)P(t)$$

и подставив $P(\Delta)$ из (16), имеем

$$P(t + \Delta) = (I + A\Delta + R(\Delta)\Delta)P(t),$$

или

$$\frac{1}{\Delta}(P(t + \Delta) - P(t)) = AP(t) + R(\Delta)P(t).$$

При $\Delta \rightarrow 0$ получаем систему обратных дифференциальных уравнений Колмогорова:

$$\frac{d}{dt} P(t) = AP(t), \quad (35)$$

или

$$p'_{ij}(t) = \sum_k a_{ik} p_{kj}(t). \quad (36)$$

Из (16) следует непрерывность $P(t)$. Каждая из систем (31), (33), (35) имеет единственное непрерывное при $t \geq 0$ решение, удовлетворяющее начальному условию $P(0) = 1$ для систем (31), (35) и $p(0)$ равно начальному распределению для системы (33).

Пример. Автоматический подсчет различных однородных событий может быть реализован на счетчиках по модулю вместо обычного сумматора. Пусть $X(t)$ — процесс Пуассона с параметром λ , $\xi(t)$ — результат приведения $X(t)$ по модулю m , т. е. $\xi(t) = 0, 1, \dots, m-1$. Тогда $\xi(t)$ — цепь Маркова с непрерывным временем. Ее начальное состояние есть 0 с вероятностью 1; матрица интенсивностей

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda & \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \lambda & 0 & 0 & \dots & -\lambda \end{pmatrix}.$$

Следовательно, система прямых уравнений Колмогорова имеет вид

$$\begin{aligned} p'_0(t) &= -\lambda p_0(t) + \lambda p_{m-1}(t), \\ p'_k(t) &= -\lambda p_k(t) + \lambda p_{k-1}(t), \quad 1 \leq k \leq m-2. \end{aligned}$$

Начальное распределение

$$p_0(0) = 1, \quad p_k(0) = 0, \quad 1 \leq k \leq m-2.$$

Применим к решению данной системы преобразование Лапласа, обозначив

$$\pi_k(z) = \int_0^\infty e^{-zt} p_k(t) dt.$$

Имеем

$$\begin{aligned} (z + \lambda) \pi_0(z) &= \lambda \pi_{m-1}(z) + 1, \\ (z + \lambda) \pi_k(z) &= \lambda \pi_{k-1}(z), \quad 1 \leq k \leq m-2. \end{aligned}$$

Выражая $\pi_k(z)$ последовательно через $\pi_0(z)$, придем к формуле

$$\pi_0(z) = (z + \lambda)^{m-1} / ((z + \lambda)^m - \lambda^m)$$

и затем к формуле

$$\pi_k(z) = \lambda^k (z + \lambda)^{m-k-1} / ((z + \lambda)^m - \lambda^m), \quad 1 \leq k \leq m-1.$$

Поскольку корни знаменателя

$$z_j = -\lambda (e^{2\pi i j/m} - 1), \quad 0 \leq j \leq m-1,$$

различны, то $p_k(t)$ выражаются в виде линейных комбинаций экспонент $e^{z_j t}$.

Интересно заметить, что $z_0 = 0$, все остальные z_k имеют отрицательную действительную часть. Отсюда после вычисления соответствующих коэффициентов получаем

$$p_k(t) = \frac{1}{m} + \gamma_k(t), \quad 0 \leq k \leq m-1,$$

где $\gamma_k(t)$ экспоненциально быстро стремятся к нулю при $t \rightarrow \infty$.

Матричное уравнение (31) при условии $P(0) = I$ имеет решение

$$P(t) = e^{At}, \quad (37)$$

или, что то же самое,

$$P(t) = \sum_{n=0}^{\infty} A^n t^n / n!, \quad (38)$$

где сходимость ряда из матриц понимается как сходимость рядов из их элементов. Формула (38) эффективно используется при малых значениях элементов матрицы A . Так если $a_i \leq \varepsilon$ при всех i , то сумма абсолютных значений элементов строки матрицы A^n не больше $(2\varepsilon)^n$.

В случаях, когда a_i не слишком сильно различаются, можно рекомендовать следующий прием. Возьмем $\lambda \geq \max_i \{a_i\}$. Тогда

$$P(t) = e^{-\lambda t} \sum_{n=0}^{\infty} (\lambda I + A)^n t^n / n!. \quad (39)$$

Сходимость ряда в правой части (39) монотонна, так как элементы матрицы $\lambda I + A$ и всех ее степеней неотрицательны.

§ 7. Эргодическое распределение

Пусть $p_i(t) \xrightarrow{t \rightarrow \infty} \pi_i$, где (π_i) — распределение вероятностей. Тогда цепь Маркова с непрерывным временем называется *эргодической*, а (π_i) — ее *эргодическим распределением*.

Если при некотором $\Delta > 0$ цепь Маркова $(\xi(n\Delta))$ имеет эргодическое распределение, то и $(\xi(t))$ имеет эргодическое распределение, совпадающее с первым. Поэтому эргодичность цепи Маркова с непрерывным временем можно установить по эргодичности цепи $(\xi(n\Delta))$. Более того, можно доказать следующее свойство. Для данных i, j либо $p_{ij}(t) > 0$ для всех $t > 0$, либо $p_{ij}(t) = 0$ для всех $t > 0$. Таким образом, если все состояния сообщаются, то для эргодичности $(\xi(t))$ достаточно положительности $(\xi(n\Delta))$. Для цепи Маркова с непрерывным временем строится *граф переходов*: от вершины i к вершине j проводится направленная дуга, если $a_{ij} > 0$. Состояние j *достижимо* из состояния i , если имеется путь от i к j . *Совокупность состояний, замкнутые классы, неразложимость* определяются по графу переходов так же, как и в случае цепей Маркова с дискретным временем.

Пусть $\pi = (\pi_i)$ — эргодическое распределение. Проинтегрируем обе части уравнения (33) от t до $t+1$:

$$p(t+1) - p(t) = \left(\int_t^{t+1} p(x) dx \right) A.$$

Левая часть при $t \rightarrow \infty$ стремится к нулевому вектору, правая — к πA . Следовательно, эргодическое распределение удовлетворяет уравнению $\pi A = 0$, или в скалярном виде

$$a_j \pi_j = \sum_{i \neq j} \pi_i a_{ij}. \quad (40)$$

К этим уравнениям добавляется условие нормирования

$$\sum_j \pi_j = 1. \quad (41)$$

Уравнение (40) удобно составлять, пользуясь графом переходов: левая часть есть произведение π_j на сумму a_{ij} по всем дугам, исходящим из вершины j , правая часть — сумма $\pi_i a_{ij}$ по всем дугам, входящим в эту вершину. Аналогия с уравнениями для электрических цепей имеет следствием совпадение расчетных методов, в которых эффективно используются понятия теории графов.

В качестве частного случая рассмотрим процесс размножения и гибели — цепь Маркова ($\xi(t)$) с непрерывным временем, множеством состояний \mathbb{Z}^+ и интенсивностями перехода a_{ij} , равными 0 при $|j - i| > 1$. Таким образом, достаточно задать $a_{i,i+1} = \lambda_i$ («интенсивности размножения») и $a_{i,i-1} = \mu_i$ («интенсивности гибели»).

Для эргодического распределения имеем уравнения

$$\pi_j (\lambda_j + \mu_j) = \lambda_{j-1} \pi_{j-1} + \mu_{j+1} \pi_{j+1}, \quad (42)$$

где $\lambda_{-1} = \pi_{-1} = 0$.

Уравнения (42) можно записать так:

$$\mu_{j+1} \pi_{j+1} - \lambda_j \pi_j = \mu_j \pi_j - \lambda_{j-1} \pi_{j-1}. \quad (43)$$

Из (43) вытекает, что $\mu_j \pi_j - \lambda_{j-1} \pi_{j-1}$ не зависит от j . Положив в (42) $j = 0$, получим $\mu_1 \pi_1 - \lambda_0 \pi_0 = 0$. Итак, $\mu_j \pi_j - \lambda_{j-1} \pi_{j-1} = 0$, т. е.

$$\pi_j = \frac{\lambda_{j-1}}{\mu_j} \pi_{j-1}. \quad (44)$$

Предположим, что имеет место один из двух случаев:

а) $\lambda_i > 0$, $i \geq 0$; $\mu_i > 0$, $i > 0$;

б) $\lambda_i > 0$, $0 \leq i \leq n-1$, $\lambda_n = 0$; $\mu_i > 0$, $1 \leq i \leq n$; $\sum_{i=0}^n p_i(0) = 1$.

В случае а) эргодическое распределение, если оно существует, имеет вид

$$\pi_j = \pi_0 \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j}, \quad j \geq 0,^1 \quad (45)$$

¹ При $j = 0$ выражение $\lambda_0 \dots \lambda_{j-1} / (\mu_1 \dots \mu_j)$ принимается равным 1, как и любое произведение по пустому множеству индексов.

где

$$\pi_0 = \left(\sum_{j=0}^{\infty} \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} \right)^{-1}; \quad (46)$$

в случае б) эргодическое распределение заведомо существует и имеет вид (45) при $0 \leq j \leq n$, где

$$\pi_0 = \left(\sum_{j=0}^n \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} \right)^{-1}. \quad (47)$$

Для существования эргодического распределения в случае а) необходимо, чтобы

$$\sum_{j=0}^{\infty} \lambda_0 \dots \lambda_{j-1} / (\mu_1 \dots \mu_j) < \infty \quad (48)$$

(иначе по формуле (46) будет $\pi_0 = 0$ и затем по формуле (47) и все вообще $\pi_j = 0$). Если $\lambda_j \leq c$, $\mu_j \leq c$, то условие (48) достаточно для существования эргодического распределения.

Докажем более частный, но довольно употребительный признак эргодичности процесса размножения и гибели: $\lambda_i > 0$, $\mu_i > 0$ для всех i ; при $i \geq n_0$ $\lambda_i \leq a$, $\mu_i \geq b$, где $a < b$.

Рассмотрим цепь Маркова с дискретным временем $\xi_n = \xi(n\Delta)$, $n \geq 0$, где $\Delta > 0$, и предположим, что $\xi_n = i$. Оценим $M(\xi_{n+1} - \xi_n | \xi_n = i)$. Для того чтобы было $\xi_{n+1} \geq j > i$, нужно, чтобы в некоторых интервалах dt_1, \dots, dt_{j-i} между $n\Delta$ и $(n+1)\Delta$ произошли переходы из i в $i+1$, из $i+1$ в $i+2$, ..., из $j-1$ в j . Вероятность такого события не больше выражения $\lambda_i dt_1 \lambda_{i+1} dt_2 \dots \lambda_{j-1} dt_{j-1}$ (при этом возможны промежуточные переходы «сверху вниз»). Проинтегрировав его по t_1, \dots, t_{j-i} , получим выражение $\lambda_i \lambda_{i+1} \dots \lambda_j \Delta^{j-i} / (j-i)!$, оценивающее сверху условную вероятность события $\{\xi_{n+1} \geq j\}$. При том же условии событие $\{\xi_{n+1} = i-1\}$ наступит, если в течение времени Δ произойдет хотя бы один скачок вниз, а скачков вверх не будет. Вероятность такого события не меньше, чем $e^{-c\Delta} (1 - e^{-\mu_i \Delta})$. При малом Δ и $i \geq n_0$ это выражение больше $(b - \varepsilon) \Delta$. При том же условии вероятность события $\{\xi_{n+1} \geq j\}$ меньше $(a\Delta)^{j-i} / (j-i)!$. Из этих оценок следует, что

$$\begin{aligned} M(\xi_{n+1} - \xi_n | \xi_n = i) &\leq \sum_{j=i+1}^{\infty} (a\Delta)^{j-i} / (j-i)! - (b - \varepsilon) \Delta = \\ &= e^{a\Delta} - 1 - (b - \varepsilon) \Delta < -\rho, \quad \rho > 0, \end{aligned} \quad (49)$$

поскольку $e^{a\Delta} \sim 1 + a\Delta$, $a < b$. Для $i < n_0$

$$M(\xi_{n+1} - \xi_n | \xi_n = i) \leq e^{A\Delta} - 1,$$

где $A = \max \{\lambda_0, \dots, \lambda_{n_0-1}\}$. Выполнены все условия эргодической теоремы § 4. Следовательно, существует эргодическое распределение.

§ 8. Ветвящийся процесс

Рассматривается популяция, в любой момент $t \geq 0$ состоящая из $\xi(t)$ элементов, которые будем условно называть частицами. В приложениях это могут быть бактерии, неисправности технической системы, трещины в металлической конструкции, завязанные траектории (в том числе шумовые) при обработке радиолокационной информации и многое другое. Пусть $\xi(t) = n$, т. е. в момент t имеется n частиц. За время dt каждая из них, независимо от остальных, может превратиться в $k+1$ частиц с вероятностью $\lambda_k dt$, $k = -1, 0, 1, \dots$. В частности $\lambda_{-1} dt$ — вероятность гибели частицы, $\lambda_1 dt$ — вероятность деления на две и т. п. От λ_0 , как легко видеть, ничто не зависит, поэтому положим $\lambda_0 = 0$. Вероятность превращения за время dt составляет λdt , где $\lambda = \lambda_{-1} + \lambda_1 + \lambda_2 + \dots$. Вероятность сохранения состояния частицы равна $1 - \lambda dt$.

Обозначим через $a_{nm} dt$ вероятность события $\{\xi(t+dt) = m\}$ при условии $\{\xi(t) = n\}$, т. е. a_{nm} — интенсивность перехода $\xi(t)$ из состояния n в состояние m . Возьмем вместо dt конечный малый интервал времени Δ . Переход из n в m за время Δ может произойти следующими несовместными способами.

Событие A_k : k -я частица превратилась в $m - n + 1$ частиц; остальные частицы сохранили прежнее состояние.

Событие B : переход из n в m произошел посредством превращения хотя бы двух частиц. В силу независимости превращения частиц

$$P(A_k) \sim \lambda_{m-n} \Delta (1 - \lambda \Delta)^{n-1} \sim \lambda_{m-n} \Delta;$$

$P(B) \leq C_n^2 \Delta^2$ (сомножитель C_n^2 появляется вследствие перебора по парам частиц, с которыми могли произойти превращения).

Отсюда вероятность искомого события

$$P(A_1) + \dots + P(A_n) + P(B) \sim n \lambda_{m-n} \Delta,$$

т. е.

$$a_{nm} = n \lambda_{m-n}.$$

Обозначим $\psi(z) = \sum_{k=-1}^{\infty} \lambda_k z^k$, $|z| \leq 1$, $p_{in}(t) = P(\xi(t) = n | \xi(0) = i)$, $\varphi_i(z, t) = \sum_{n=0}^{\infty} p_{in}(t) z^n$, $|z| \leq 1$.

Назовем потомством данной частицы ее саму, те частицы, в которые она превратилась, и все следующие — поколения. Тогда при $\xi(0) = n$ имеем $\xi(t) = \xi_1(t) + \dots + \xi_n(t)$, где $\xi_j(t)$ — объем в момент t потомства j -й частицы из присутствовавших вначале n частиц. Поскольку поколения различных частиц превращаются независимо, то $\xi_1(t)$, ..., $\xi_n(t)$ — независимые случайные величины. Отсюда

$$\varphi_i(z, t) = (\varphi_1(z, t))^i, \quad i = 1, 2, \dots; \quad (1)$$

следовательно, основной характеристикой является $\varphi_1(z, t)$. Итак, пусть $\xi(0) = 1$. С вероятностью $1 - \lambda dt$ будет $\xi(dt) = 1$ и в этом

случае $\xi(t)$ имеет то же распределение, что и $\xi(t - dt)$: превращения частицы за время от 0 до dt не засчитываются. С вероятностью $\lambda_k dt$ частица за время dt превратится в $k + 1$ частиц, так что в момент dt начинается независимое поведение k потомков данной частицы. С учетом сказанного

$$\begin{aligned}\varphi_1(z, t) &= Mz^{\xi(t)} = (1 - \lambda dt) Mz^{\xi(t-dt)} + \sum_{k=-1}^{\infty} \lambda_k dt (Mz^{\xi(t-dt)})^{k+1} = \\ &= (1 - \lambda dt) \varphi_1(z, t - dt) + \sum_{k=-1}^{\infty} \lambda_k \varphi_1^{k+1}(z, t - dt) dt.\end{aligned}$$

Выражение $(1 - \lambda dt) \varphi_1(z, t - dt)$ запишем в виде

$$\varphi_1(z, t) - \lambda \varphi_1(z, t) dt - \frac{\partial \varphi_1}{\partial t} dt,$$

выражение $\sum_{k=-\infty}^{\infty} \lambda_k \varphi_1^{k+1}(z, t)$ — в виде $\varphi_1(z, t) \psi(\varphi_1(z, t))$ (замена $t - dt$ на t возможна, так как все выражение потом умножается на dt). Итак,

$$\varphi_1(z, t) = \varphi_1(z, t) - \lambda \varphi_1(z, t) dt - \frac{\partial \varphi_1}{\partial t} dt + \varphi_1(z, t) \psi(\varphi_1(z, t)) dt.$$

Данное уравнение можно записать следующим образом:

$$-\frac{\partial \varphi_1}{\partial t} + \varphi_1(\lambda - \psi(\varphi_1(z, t))) = 0. \quad (2)$$

Из него можно, в частности, найти $a(t) = M(\xi(t) | \xi(0) = 1)$. Имеем

$$-\frac{\partial \varphi}{\partial z} = \sum_{n=0}^{\infty} n p_{1n}(t) z^{n-1},$$

откуда

$$a(t) = \left. \frac{\partial \varphi(z, t)}{\partial z} \right|_{z=1}.$$

В то же время

$$\varphi_1(1, t) = \sum_{n=0}^{\infty} p_{1n}(t) = 1,$$

так что

$$\psi(\varphi_1(1, t)) = \psi(1) = \sum_{k=-1}^{\infty} \lambda_k = \lambda; \quad (3)$$

$$\psi'(1) = \sum_{k=-1}^{\infty} k \lambda_k = \alpha.$$

С учетом этих равенств, продифференцировав (2) по z и положив $z = 1$, получим уравнение

$$a'(t) = \alpha a(t). \quad (4)$$

Поскольку, очевидно, $a(0) = 1$, то из (4) получаем

$$a(t) = e^{\alpha t}. \quad (5)$$

Итак, при конечном α математическое ожидание числа частиц изменяется с течением времени по экспоненциальному закону. При $\alpha > 0$ $a(t)$ возрастает до бесконечности, при $\alpha < 0$ убывает к нулю, при $\alpha = 0$ остается постоянным. Это учитывается при решении практических задач. Например, при обработке радиолокационной информации выбирают пороговый уровень таким образом, чтобы цепочки шумовых импульсов не приводили к размножению ложных трасс. Еще пример: при техническом обслуживании систем устранение одних отказов может приводить к другим, так что в результате получается ветвящийся процесс.

Говорят, что ветвящийся процесс вырождается, если $\xi(t) = 0$, начиная с некоторого (случайного) t . При $\alpha < 0$ процесс вырождается с вероятностью 1: из (5) имеем $P(\xi(t) \geq 1) \leq a(t) \xrightarrow{t \rightarrow \infty} 0$. Обозначим через β *вероятность вырождения* ветвящегося процесса, считая $\alpha \geq 0$. Для нахождения β используем тот же прием, что и при выводе уравнения (2). Либо начальная частица за время dt не превратится, и тогда вероятность вырождения процесса в момент dt снова будет равна β , либо превратится в $k+1$ частиц. В последнем случае для вырождения нужно, чтобы потомство всех этих частиц выродилось. Итак,

$$\beta = (1 - \lambda dt) \beta + \sum_{k=1}^{\infty} \lambda_k \beta^{k+1} dt,$$

или

$$\lambda \beta = \beta \psi(\beta). \quad (6)$$

Если $\lambda_1 = 0$, то $\beta = 0$: $\xi(t)$ в этом случае не убывает. Если $\lambda_1 > 0$, то $\beta > 0$: с вероятностью, большей $(1 - \exp\{-\lambda_1\}) \times \exp\{-(\lambda - \lambda_1)\}$, за время 1 произойдет вырождение. Этот последний случай мы и рассмотрим. Сократив (6) на β , получим уравнение

$$\psi(\beta) = \lambda. \quad (7)$$

Одним его корнем является $\beta = 1$. Далее, при $0 < z < 1$ имеем $\psi'(z) = \sum k \lambda_k z^{k-1}$, $\psi'(1) = \alpha$ (это уже нами отмечалось) $\psi''(z) = \sum k(k-1) \lambda_k z^{k-2} > 0$. Итак, $\psi(z)$ — выпуклая функция, в точке $z = 1$ равная λ и имеющая производную α . Если $\alpha = 0$, то $\psi'(z) < 0$ при $0 < z < \beta$ (производная выпуклой функции возрастает, а следовательно, $\psi(z) > \lambda$ при $0 \leq z < 1$). Итак, корень $\beta = 1$ — единственный. При $\alpha > 1$, так как $\psi(+0) = \infty$, существует корень β_0 , $0 < \beta_0 < 1$, уравнения (7). Докажем, что именно β_0 , а не 1, его вероятность вырождения.

Пусть ξ_n — объем n -го поколения начальной частицы. Легко видеть, что каждая частица n -го поколения превращается в $k+1$ частиц $(n+1)$ -го поколения с вероятностью $f_k = \lambda_k/k$. При $\alpha > 0$ имеем $\sum_k k f_k > 0$, т. е. $\sum_{k=1}^m k f_k = 2\delta > 0$ при некотором m . Из гл. 2 при $\xi_0 = N$ имеем

$$P(\xi_1 \geq N(1 + \delta)) \geq 1 - \rho^N, \quad 0 < \rho < 1.$$

Повторяя те же рассуждения, получаем

$$P(\xi_2 \geq N(1 + \delta)^2) \geq (1 - \rho^N)(1 - \rho^{(1+\delta)N})$$

и т. д. Следовательно,

$$P(\lim_{n \rightarrow \infty} \xi_n = 0) \leq \rho^N + \rho^{(1+\delta)N} + \dots < \rho^N / (1 - \rho^{\delta N}).$$

Таким образом, вероятность вырождения поколения, содержащего достаточно много частиц, не равна единице. Но эта вероятность равна β^N . Следовательно, $\beta < 1$. Остается единственная возможность: $\beta = \beta_0$.

Пример. Пусть частица может лишь превратиться в две с интенсивностью $(1 + \alpha)/2$ либо погибнуть, не дав потомства, с интенсивностью $(1 - \alpha)/2$. Поскольку $\lambda = 1$, то $f_1 = (1 + \alpha)/2$, $f_{-1} = (1 - \alpha)/2$, $\Sigma k f_k = \alpha$. В данном случае

$$\psi(z) = z(1 + \alpha)/2 + z^{-1}(1 - \alpha)/2 = (z^2(1 + \alpha) + 1 - \alpha)/2z.$$

Уравнение (7) принимает вид $z^2(1 + \alpha) - 2z + 1 - \alpha = 0$. Разделив это выражение на $z - 1$, получим $z(1 + \alpha) + \alpha - 1 = 0$. Корень $\beta_0 = (1 - \alpha)/(1 + \alpha)$ есть вероятность вырождения при $0 \leq \alpha \leq 1$. При $\alpha < 0$, как следует из общего случая, $\beta = 1$.

Задачи

1. Матрица перехода цепи Маркова имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,5 & 0,5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Все ли состояния сообщающиеся? Найти явные формулы для $p_{ij}^{(n)}$ и пределы этих выражений при $n \rightarrow \infty$.

2. Матрица перехода цепи Маркова имеет вид

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0,4 & 0,6 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Найти пределы

$$\lim_{n \rightarrow \infty} \frac{1}{n} (p_{ij}^{(1)} + p_{ij}^{(2)} + \dots + p_{ij}^{(n)}).$$

3. Пусть $X(t)$ — процесс Пуассона с параметром λ . Найти матрицу перехода цепи Маркова $\xi_n = X(nh)$, где $h > 0$ — постоянная.

4. Пусть ξ_n, η_n — независимые в совокупности случайные величины, распределенные по закону Пуассона с параметром a . Введем случайные величины

$$\zeta_1 = \xi_1 + \xi_2 + \dots + \xi_{\eta_1},$$

$$\zeta_2 = \xi_1 + \xi_2 + \dots + \xi_{\eta_1 + \eta_2},$$

$$\dots \dots \dots$$

Доказать, что ζ_n — цепь Маркова, и найти ее матрицу перехода.

5. Матрица интенсивностей перехода цепи Маркова с непрерывным временем имеет вид

$$\Lambda = \begin{pmatrix} -(\lambda_0 + \lambda_1) & \lambda_0 & \lambda_1 \\ 0 & -\lambda_2 & \lambda_2 \\ \mu\theta & 0 & -\mu\theta \end{pmatrix},$$

где $\lambda_0, \lambda_1, \lambda_2, \mu$ — положительные числа. Найти эргодическое распределение при $\theta = 1$ и $\theta = 0$.

6. Имеется цепь Маркова ξ_n с конечным множеством состояний и матрицей перехода P , $X(t)$ — независимый от (ξ_n) процесс Пуассона с параметром λ . Положим

$$\eta(t) = \xi_{X(t)}.$$

Доказать, что $\eta(t)$ — цепь Маркова с непрерывным временем, и найти ее интенсивности перехода. Доказать, что если все состояния ξ_n сообщаются, то $\eta(t)$ имеет эргодическое распределение.

7. Пусть $(\xi_n^{(e)})$ — цепь Маркова с состояниями 1, 2, 3, 4, начальным состоянием 1 и матрицей перехода

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} - \varepsilon & \varepsilon & 0 \\ \frac{1}{3} - 2\varepsilon & \frac{2}{3} & 0 & 2\varepsilon \\ 0 & \varepsilon & 0 & 1 - \varepsilon \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

где $0 < \varepsilon \leq \frac{1}{6}$. Положим $\eta_n^{(e)} = 1$ при $\xi_n^{(e)}$ равном 1 или 2, $\eta_n^{(e)} = 2$ при $\xi_n^{(e)}$ равном 3 или 4. Найти пределы

$$p_j(t) = \lim_{\varepsilon \rightarrow 0} P(\eta_{[t/\varepsilon]}^{(e)} = j), \quad t > 0, \quad j = 1, 2.$$

8. Пусть $(\xi_n^{(e)})$ — цепь Маркова с состояниями 1, 2, 3, 4, 5, начальным состоянием 1 и матрицей перехода

$$\begin{pmatrix} 0,5 - \varepsilon^2 & 0,5 & 0 & 0 & \varepsilon^2 \\ 0,9 & 0,1 - 2\varepsilon & 0 & 2\varepsilon & 0 \\ 0 & 0 & 0,7 & 0,3 - \varepsilon & \varepsilon \\ 0 & 0 & 0,2 & 0,8 - 2\varepsilon & 2\varepsilon \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

где $\varepsilon > 0$. Найти

$$p(t) = \lim_{\varepsilon \rightarrow 0} P(\xi_{[t/\varepsilon^2]}^{(e)} = 5).$$

9. Пусть G — аддитивная абелева группа, (g_z) — распределение вероятностей на G ; $p_z^{(0)}$ равно 1 при $z = 0$ и равно 0 при $z \neq 0$. Определим $p_z^{(n)}$, $z \in G$, рекуррентной формулой

$$p_z^{(n)} = \sum_{x \in G} p_x^{(n-1)} g_{z-x}.$$

Доказать, что существует предел $(p_z^{(n)})$ при $n \rightarrow \infty$, и найти его вид. Когда

$$p_z^{(n)} \xrightarrow{n \rightarrow \infty} \frac{1}{m}, \quad z \in G, \quad \text{где } m — \text{порядок группы } G?$$

§ 1. Пространства и сходимость

В математическом анализе встречаемся прежде всего с понятием сходимости последовательности (x_n) :

$$x_n \xrightarrow{n \rightarrow \infty} a \Leftrightarrow \forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \{|x_n - a| < \varepsilon\}.$$

Аналогичное понятие вводится и для непрерывной переменной t вместо n :

$$x(t) \xrightarrow{t \rightarrow t_0} a \Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall t: |t - t_0| < \delta \{|x(t) - a| < \varepsilon\}.$$

Это понятие переносится и на последовательности (x_n) , и на функции $x(t)$ со значениями в любом метрическом пространстве X , т. е. пространстве, где введено понятие расстояния $\rho(x, y)$ между любыми $x \in X, y \in X$, а именно:

$$x_n \xrightarrow{n \rightarrow \infty} a \Leftrightarrow \forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \{\rho(x_n, a) < \varepsilon\}$$

и соответственно

$$x(t) \xrightarrow{t \rightarrow t_0} a \Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall t: |t - t_0| < \delta \{\rho(x(t), a) < \varepsilon\}.$$

Например, если $X = \mathbb{R}^m$, то $\rho(x, y) = \left(\sum_{i=1}^m (x_i - y_i)^2\right)^{1/2}$, где x_i ,

y_i — компоненты векторов x, y , и тогда $x(t) \rightarrow a = (a_1, \dots, a_m)$ означает, что отклонение $x(t)$ от a при $|t - t_0| < \delta$ попадает в «трубку»

$$\sum_{i=1}^m (x_i(t) - a_i)^2 < \varepsilon^2.$$

Поскольку в «круговую» окрестность можно вписать «прямоугольную» и в последнюю — «круговую» (например, при $m = 2$ $\{|x - x_0| < \varepsilon, |y - y_0| < \varepsilon\} \subset \{(x - x_0)^2 + (y - y_0)^2 \leq (\varepsilon\sqrt{2})^2\} \subset \{|x - x_0| < \varepsilon\sqrt{2}, |y - y_0| < \varepsilon\sqrt{2}\}$), то сходимость $x(t) \rightarrow a$ эквивалентна сходимости $x_i(t) \rightarrow a_i$ для всех $i, 1 \leq i \leq m$. Это позволяет при рассмотрении вопросов сходимости изучать одномерные последовательности и функции вместо многомерных.

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ — вероятностное пространство, ξ, η — одномерные случайные величины: $\xi = f(\omega), \eta = g(\omega), \omega \in \Omega$. Введя то или иное понятие расстояния между функциями ω , получим соответствующее понятие расстояния между случайными величинами ξ и η .

Основную роль играют следующие три понятия, которыми мы и ограничимся:

$$\rho_0(\xi, \eta) = \min\{\varepsilon: \mathbf{P}(|\xi - \eta| > \varepsilon) \leq \varepsilon\}, \quad (1)$$

$$\rho_1(\xi, \eta) = \mathbf{M}|\xi - \eta|, \quad (2)$$

$$\rho_2(\xi, \eta) = (\mathbf{M}(\xi - \eta)^2)^{1/2}. \quad (3)$$

Расстояние в метрическом пространстве должно удовлетворять неравенству треугольника

$$\rho(\xi, \zeta) \leq \rho(\xi, \eta) + \rho(\eta, \zeta). \quad (4)$$

Проверим это неравенство для расстояний ρ_0, ρ_1, ρ_2 .

Пусть $\rho_0(\xi, \eta) = a, \rho_0(\eta, \zeta) = b$. Это означает, что $P(|\xi - \eta| > a) \leq a, P(|\eta - \zeta| > b) \leq b$. Тогда

$$\begin{aligned} P(|\xi - \zeta| > a + b) &= P(|(\xi - \eta) + (\eta - \zeta)| > a + b) \leq \\ &\leq P(|\xi - \eta| > a) + P(|\eta - \zeta| > b) \leq a + b, \end{aligned}$$

а это и означает, что $\rho_0(\xi, \zeta) \leq a + b$.

Поскольку $|\xi - \zeta| = |(\xi - \eta) + (\eta - \zeta)| \leq |\xi - \eta| + |\eta - \zeta|$, то по свойству математического ожидания $M|\xi - \zeta| \leq M|\xi - \eta| + M|\eta - \zeta|$, что означает неравенство треугольника для расстояния ρ_1 .

Обозначим $u = \xi - \eta, v = \eta - \zeta$. Тогда

$$\rho_2^2(\xi, \eta) = M(u + v)^2 = Mu^2 + Mv^2 + 2Muv.$$

Поскольку, вследствие неравенства Коши — Буняковского,

$$Muv \leq (Mu^2)^{1/2} (Mv^2)^{1/2},$$

то

$$Mu^2 + Mv^2 + 2Muv \leq ((Mu^2)^{1/2} + (Mv^2)^{1/2})^2 = (\rho_2(\xi, \eta) + \rho_2(\eta, \zeta))^2.$$

Для получения неравенства треугольника для ρ_2 достаточно извлечь квадратный корень.

Итак, все соотношения (1) — (3) могут быть приняты в качестве расстояния между ξ и η как функций элементарного события ω .

Пусть ξ_n, η — случайные величины. Если $\rho_0(\xi_n, \eta) \xrightarrow{n \rightarrow \infty} 0$, то последовательность (ξ_n) называется *сходящейся по вероятности* к случайной величине η . Сходимость по вероятности обозначается следующим образом:

$$\xi_n \xrightarrow{p} \eta, \quad \eta = p \lim_{n \rightarrow \infty} \xi_n.$$

Пусть L^1 — множество случайных величин с конечным математическим ожиданием. Пусть $\eta \in L^1$. Последовательность случайных величин (ξ_n) называется *сходящейся к η в среднем*, если

$$\rho_1(\xi_n, \eta) \xrightarrow{n \rightarrow \infty} 0.$$

Символическое обозначение

$$\xi_n \xrightarrow{(1)} \eta, \quad \lim_{n \rightarrow \infty} \xi_n = \eta \text{ (в ср.)}.$$

Пусть, наконец, η принадлежит множеству L^2 случайных величин с конечной дисперсией. Тогда если для последовательности (ξ_n) случайных величин $\rho_2(\xi_n, \eta) \xrightarrow{n \rightarrow \infty} 0$, то (ξ_n) называется *сходящейся к η в среднем квадратическом*, а η — *среднеквадратическим пределом* ξ_n при $n \rightarrow \infty$. Символическое обозначение

$$\xi_n \xrightarrow{(2)} \eta, \quad \lim_{n \rightarrow \infty} \xi_n = \eta \text{ (с. к.)}.$$

Объясним введенные понятия сходимости.

Сходимость по вероятности равносильна тому, что для любого $\varepsilon > 0$

$$P(|\xi_n - \eta| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0. \quad (5)$$

В самом деле, если выполнено (5), то, начиная с некоторого n , $P(|\xi_n - \eta| > \varepsilon) \leq \varepsilon$, т. е. $\rho_0(\xi_n, \eta) \leq \varepsilon$. Кроме того, если $\rho_0(\xi_n, \eta) \rightarrow 0$, то, начиная с некоторого n , $\rho_0(\xi_n, \eta) = \varepsilon_n < \varepsilon$, и тогда

$$P(|\xi_n - \eta| > \varepsilon) \leq P(|\xi_n - \eta| > \varepsilon_n) \leq \varepsilon_n \rightarrow 0,$$

что равносильно (5).

Закон больших чисел

$$P(|\bar{x}_n - a| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0,$$

где \bar{x}_n — среднее арифметическое n независимых случайных величин с математическим ожиданием a , можно теперь выразить так:

$$\bar{x}_n \xrightarrow{P} a.$$

Если $\xi_n \xrightarrow{(1)} \eta$, то $M\xi_n$ существует для достаточно большого n , так как вследствие неравенства треугольника

$$M\xi_n = \rho_1(\xi_n, 0) \leq \rho_1(\xi_n, \eta) + \rho_1(\eta, 0) = \rho_1(\xi_n, \eta) + M|\eta|.$$

Второе слагаемое конечно по условию, а первое стремится к нулю; следовательно, $M\xi_n$ конечно для достаточно большого n .

Аналогично доказывается, что если $\xi_n \rightarrow \eta$ (с. к.), то $D\xi_n < \infty$ для достаточно большого n .

Вследствие неравенства

$$P(|\xi_n - \eta| > \varepsilon) \leq M|\xi_n - \eta|/\varepsilon \quad (6)$$

из $\xi_n \xrightarrow{(1)} \eta$ следует, что $\xi_n \xrightarrow{P} \eta$. Поскольку вследствие неравенства Коши — Буняковского

$$(M|\xi_n - \eta|)^2 = (M(|\xi_n - \eta| \cdot 1))^2 \leq M(\xi_n - \eta)^2 \cdot M1^2 = \rho_2^2(\xi_n, \eta), \quad (7)$$

то из $\xi_n \xrightarrow{(2)} \eta$ следует, что $\xi_n \xrightarrow{(1)} \eta$. Итак, установлено логическое следование

$$\xi_n \xrightarrow{(2)} \eta \Rightarrow \xi_n \xrightarrow{(1)} \eta \Rightarrow \xi_n \xrightarrow{P} \eta. \quad (8)$$

Множество всех одномерных случайных величин с расстоянием $\rho_0(\xi, \eta)$ назовем *пространством* L^0 , множество случайных величин с конечным математическим ожиданием (дисперсией) с расстоянием $\rho_1(\xi, \eta)$ ($\rho_2(\xi, \eta)$) — *пространством* L^1 (L^2).

Важнейшим свойством пространств L^r , $r = 0, 1, 2$, является *полнота*: если последовательность (ξ_n) случайных величин из L^r фундаментальна, т. е. $\rho_r(\xi_n, \xi_m) \rightarrow 0$ при $n \rightarrow \infty$, $m \rightarrow \infty$, то существует случайная величина η из L^r , к которой сходится (ξ_n) :

$$\rho_r(\xi_n, \eta) \xrightarrow{n \rightarrow \infty} 0.$$

Ограничимся доказательством полноты L^0 . Выберем n_k из условия, что $\rho_0(\xi_n, \xi_m) < 2^{-k}$ при $n \geq n_k, m \geq n_k$, и положим $\eta_k = \xi_{n_k}$. Пусть A_k — множество тех ω , для которых $|\eta_i - \eta_j| \leq 2^{1-i}$ при любых $j > i \geq k$. При $\omega \in A_k$ числовая последовательность (η_k) фундаментальна, а значит, существует

$$\eta = \eta(\omega) = \lim_{k \rightarrow \infty} \eta_k.$$

Поскольку

$$\mathbf{P}(A_k) \geq 1 - \sum_{i=k}^{\infty} \mathbf{P}(|\eta_i - \eta_{i+1}| \geq 2^{-i}) \geq 1 - \sum_{i=k}^{\infty} \rho_0(\eta_i, \eta_{i+1}) \geq 1 - 2^{1-k},$$

то по аксиоме непрерывности $\mathbf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = 1$, т. е. $\eta = \lim_{n \rightarrow \infty} \eta_k$ существует с вероятностью 1.

При $\omega \in A_k$

$$|\xi_n - \eta| \leq |\xi_n - \eta_k| + |\eta_k - \eta| \leq |\xi_n - \eta_k| + 2^{1-k} \leq 2^{2-k}$$

при достаточно большом n . Следовательно, $\rho_0(\xi_n, \eta) \leq 2^{2-k}$, откуда $\rho_0(\xi_n, \eta) \rightarrow 0$, что и требовалось доказать.

То, что η — случайная величина, следует из соотношения

$$\{\xi < x\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{\eta_k < x\}.$$

Если ξ_n, η — случайные величины со значениями в \mathbb{R}^m , то понятия сходимости $\xi_n \xrightarrow{p} \eta$, $\xi_n \xrightarrow{(1)} \eta$, $\xi_n \xrightarrow{(2)} \eta$ вводятся как соответствующая сходимость для каждой из m компонент m -мерных случайных величин в отдельности.

Пусть $\zeta = \alpha + i\beta$ — комплексная случайная величина, α и β — одномерные действительные случайные величины. В этом случае чаще всего используется *среднеквадратическая сходимость*

$$\zeta_n \xrightarrow{(2)} \gamma.$$

Вместо проверки условий сходимости действительных и мнимых частей ζ_n в отдельности используют понятие *расстояния* в пространстве комплексных случайных величин

$$\rho(\zeta_1, \zeta_2) = \|\zeta_1 - \zeta_2\|,$$

где

$$\|\zeta\| = (\mathbf{M}|\zeta|^2)^{1/2} = (\mathbf{M}\zeta\bar{\zeta})^{1/2} \quad (9)$$

и в свою очередь $\bar{\zeta} = \alpha - i\beta$ — величина, комплексно сопряженная с ζ . Таким образом, $\zeta_n \xrightarrow{(2)} \gamma$ в том и только том случае, если

$$\|\zeta_n - \gamma\| = (\mathbf{M}|\zeta_n - \gamma|^2)^{1/2} \xrightarrow{n \rightarrow \infty} 0. \quad (10)$$

Поскольку

$$\mathbf{M}|\zeta|^2 = \mathbf{M}\zeta\bar{\zeta} = \mathbf{M}\alpha^2 + \mathbf{M}\beta^2, \quad (11)$$

то соотношение (10) и означает, что действительные и мнимые части ξ_n сходятся в среднеквадратическом смысле к соответствующим частям γ .

Если $\xi_n \xrightarrow{(2)} \xi$, то для $\gamma \in L^2$

$$M \xi_n \bar{\gamma} \xrightarrow{n \rightarrow \infty} M \xi \bar{\gamma}. \quad (12)$$

Действительно, квадрат разности левой и правой частей (12)

$$(M(\xi_n - \xi) \bar{\gamma})^2 \leq M |\xi_n - \xi|^2 M |\gamma|^2.$$

По условию первый сомножитель бесконечно мал, а второй конечен. Требуемое доказано.

Пусть $(P_n(A))$ — последовательность распределений в \mathbb{R}^m , $P_0(A)$ — распределение в \mathbb{R}^m . По определению $(P_n(A))$ слабо сходится к $(P(A))$, или символическое обозначение

$$P_n(A) \xrightarrow{w} P(A),$$

если

$$\int f(x) P_n(dx) \xrightarrow{n \rightarrow \infty} \int f(x) P(dx) \quad (13)$$

для любой непрерывной ограниченной функции $f(x)$, $x \in \mathbb{R}^m$.

Пусть распределениям $P_n(A)$ и $P(A)$ соответствуют функции распределения $F_n(x)$, $F(x)$, $x \in \mathbb{R}^m$. Тогда слабая сходимость может быть также определена условием

$$F_n(x) \xrightarrow{n \rightarrow \infty} F(x) \quad (14)$$

в любой точке x , в которой $F(x)$ непрерывна.

Определения (13) и (14) эквивалентны. Покажем, например, что из (13) вытекает (14).

Обозначим $\bar{\varepsilon} = (\varepsilon, \dots, \varepsilon)$, где $\varepsilon > 0$, и построим такие непрерывные функции $f(x)$, $g(x)$, что

$$I(x < x_0 - \bar{\varepsilon}) \leq g(x) \leq I(x < x_0) \leq f(x) \leq I(x < x_0 + \bar{\varepsilon}).$$

Тогда

$$\begin{aligned} F(x_0 - \bar{\varepsilon}) &\leq \int g dF = \lim \int g dF_n \leq \lim_{n \rightarrow \infty} F_n(x_0) \leq \\ &\leq \lim_{n \rightarrow \infty} F_n(x_0) \leq \lim_{n \rightarrow \infty} \int f dF_n = \int f dF \leq F(x_0 + \bar{\varepsilon}). \end{aligned}$$

Если x_0 — точка непрерывности $F(x)$, то $F(x \pm \bar{\varepsilon})$ сколь угодно мало отклоняются от $F(x_0)$, а следовательно, этим же свойством обладают и $\lim_{n \rightarrow \infty} F_n(x_0)$, $\lim_{n \rightarrow \infty} F_n(x_0)$. Отсюда $\lim_{n \rightarrow \infty} F_n(x_0) = F(x_0)$, что и требовалось доказать.

Если $\xi_n \xrightarrow{p} \xi$, то

$$F_{\xi_n}(x) \xrightarrow{w} F_{\xi}(x). \quad (15)$$

Действительно, из сходимости ξ_n и ξ по вероятности следует, что

$$P(|\xi_n^{(k)} - \xi^{(k)}| > \varepsilon) \rightarrow 0, \quad 1 \leq k \leq m,$$

где индексом k обозначены k -е компоненты ξ_n , ξ .

Поскольку

$$\begin{aligned} \{\xi < x - \bar{\varepsilon}; |\xi_n^{(k)} - \xi^{(k)}| \leq \varepsilon, 1 \leq k \leq m\} &\subset \{\xi_n < x\} \subset \\ &\subset \{\xi < x\} \cup \left\{ \bigcup_{k=1}^m \{|\xi_n^{(k)} - \xi^{(k)}| > \varepsilon\} \right\}, \end{aligned}$$

то

$$F(x - \bar{\varepsilon}) - \sum_{k=1}^m \rho_0(\xi_n^{(k)}, \xi^{(k)}) \leq F_n(x) \leq F(x + \bar{\varepsilon}) + \sum_{k=1}^m \rho_0(\xi_n^{(k)}, \xi^{(k)}).$$

Теперь достаточно заметить, что слагаемые сумм бесконечно малы при $n \rightarrow \infty$, а $F(x \pm \bar{\varepsilon})$ можно сделать сколь угодно близкими к $F(x)$, если x — точка непрерывности этой функции.

Пример 1. Пусть $F_n(x)$ — функция распределения случайной величины $(S_n - np)/\sqrt{npq}$, где $S_n = B(n, p)$, т. е. S_n — число успехов в n независимых испытаниях с вероятностью успеха p . Тогда из предельной теоремы Лапласа следует, что

$$F_n(x) \xrightarrow{w} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt. \quad (16)$$

Поскольку $\Phi(x)$ непрерывна в любой точке x , то значок « w » можно убрать: сходимость в обычном смысле имеет место для любого $x \in \mathbb{R}$.

Пример 2. Пусть $E(x) = 0$ при $x \leq 0$, $E(x) = 1$ при $x > 0$, т. е. $E(x)$ — функция распределения величины, равной 0 с вероятностью 1.

Рассмотрим функции $F_n(x) = \Phi(x/\sigma_n)$, где $\sigma_n \rightarrow 0$ при $n \rightarrow \infty$. При $x = 0$ $F_n(x) = 1/2$, в то время как $E(x) = 0$. При любых $x \neq 0$ $F_n(x) \rightarrow E(x)$ при $n \rightarrow \infty$ (рис. 25). Имеем слабую сходимость: $F_n(x) \xrightarrow{w} E(x)$.

Очевидно, соотношение $\rho_0(\bar{x}_n, a) \xrightarrow{n \rightarrow \infty} 0$ (закон больших чисел для \bar{x}_n) равносильно соотношению

$$F_{\bar{x}_n}^-(x) \xrightarrow{w} E(x - a). \quad (17)$$

Пример 3. В физической лаборатории многократно взвешивается один и тот же образец. Показания весов при k -м взвешивании дают погрешность $\alpha + \beta_k$, где α — постоянная погрешность (например, из-за наклона основания весов), β_k — независимые случайные величины с математическим ожиданием 0 и дисперсией σ^2 . Тогда среднее арифметическое \bar{x}_n равно $\alpha + \bar{\beta}_n$, где $\bar{\beta}_n = (\beta_1 + \dots + \beta_n)/n$.

Имеем $M\bar{\beta}_n = 0$, $D\bar{\beta}_n = \sigma^2/n$. Поскольку

$$M(\bar{x}_n - \alpha)^2 = \sigma^2/n \xrightarrow{n \rightarrow \infty} 0,$$

то по неравенству Чебышева

$$\bar{x}_n \xrightarrow{p} \alpha.$$

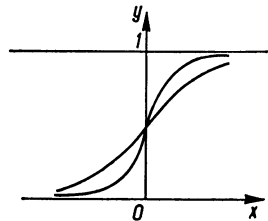


Рис. 25

§ 2. Теорема обращения. Теоремы непрерывности. Закон больших чисел

Рассмотрим вопрос о восстановлении распределения по характеристической функции.

Пусть одномерная случайная величина имеет плотность $p(x)$, функцию распределения $F(x)$ и характеристическую функцию $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx$. Тогда, если $\varphi(t)$ абсолютно интегрируема в интервале $(-\infty; \infty)$, то в точках, в которых $p(x)$ непрерывна, имеет место *формула обращения*

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt. \quad (1)$$

Проинтегрировав (1) в пределах от a до b , получим

$$\begin{aligned} F(b) - F(a) &= \int_a^b p(x) dx = \frac{1}{2\pi} \int_a^b dx \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_a^b e^{-itx} dx \right) \varphi(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) dt. \end{aligned} \quad (2)$$

Пусть теперь ξ — случайная величина с функцией распределения $F(x)$ и характеристической функцией $\varphi(t)$, а плотность существует не обязательно. Рассмотрим свертку функции $F(x)$ с распределением нормальной случайной величины η , $M\eta = 0$, $D\eta = \sigma^2$. Тогда характеристической функцией величины $\zeta = \xi + \eta$ будет $\varphi(t) e^{-\sigma^2 t^2/2}$, т. е. абсолютно интегрируемая функция. Величина ζ , как можно показать, имеет ограниченную плотность. По формуле (2)

$$F_{\zeta}(b) - F_{\zeta}(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) e^{-\sigma^2 t^2/2} dt.$$

Поскольку $\zeta \xrightarrow{p} \xi$ при $\sigma \rightarrow 0$, то $F_{\zeta}(x) \rightarrow F(x)$ в любой точке x непрерывности $F(x)$. Следовательно, если $F(x)$ непрерывна в точках a и b , то имеет место формула обращения

$$F(b) - F(a) = \lim_{\sigma \rightarrow 0} \int_{-\infty}^{\infty} \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) e^{-\sigma^2 t^2/2} dt. \quad (3)$$

Вместо η можно было бы взять любую другую величину, зависящую от некоторого параметра, сходящуюся к нулю по вероятности при стремлении параметра к предельному значению. Возьмем, например, случайную величину η с плотностью

$$p_{\eta}(x) = (1 - \cos(cx))/(\pi cx^2), \quad x \in \mathbb{R}, \quad c > 0.$$

Ее характеристическая функция имеет вид $1 - |t|/c$ при $|t| \leq c$ и равна 0 при $|t| > c$. В этом проще всего убедиться, вычислив ин-

теграл

$$\frac{1}{2\pi} \int_{-c}^c (1 - |t|/c) e^{-itx} dt.$$

Поскольку, очевидно, $\eta \xrightarrow{p} 0$ при $c \rightarrow \infty$, то подобно (3) получаем формулу обращения

$$F(b) - F(a) = \lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) (1 - |t|/c) dt. \quad (4)$$

Справедлива и более простая формула

$$F(b) - F(a) = \lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) dt. \quad (5)$$

Действительно, правые части (4) и (5) различаются выражением $\lim_{c \rightarrow \infty} (G(a, c) - G(b, c))$, где

$$\begin{aligned} G(a, c) &= \frac{1}{2\pi i} \int_{-c}^c e^{-ita} \varphi(t) (|t|/ct) dt = \\ &= \frac{1}{2\pi i} \int_{-c}^c e^{-ita} \left(\int e^{tz} dF(z) \right) (|t|/ct) dt = \\ &= \frac{1}{2\pi i} \int dF(z) \int_{-c}^c e^{it(z-a)} (|t|/ct) dt = \\ &= \frac{1}{\pi} \int dF(z) \left[\frac{1}{c} \int_0^c \sin t(z-a) dt \right]. \end{aligned}$$

Интеграл в квадратных скобках по модулю меньше или равен 1, откуда

$$\frac{1}{\pi} \int_{|z-a| < \delta} dF(z) [\dots] \leq \frac{1}{\pi} (F(a+\delta) - F(a-\delta));$$

это выражение сколь угодно мало при достаточно малом $\delta > 0$, так как $F(x)$ непрерывна в точке a . Подобный же интеграл по области $|z-a| \geq \delta$ бесконечно мал при $c \rightarrow \infty$, поскольку в этой области

$$|[\dots]| = \frac{1}{c} \left| \frac{1}{z-a} (\cos c(z-a) - 1) \right| \leq 2/c\delta.$$

Следовательно, $G(a, c)$, $G(b, c)$ бесконечно малы, так что (5) следует из (4).

При произвольном t выполняется следующая формула обращения. Вероятность попадания случайной величины в гиперпарал-

лелепипед имеет вид

$$\begin{aligned} P(a \leq \xi < b) &= P(a_i \leq \xi_i < b_i, \quad 1 \leq i \leq m) = \\ &= \frac{1}{(2\pi)^m} \lim_{c_1, \dots, c_m \rightarrow \infty} \int_{-c_1}^{c_1} dt_1 \int_{-c_2}^{c_2} dt_2 \dots \\ &\dots \int_{-c_m}^{c_m} \varphi(t_1, \dots, t_m) \prod_{k=1}^m \left(\frac{e^{-it_k b_k} - e^{-it_k a_k}}{-it_k} \right) dt_m, \end{aligned} \quad (6)$$

если a и b — точки непрерывности этой вероятности.

Справедливы важные *теоремы непрерывности для характеристических функций*. Пусть $F_n(t)$, $F(t)$ — функции распределения, $\varphi_n(t)$, $\varphi(t)$ — соответствующие им характеристические функции.

Прямая теорема непрерывности. Если $F_n \xrightarrow{w} F$, то $\varphi_n(t) \xrightarrow{n \rightarrow \infty} \varphi(t)$ равномерно на любом отрезке $|t| \leq T$.

Обратная теорема непрерывности. Если $\varphi_n(t) \xrightarrow{n \rightarrow \infty} \varphi(t)$ равномерно на любом отрезке $|t| \leq T$, то $F_n \xrightarrow{w} F$.

Прямая теорема непосредственно следует из определения слабой сходимости, так как $\varphi(t) = \int e^{it(x)} dF(x)$ — интеграл непрерывной ограниченной функции $e^{it(x)}$. Обратная теорема доказывается предельным переходом. Рассмотрим лишь случай $m = 1$: в общем случае — то же самое, но с более громоздкой символикой. Пусть ξ_n , ξ — случайные величины с распределениями $F_n(x)$, $F(x)$, η — случайная величина с характеристической функцией, равной $1 - |t|/c$ при $|t| \leq c$ и равной 0 при $|t| > c$. Тогда $\rho_0(\xi + \eta, \xi) = \rho_0(\eta, 0) \xrightarrow{c \rightarrow \infty} 0$. Это выражение от ξ не зависит; следовательно, при достаточно большом c $\rho_0(\xi + \eta, \xi) < \varepsilon$ и $\rho_0(\xi_n + \eta, \xi_n) < \varepsilon$ для всех n . Далее,

$$F_{\xi_n + \eta}(b) - F_{\xi_n + \eta}(a) = \frac{1}{2\pi} \int_{-c}^c \varphi_n(t) \frac{e^{-itb} - e^{-ita}}{-it} (1 - |t|/c) dt.$$

В силу равномерности сходимости $\varphi_n(t)$ к $\varphi(t)$ при $|t| \leq c$ записанный интеграл стремится при $n \rightarrow \infty$ к аналогичному интегралу с $\varphi(t)$ вместо $\varphi_n(t)$, т. е.

$$F_{\xi_n + \eta}(b) - F_{\xi_n + \eta}(a) \xrightarrow{n \rightarrow \infty} F_{\xi + \eta}(b) - F_{\xi + \eta}(a). \quad (7)$$

Положим $a = -A$, $b = A$, где A — число, для которого $F_{\xi + \eta}(A) - F_{\xi + \eta}(-A) > 1 - \varepsilon$. Тогда из (7) следует, что для достаточно большого n $F_{\xi_n + \eta}(A) - F_{\xi_n + \eta}(-A) > 1 - \varepsilon$, откуда $F_{\xi_n + \eta}(-A) < \varepsilon$, $F_{\xi + \eta}(-A) < \varepsilon$.

Пусть теперь x — любое число. Тогда

$$\begin{aligned} |F_{\xi_n + \eta}(x) - F_{\xi + \eta}(x)| &= |(F_{\xi_n + \eta}(x) - F_{\xi_n + \eta}(-A)) - (F_{\xi + \eta}(x) - \\ &- F_{\xi + \eta}(-A)) - (F_{\xi_n + \eta}(-A) - F_{\xi + \eta}(-A))|. \end{aligned}$$

Разность первых двух выражений в скобках при $n \rightarrow \infty$ стремится к нулю вследствие (7), а третье выражение меньше ε . Следова-

вательно, $\rho_0(\xi_n + \eta, \xi + \eta) \rightarrow 0$ при $n \rightarrow \infty$. Окончательно получаем

$$\rho_0(\xi_n, \xi) \leq \rho_0(\xi_n, \xi_n + \eta) + \rho_0(\xi_n + \eta, \xi + \eta) + \rho_0(\xi + \eta, \xi),$$

где все три слагаемые сколь угодно малы. Таким образом, $\rho_0(\xi_n, \xi) \rightarrow 0$ при $n \rightarrow \infty$, что и означает $F_n \xrightarrow{w} F$.

Пусть $f(x) = 1$ при $a \leq x < b$, $f(x) = 0$ в противном случае. Тогда преобразование Фурье функции $f(x)$ имеет вид

$$\tilde{f}(t) = \int_{-\infty}^{\infty} e^{-it(x)} f(x) dx = \prod_{k=1}^m \frac{e^{-it_k b_k} - e^{-it_k a_k}}{-it_k}; \quad (8)$$

именно это выражение находится под знаком интеграла в равенстве (6). Кроме того, $Mf(\xi) = P(a \leq \xi < b)$. Таким образом,

$$Mf(\xi) = \frac{1}{(2\pi^m)} \int_{-\infty}^{\infty} \varphi(t) \tilde{f}(t) dt. \quad (9)$$

Эта формула, называемая *равенством Парсеваля*, выполняется не только для $\tilde{f}(t)$ рассмотренного частного вида, но и для любой непрерывной ограниченной $f(x)$ с конечными $\int_{-\infty}^{\infty} |f(x)| dx, \int_{-\infty}^{\infty} |\tilde{f}(t)| dt$.

Аппарат характеристических функций значительно упрощает вывод различных предельных закономерностей.

Применим его к доказательству *закона больших чисел в форме Хинчина*: если (ξ_n) — последовательность независимых одинаково распределенных случайных величин с математическим ожиданием a , $\bar{x}_n = \frac{1}{n}(\xi_1 + \dots + \xi_n)$, то

$$\bar{x}_n \xrightarrow{p} a, \quad n \rightarrow \infty, \quad (10)$$

или, что то же самое, для любого $\varepsilon > 0$

$$P\left(\left|\frac{1}{n} \sum_{k=1}^n \xi_k - a\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0. \quad (11)$$

Для доказательства заметим, что соотношение (1) равносильно соотношению $\bar{x}_n \xrightarrow{p} 0$, где, как обычно, $\bar{x}_n^0 = \frac{1}{n}(\xi_1^0 + \dots + \xi_n^0)$.

Пусть $\psi(t)$ — характеристическая функция ξ_k^0 . Тогда, поскольку $\psi'(0) = iM\xi_k^0 = 0$, то $\psi(t) = 1 + \gamma(t)$, где $\gamma(t)/t \rightarrow 0$ при $t \rightarrow 0$. Характеристической функцией \bar{x}_n^0 будет $\varphi_n(t) = M \exp\{it \times \times (\xi_1^0 + \dots + \xi_n^0)/n\} = \psi^n(t/n)$. Следовательно,

$$\varphi_n(t) = \left(1 + \gamma\left(\frac{t}{n}\right)\right)^n \sim e^{n\gamma(t/n)} \xrightarrow{n \rightarrow \infty} 1$$

равномерно относительно t при $|t| \leq T$. Вспомнив, что

$$1 = \varphi(t) = M \exp\{it \cdot 0\},$$

на основании обратной теоремы непрерывности для характеристических функций убеждаемся в том, что функции распределения $F_n(x)$ случайных величин x_n^0 слабо сходятся при $n \rightarrow \infty$ к $E(x)$, где $E(x) = 0$ при $x \leq 0$, $E(x) = 1$ при $x \geq 0$. Поскольку $\pm \varepsilon$ — точки непрерывности $E(x)$, то

$$P(\bar{x}_n^0 < -\varepsilon) = F_n(-\varepsilon) \rightarrow E(-\varepsilon) = 0,$$

$$P(\bar{x}_n^0 \geq \varepsilon) = 1 - F_n(\varepsilon) \rightarrow 1 - E(\varepsilon) = 0.$$

Сочетание последних двух соотношений и приводит к (11).

Замечание. В доказательстве использовался лишь факт существования $\psi'(0)$. Поэтому закон больших чисел в форме (10), (11) выполняется для любой последовательности независимых одинаково распределенных случайных величин (ξ_n), для которых

$$\varphi'_{\xi_k}(0) = a.$$

(Оказывается, что при этом $M\xi_k$ может не существовать.)

§ 3. Усиленный закон больших чисел

Вначале установим важное *неравенство Колмогорова* для независимых случайных величин ξ_k с конечными дисперсиями σ_k^2 . Возьмем центрированные величины $\bar{\xi}_k^0 = \xi_k - M\xi_k$ и образуем суммы $s_k^0 = \bar{\xi}_1^0 + \dots + \bar{\xi}_k^0$. Тогда

$$P(|s_k^0| < \varepsilon, \quad 1 \leq k \leq n) \geq 1 - (\sigma_1^2 + \dots + \sigma_n^2)/\varepsilon^2 \quad (1)$$

или, что то же самое,

$$P(\max_{1 \leq k \leq n} |s_k^0| \geq \varepsilon) \leq (\sigma_1^2 + \dots + \sigma_n^2)/\varepsilon^2. \quad (2)$$

Прежде чем доказывать неравенство Колмогорова, проиллюстрируем его применение.

Пример 1. Измерительное устройство используется для измерения сигналов, поступающих в моменты $\Delta, 2\Delta, \dots, n\Delta$. Параметр $s(t)$, характеризующий точность устройства в момент t , изменяется случайным образом, причем так, что $s(0) = 0$, $\xi_k = s(k\Delta) - s((k-1)\Delta)$ — независимые случайные величины; $M\xi_k = 0$, $D\xi_k = \sigma^2$. Надежность устройства характеризуется вероятностью того, что $|s(k\Delta)| < \varepsilon$ при всех k , $1 \leq k \leq n$. Тогда из (1) получаем оценку: надежность больше или равна $1 - n\sigma^2/\varepsilon^2$.

Пример 2. Параметр $s(t)$, описываемый тем же законом, что и в примере 1, характеризует отклонение режима установки от номинального. Если $|s(k\Delta)| \geq \varepsilon$, происходит авария. Для предотвращения аварии значения $s(k\Delta)$ контролируются при k , кратных n . Применяются так называемые *упреждающие допуски*: если в момент контроля $|s(k\Delta)| \geq \delta$, установка останавливается и $s(k\Delta)$ регулируется «в номинал», т. е. новое значение этой переменной равно 0 (для простоты полагаем время регулировки равным 0). Какое значение δ следует взять, чтобы $P(|s(k\Delta)| \leq \varepsilon, \quad 1 \leq k \leq nr) \geq 1 - \gamma$?

Пусть в момент $mn\Delta$ контроля $s(mn\Delta) = z$. При этом условии

$$\begin{aligned} P(\max_{mn\Delta < k \leq (m+1)n\Delta} |s(k\Delta)| \geq \varepsilon) &\leq P(\max_{mn\Delta < k} |s(k\Delta) - s(mn\Delta)| \geq \\ &\geq \varepsilon - \delta) \leq n\sigma^2/(\varepsilon - \delta)^2. \end{aligned}$$

Выход за допуски $\pm \varepsilon$ может произойти в r интервалах между контролями,

Вероятность отсутствия аварии удовлетворяет неравенству

$$P \geq (1 - n\sigma^2/(\varepsilon - \delta)^2)^r.$$

Отсюда получаем для значений δ , удовлетворяющих условию задачи, неравенство

$$(\varepsilon - \delta)^2 \geq n\sigma^2/(1 - (1 - \gamma)^{1/r})^1.$$

Для доказательства неравенства Колмогорова обозначим через I_k индикатор события $\{|s_k^0| \geq \varepsilon; |s_i^0| < \varepsilon, 1 \leq i < k\}$. Тогда $\sum_{k=1}^n I_k$ есть индикатор события $\{\max_{1 \leq k \leq n} |s_k^0| \geq \varepsilon\}$. Имеем

$$M(s_n^0)^2 \geq M(s_n^0)^2 \sum_{k=1}^n I_k = \sum_{k=1}^n M(s_n^0)^2 I_k.$$

Рассмотрим k -е слагаемое, обозначив

$$s_n^0 = s_k^0 + t_{n-k}^0,$$

где $t_{n-k}^0 = \xi_{k+1}^0 + \dots + \xi_n^0$ — величина, не зависящая от $(\xi_1^0, \dots, \xi_k^0)$:

$$M(s_n^0)^2 I_k = M(s_k^0 + t_{n-k}^0)^2 I_k = M(s_k^0)^2 I_k + 2M(s_k^0 I_k) t_{n-k}^0 + M(t_{n-k}^0)^2 \times \\ \times I_k \geq \varepsilon^2 M I_k,$$

поскольку: а) $(s_k^0)^2 I_k \geq \varepsilon^2 I_k$; б) удвоенная сумма равна 0 вследствие независимости $s_k^0 I_k$ как функции $(\xi_1^0, \dots, \xi_k^0)$ от t_{n-k}^0 ; в) последнее слагаемое неотрицательно. Просуммировав по k , получим

$$\sigma_1^2 + \dots + \sigma_n^2 = M(s_n^0)^2 \geq \varepsilon^2 \sum_{k=1}^n M I_k = \varepsilon^2 M \sum_{k=1}^n I_k = \\ = \varepsilon^2 P(\max_{1 \leq k \leq n} |s_k^0| \geq \varepsilon),$$

что и представляет собой неравенство Колмогорова.

Используем это неравенство для доказательства *усиленного закона больших чисел (теорема Колмогорова)*. Пусть ξ_k , $k \geq 1$, — независимые одномерные случайные величины с одним и тем же распределением $F(x)$, $M\xi_k = a$;

$$s_n = \xi_1 + \dots + \xi_n, \quad n \geq 1.$$

Тогда

$$P(s_n/n \xrightarrow{n \rightarrow \infty} a) = 1. \quad (3)$$

Поскольку $s_n/n \rightarrow a$ равносильно соотношению $s_n^0/n \rightarrow 0$, то можно без ограничения общности считать, что $a = 0$. Положим $\eta_n = \xi_n$ при $|\xi_n| \leq n$, $\eta_n = 0$ в противном случае и рассмотрим суммы $t_n = \eta_1^0 + \dots + \eta_n^0$, $n \geq 1$. Имеем оценку

$$D\eta_n \leq M\eta_n^2 = \int_{-n}^n x^2 dF(x) \leq \sum_{i=1}^n i^2 p_i, \quad (4)$$

¹ Данная оценка весьма грубая, так как фактически не учитывает регуляторов «в номинал».

где

$$p_i = \int_{i-1 < |x| \leq i} dF(x).$$

Заметим сразу, что

$$\sum_{i=1}^{\infty} i p_i \leq \int |x| dF(x) + 1 < \infty. \quad (5)$$

Пусть A_r — событие, состоящее в том, что $|t_n/n| \geq \varepsilon$ хотя бы для одного n , $2^{r-1} \leq n \leq 2^r$. По неравенству Колмогорова

$$P(A_n) \leq P\left(\bigcup_{n=2^{r-1}}^{2^r} \{|t_n/2^{r-1}| \geq \varepsilon\}\right) \leq 4^{1-r} D t_{2^r} / \varepsilon^2. \quad (6)$$

Предположим, что

$$\sum_{r=1}^{\infty} 4^{-r} D t_{2^r} < \infty. \quad (7)$$

Тогда из (6) вытекает, что $P\left(\bigcup_{n=N}^{\infty} A_n\right) \xrightarrow{N \rightarrow \infty} 0$, а следовательно, по аксиоме непрерывности, начиная с некоторого номера n_0 , все $|t_n/n| < \varepsilon$, т. е. $t_n/n \rightarrow 0$ с вероятностью 1. Остается оценить левую часть (7). Вследствие (4)

$$D t_n = \sum_{k=1}^n \sum_{i=1}^k i^2 p_i \leq n \sum_{i=1}^n i^2 p_i,$$

откуда

$$\begin{aligned} \sum_{r=1}^{\infty} 4^{-r} \cdot 2^r \sum_{i=1}^{2^r} i^2 p_i &= \sum_{i=1}^{\infty} i^2 p_i \sum_{2^{-r} \geq i} 2^{-r} \leq \sum_{i=1}^{\infty} i^2 p_i \cdot 2/i = \\ &= 2 \sum_{i=1}^{\infty} i p_i < \infty \end{aligned}$$

(последнее вследствие (5)). Итак, $t_n/n \rightarrow 0$ с вероятностью 1.

Можно записать

$$\frac{1}{n} s_n = \frac{1}{n} t_n + \frac{1}{n} \sum_{k=1}^n M \eta_k + \frac{1}{n} \sum_{k=1}^n (\xi_k - \eta_k). \quad (8)$$

Первое слагаемое правой части (8) стремится к нулю с вероятностью 1. Второе слагаемое имеет вид

$$\frac{1}{n} \sum_{k=1}^n \int_{|x| \geq k} x dF(x) \leq \frac{1}{n} \sum_{k=1}^n \sum_{i=k+1}^{\infty} i p_i \xrightarrow{n \rightarrow \infty} 0;$$

если k -й член последовательности $\left(\sum_{i=k+1}^{\infty} i p_i\right)$ стремится к 0, то и среднее арифметическое n членов этой последовательности стремится к 0 при $n \rightarrow \infty$.

Переходя к оценке последней суммы в правой части (8), введем событие $B_k = \{\xi_k - \eta_k \neq 0\}$. Тогда

$$\mathbf{P}(B_k) = \int_{|x| \geq k} dF(x) \leq \sum_{i=k+1}^{\infty} p_i.$$

Следовательно,

$$\sum_{k=1}^{\infty} \mathbf{P}(B_k) \leq \sum_{k=1}^{\infty} \sum_{i=k+1}^{\infty} p_i = \sum_{i=2}^{\infty} p_i \sum_{k=1}^{i-1} 1 = \sum_{i=2}^{\infty} (i-1) p_i < \infty,$$

откуда

$$\mathbf{P}\left(\bigcup_{k=N}^{\infty} B_k\right) \leq \sum_{k=N}^{\infty} \mathbf{P}(B_k) \xrightarrow{N \rightarrow \infty} 0,$$

т. е. по аксиоме непрерывности найдется такое N , что ни одно из событий B_N, B_{N+1}, \dots не наступит. Это означает, что с вероятностью 1 лишь конечное число разностей $\xi_k - \eta_k$ не равны нулю, а следовательно, последняя сумма в правой части (8) стремится к нулю при $n \rightarrow \infty$. Итак, $s_n/n \rightarrow 0$. Теорема Колмогорова доказана.

Докажем следующее важное предложение. Пусть (ξ_n) — последовательность независимых действительных или комплексных случайных величин, $a_n = \mathbf{M}\xi_n$, $\sigma_n^2 = \mathbf{D}\xi_n$ в действительном случае, $\sigma_n^2 = \mathbf{M}|\xi_n^0|^2$ в комплексном случае. Тогда, если

$$\sum_{n=1}^{\infty} a_n = A, \quad \sum_{n=1}^{\infty} \sigma_n^2 = B^2,$$

то с вероятностью 1

$$\sum_{n=1}^{\infty} \xi_n = \eta, \quad (9)$$

где η — случайная величина, $\mathbf{M}\eta = A$, $\mathbf{M}|\eta^0|^2 = B^2$, причем.

$$\mathbf{M}\left|\sum_{n=1}^N \xi_n - \eta\right|^2 \xrightarrow{N \rightarrow \infty} 0. \quad (10)$$

То, что $\sum_{n=1}^N \xi_n \xrightarrow[N \rightarrow \infty]{(2)} \eta$, очевидно, поскольку последовательность частных сумм ряда фундаментальна:

$$\mathbf{M}\left|\sum_{n=N}^M \xi_n\right|^2 = \sum_{n=N}^M |a_n + \xi_n^0|^2 = \left|\sum_{n=N}^M a_n\right|^2 + \sum_{n=N}^M \mathbf{M}|\xi_n^0|^2 \xrightarrow{n \rightarrow \infty} 0.$$

Пусть вначале ξ_n — действительные случайные величины. Вследствие неравенства Колмогорова

$$\mathbf{P}\left(\left|\sum_{n=N}^r \xi_n^0\right| < \varepsilon, N \leq r \leq N'\right) \geq 1 - \sum_{n=N}^{N'} \sigma_n^2/\varepsilon^2 \geq 1 - \sum_{n=N}^{\infty} \sigma_n^2/\varepsilon^2.$$

По аксиоме непрерывности

$$\mathbf{P}\left(\left|\sum_{n=N}^r \xi_n^0\right| < \varepsilon, r \geq N\right) \geq 1 - \sum_{n=N}^{\infty} \sigma_n^2/\varepsilon^2,$$

а тогда снова по этой аксиоме с вероятностью 1 ряд $\sum \xi_n^0$ сходится.

Вместе с ним сходится и ряд $\sum \xi_n$, так как ξ_n^0 и ξ_n различаются членами сходящегося ряда. Сумму этого ряда обозначим через η' . Остается выяснить, совпадает ли η' со среднеквадратическим пределом η .

Поскольку $P\left(\left|\sum_{n=1}^N \xi_n - \eta'\right| > \varepsilon\right) \rightarrow 0$, то для некоторого достаточно большого N

$$P\left(\left|\sum_{n=1}^N \xi_n - \eta'\right| > \varepsilon\right) < \varepsilon. \quad (11)$$

В то же время по аксиоме непрерывности

$$P\left(\left|\sum_{n=1}^N \xi_n - \eta\right| > \left|\sum_{n=N+1}^{\infty} a_n\right| + \varepsilon\right) \leq \sum_{n=N+1}^{\infty} \sigma_n^2.$$

Взяв несколько большее N , чтобы $\left|\sum_{n=N+1}^{\infty} a_n\right| < \varepsilon$, $\sum_{n=N+1}^{\infty} \sigma_n^2 < \varepsilon$, получим

$$P\left(\left|\sum_{n=1}^N \xi_n - \eta\right| > 2\varepsilon\right) < \varepsilon. \quad (12)$$

Из (11) и (12) вытекает, что

$$P(|\eta' - \eta| > 3\varepsilon) < 2\varepsilon.$$

Поскольку $\varepsilon > 0$ произвольно, то отсюда вытекает, что $\eta' = \eta$ с вероятностью 1, а тогда имеет место и равенство (10).

Доказательство в комплексном случае легко сводится к доказательству отдельно для действительной и мнимой частей.

§ 4. Стационарные и эргодические последовательности

Пусть (ξ_n) — последовательность случайных величин, заданных на одном и том же вероятностном пространстве (Ω, \mathcal{F}, P) . Эта последовательность называется *стационарной* (более подробно: *стационарной в узком смысле случайной последовательностью*), если при любом n распределение случайной величины $(\xi_{k+1}, \dots, \xi_{k+n})$ не зависит от k .

Например, (η_n) — бернуллиева последовательность, $\xi_n = f(\eta_n, \eta_{n+1}, \dots, \eta_{n+m-1})$. Тогда (ξ_1, \dots, ξ_n) определяется исходами испытаний с номерами от 1 до $n + m - 1$, $(\xi_{k+1}, \dots, \xi_{k+n})$ — соответствующими исходами испытаний с номерами, измененными на k ; ясно, что вероятности не зависят от сдвига номеров испытаний.

Из стационарности следует, что если $f((\xi_n))$ — случайная величина, определенная последовательностью (ξ_n) , то

$$Mf((\xi_n)) = Mf((\xi_{n+k}));$$

более точно, если одно из этих чисел существует, то существует и другое, и оба совпадают.

Обычно последовательность (ξ_n) определяется при $n \in \mathbb{Z}$.

Фундаментальным результатом теории стационарных случайных последовательностей является следующая *эргодическая теорема Биркгофа — Хинчина*.

Пусть (ξ_n) — стационарная в узком смысле последовательность одномерных случайных величин. Тогда если существует $M\xi_0$, то последовательность средних арифметических

$$\bar{x}_n = \frac{1}{n} (\xi_0 + \dots + \xi_{n-1})$$

при $n \rightarrow \infty$ сходится с вероятностью 1 к некоторой случайной величине η :

$$P(\bar{x}_n \rightarrow \eta) = 1. \quad (1)$$

Поскольку последовательность $\xi_n = f((\xi_{n+j}, j \in \mathbb{Z}))$, где ξ_n — одномерные случайные величины, сама стационарна, то при условии $|M\xi_0| < \infty$ теорема Биркгофа — Хинчина относится и к ней:

$$P\left(\frac{1}{n} (\xi_0 + \dots + \xi_{n-1}) \rightarrow \gamma\right) = 1, \quad (2)$$

где γ — некоторая случайная величина.

Полное доказательство имеется в учебнике Б. В. Гнеденко «Курс теории вероятностей» (М., 1988). Приведем лишь основные моменты доказательства.

Пусть K — событие, состоящее в том, что предел (\bar{x}_n) не существует. Требуется доказать, что $P(K) = 0$.

Имеем

$$P(K) \leq \sum_{\alpha < \beta} P(\lim_{n \rightarrow \infty} \bar{x}_n < \alpha, \overline{\lim}_{n \rightarrow \infty} \bar{x}_n > \beta),$$

где α, β — всевозможные рациональные числа. Поскольку множество рациональных чисел счетно, то, для того чтобы сумма была равна 0, достаточно, чтобы каждое слагаемое было равно 0. Итак, если $K_{\alpha\beta}$ — событие, состоящее в том, что

$$\lim_{n \rightarrow \infty} \bar{x}_n < \alpha, \quad \overline{\lim}_{n \rightarrow \infty} \bar{x}_n > \beta, \quad (3)$$

где $\alpha < \beta$ — фиксированные числа, то нам достаточно показать, что $P(K_{\alpha\beta}) = 0$. Предположим противное: $P(K_{\alpha\beta}) > 0$, и будем рассматривать все вероятности при условии $K_{\alpha\beta}$. Тогда с вероятностью 1 последовательность (\bar{x}_n) бесконечно много раз попадает выше уровня β и ниже уровня α .

В частности, найдутся такие $a \leq 0 < b$, что $(\xi_a + \xi_{a+1} + \dots + \xi_{b-1}) / (b - a) > \beta$. Пусть $\xi_a, \xi_{a+1}, \dots, \xi_{b-1}$ приняли фиксированные значения x_0, x_1, \dots, x_{b-a} . Скажем, что точка 0 покрывается отрезком $[x_0; x_1; \dots; x_{b-a}]$. Каким дополнительным свойствам должны удовлетворять подобные отрезки (могут ли они пересекаться, включать один в другой и т. д.), этот вопрос мы опускаем. Заметим лишь следующее. Если точка 0 покрывается упомянутым отрез-

ком, то в силу стационарности последовательности одинаково вероятны следующие возможности:

$$\begin{aligned}\xi_0 &= x_0, & \xi_1 &= x_1, & \dots, & \xi_{b-a-1} &= x_{b-a-1}; \\ \xi_{-1} &= x_0, & \xi_0 &= x_1, & \dots, & \xi_{b-a-2} &= x_{b-a-1}; \\ & \dots & & \dots & & & \dots \\ \xi_{-(b-a-1)} &= x_0, & \xi_{-(b-a-2)} &= x_1, & \dots, & \xi_0 &= x_{b-a-1}.\end{aligned}$$

Таким образом, условное математическое ожидание ξ_0 равно $(x_0 + x_1 + \dots + x_{b-a-1})/(b-a) > \beta$. Усреднив по всем возможным (x_i) , получим $M(\xi_0 | K_{\alpha\beta}) > \beta$. Однако точно так же доказывается и соотношение $M(\xi_0 | K_{\alpha\beta}) < \alpha$. Поскольку $\alpha < \beta$, то мы пришли к противоречию, что и доказывает эргодическую теорему Биркгофа — Хинчина.

Из стационарных последовательностей (ξ_n) наибольшее значение имеют эргодические последовательности. Последовательность (ξ_n) называется *эргодической* относительно функции $f(\dots, x_{-1}, x_0, x_1, \dots)$, если $\zeta_n = f(\dots, \xi_{n-1}, \xi_n, \xi_{n+1}, \dots)$ ¹ — случайные величины с математическим ожиданием a и с вероятностью 1

$$\lim_{n \rightarrow \infty} \frac{1}{n} (\zeta_0 + \zeta_1 + \dots + \zeta_{n-1}) = a. \quad (4)$$

Очевидно, если ζ_n — случайные величины (кстати, для этого достаточно, чтобы ζ_0 была случайной величиной), то, очевидно, (ζ_n) — стационарная последовательность, и по теореме Биркгофа — Хинчина средние арифметические сходятся с вероятностью 1 к некоторой случайной величине η . Для доказательства того, что $\eta = a$, достаточно показать, что

$$D\left(\frac{1}{n} (\zeta_0 + \dots + \zeta_{n-1})\right) \rightarrow 0, \quad (5)$$

если ζ_n — одномерные величины, и соответствующие равенства для компонент ζ_n , если они многомерные; этот критерий большей частью и применяется.

Назовем *корреляционной функцией* последовательности (ζ_n) одномерных случайных величин функцию

$$R(n) = \text{cov}(\zeta_k, \zeta_{n+k}), \quad n \in \mathbb{Z}. \quad (6)$$

Отметим, что $|R(n)| \leq R(0) = D\zeta_0$ и $R(n) = R(-n)$. Имеем

$$\begin{aligned}D\left(\frac{1}{n} (\zeta_0 + \dots + \zeta_{n-1})\right) &= \frac{1}{n^2} M(\zeta_0^2 + \dots + \zeta_{n-1}^2) = \\ &= \frac{1}{n^2} \sum_{k,j=0}^{n-1} M\zeta_k^0 \zeta_j^0 = \frac{1}{n^2} \sum_{k,j=0}^{n-1} R(k-j) = \\ &= \frac{1}{n^2} \sum_{m=-(n-1)}^{n-1} (n - |m|) R(m),\end{aligned}$$

¹ Имеется в виду, что вместо каждой переменной x_k подставлена ξ_{n+k} .

или

$$\begin{aligned} D\left(\frac{1}{n}(\xi_0 + \dots + \xi_{n-1})\right) &= \frac{1}{n} \sum_{m=-(n-1)}^{n-1} (1 - |m|/n) R(m) = \\ &= \frac{1}{n} \left(R(0) + 2 \sum_{m=1}^{n-1} (1 - m/n) R(m) \right). \end{aligned} \quad (7)$$

Таким образом, если правая часть (7) стремится к 0 при $n \rightarrow \infty$, то (ξ_n) эргодична относительно функции f . Это условие выполняется, в частности, при $R(m) \xrightarrow{m \rightarrow \infty} 0$. В этом случае $|R(m)| < \varepsilon$ при $m \geq N$, и тогда правая часть (7) по абсолютной величине меньше

$$\frac{1+2N}{n} R(0) + 2\varepsilon < 3\varepsilon, \quad n > (1+2N) R(0).$$

Эргодичность стационарной последовательности имеет глубокий смысл. Допустим, что нужно определить из опыта значение $M\xi_0 = a$. Имеем $\xi_0 = \xi_0(\omega)$, где ω — элементарное событие. Повторяя опыт в независимых условиях, наблюдаем независимые значения $\xi_0^{(k)} = \xi_0(\omega_k)$. По усиленному закону больших чисел с вероятностью 1

$$\frac{1}{n} (\xi_0^{(1)} + \dots + \xi_0^{(n)}) \rightarrow a, \quad (8)$$

т. е. в принципе можно вычислить искомое число a путем усреднения по множеству наблюдений (реализаций) случайной величины ξ_0 при различных значениях ω . Если последовательность эргодична, то этот же результат получится при усреднении по времени, т. е. взятии средних арифметических ξ_0, \dots, ξ_{n-1} при одном и том же значении ω . Указанное свойство называют равенством *среднего по множеству реализаций и среднего по времени (эргодического среднего)*.

Пример. В канале радиоприемного устройства наблюдаются независимые значения $\xi(n\Delta)$ амплитуды в моменты $n\Delta$, $n \in \mathbb{Z}$. Требуется вычислить вероятность события $\{\xi(0) + \xi(\Delta) + \dots + \xi((m-1)\Delta) \geq z\}$, называемую вероятностью ложной тревоги.

Обозначим через ξ_n индикатор ложной тревоги в момент $n\Delta$. Поскольку $\xi(n\Delta)$ независимы, то ξ_n и ξ_k некоррелированы при $|n-k| \geq m$, т. е. $R(n) = 0$ при $n \geq m$. Следовательно, существует среднее по времени, т. е. вероятность ложной тревоги можно найти по наблюдению последовательности (ξ_n) .

Задачи

1. Пусть $F_n(x)$, $F(x)$ — одномерные распределения. Показать, что $F_n(x) \xrightarrow[n \rightarrow \infty]{w} F(x)$ в том и только том случае, если для произвольного $\Delta > 0$

$$\int_{x_0-\Delta}^{x_0+\Delta} F_n(x) dx \xrightarrow{n \rightarrow \infty} \int_{x_0-\Delta}^{x_0+\Delta} F(x) dx$$

при любых x_0 .

2. С помощью формулы обращения вывести формулу для интеграла функции распределения $F(x_1, \dots, x_m)$ по параллелепипеду $a_1 < x_1 < b_1, \dots, a_m < x_m < b_m$.

3. Пусть s_n — число успехов в первых n испытаниях по схеме Бернулли. Назовем s_n рекордом, если $s_n > s_k$ при всех $k < n$. Обозначим через m число рекордов до момента n включительно. Доказать, что m_n/n сходится по вероятности к нулю при $n \rightarrow \infty$.

4. Пусть $X(t)$ — процесс Пуассона с параметром λ . Доказать, что с вероятностью 1

$$\frac{1}{t} X(t) \xrightarrow{t \rightarrow \infty} \lambda.$$

5. Пусть характеристическая функция $\varphi(t)$ m -мерной случайной величины (ξ_1, \dots, ξ_m) дважды дифференцируема в точке $t = 0$. Доказать, что в этом случае $D\xi_k < \infty$, $1 \leq k \leq m$.

6. Пусть $f(z) = Me^{z\xi}$, где ξ — одномерная случайная величина с плотностью $p(x)$. Доказать, что если $p(x)$ убывает, как экспонента, при $|x| \rightarrow \infty$, то $f(z)$ существует при $|z| < \delta$, $\delta > 0$. Показать, что обратное не верно.

7. Пусть $F_n(x)$, $F(x)$ — функции распределения одномерных случайных величин ξ_n , ξ ; $Me^{z\xi_n} \xrightarrow{n \rightarrow \infty} Me^{z\xi}$ для любых $z = 1, 2, \dots$, причем известно, что $Me^{\pm e\xi}$ существуют при некотором $e > 0$. Доказать, что в этом случае $F_n \xrightarrow{w} F$.

8. Пусть ξ_n — бернуллиева последовательность; $P(\xi_n = 1) = P(\xi_n = 0) = 1/2$. Назовем «словом» цепочку двоичных символов $\bar{e} = (e_1, \dots, e_m)$ и скажем, что данное слово появилось в бернуллиевой последовательности в момент t , если $\xi_{n-m+k} = e_k$, $1 \leq k \leq m$. Обозначим через $s_N(\bar{e})$ число появлений слова \bar{e} за время N , т. е. в моменты $t \leq N$. Доказать следующее утверждение: если переменные $m \geq 1$ и N изменяются так, что $2^{-m}N \rightarrow \infty$, то

$$2^m s_N(\bar{e}) N^{-1} \xrightarrow{p} 1.$$

У к а з а н и е. Применить неравенство Чебышева.

одномерная

Глава 6

ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

§ 1. Асимптотическая нормальность

Пусть (η_n) — последовательность случайных величин со значениями в \mathbb{R}^m , ξ — m -мерная невырожденная случайная величина с математическим ожиданием $a = (a_1, \dots, a_m)$ и корреляционной матрицей $\Sigma = (\text{cov}(\xi_i, \xi_j))$. Случайная величина η_n называется *асимптотически нормальной* с параметрами a , Σ , если для любого $x \in \mathbb{R}^m$

$$F_{\eta_n}(x) \xrightarrow{n \rightarrow \infty} F_{\xi}(x). \quad (1)$$

В частности, при $m = 1$ случайная величина η_n — общий член последовательности (η_n) — называется асимптотически нормальной с параметрами a , σ^2 , если при любых x

$$F_{\eta_n}(x) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-(u-a)^2/2\sigma^2} du = \Phi((x-a)/\sigma). \quad (2)$$

Так, пусть s_n — число успехов в n независимых испытаниях с вероятностью p успеха в каждом, $\eta_n = (s_n - np)/\sqrt{npq}$. Тогда

из интегральной теоремы Лапласа (см. § 2.7) вытекает, что η_n асимптотически нормальна с параметрами 0, 1.

Свойство асимптотической нормальности символически записывают следующим образом:

$$\eta_n \sim N(a, \Sigma), \quad \eta_n \sim N(a, \sigma^2).$$

Подчеркнем, что асимптотическая нормальность — свойство не индивидуальной случайной величины η_n , а члена последовательности (η_n) . Если же $F_{\eta}(x) \approx F_{\xi}(x)$, то говорят, что η приблизительно нормальна. Так, в определенном практическом расчете можно вместо распределения η_{100} использовать нормальное распределение $\Phi(x)$, т. е. в пределах допустимой погрешности вычислений $P(\eta_{100} < x) \approx \Phi(x)$.

Важно заметить, что сходимость в (1), (2) равномерна по x . В случае $m = 1$ это следует из того, что в точках x_k , в которых $F_{\xi}(x_k) = k/N$ ($1 \leq k \leq N - 1$), значения $F_{\eta_n}(x)$ стремятся к $F_{\xi}(x)$, а в интервалах между ними обе функции монотонны, так что в пределе отклонение не больше $1/N$. Ввиду произвольности N заключаем, что $\max_x |F_{\eta_n}(x) - F_{\xi}(x)| \xrightarrow{n \rightarrow \infty} 0$. При произвольном m доказательство отличается тем, что вместо отрезков $(x_k; x_{k+1})$ будут гиперкубы пространства \mathbb{R}^m .

Пусть при любом n имеются случайная величина s_n и невырожденная нормальная случайная величина ξ_n с математическим ожиданием a_n и корреляционной матрицей Σ_n . Тогда s_n называется *асимптотически нормальной* с параметрами a_n, Σ_n , если

$$\sup_{x \in \mathbb{R}^m} |F_{s_n}(x) - F_{\xi_n}(x)| \xrightarrow{n \rightarrow \infty} 0. \quad (3)$$

В частности, при $m = 1$ асимптотическая нормальность с параметрами a_n, σ_n^2 означает, что

$$\sup_{-\infty < x < \infty} |F_{\xi_n}(x) - \Phi((x - a_n)/\sigma_n)| \xrightarrow{n \rightarrow \infty} 0. \quad (4)$$

Символическое обозначение (3), (4):

$$s_n \sim N(a_n, \Sigma_n), \quad s_n \sim N(a_n, \sigma_n^2), \quad n \rightarrow \infty. \quad (5)$$

Так, для величины s_n , связанной с числом успехов, имеем

$$s_n \sim N(np, npq), \quad s_n - np \sim N(0, npq), \\ s_n/n \sim N(p, pq/n), \quad (s_n - np)/n \sim N(0, pq/n).$$

Асимптотическая нормальность играет большую роль в теории и практике. Она позволяет заменять сложные и часто неопределенные распределения случайных величин хорошо изученным нормальным распределением. Основным аппаратом доказательства асимптотической нормальности является аппарат характеристических функций. Из обратной теоремы непрерывности (см. гл. 5) получаем следующий вывод. Пусть $\varphi_n(t) = \varphi_n(t_1, \dots, t_m)$ — характеристическая функция m -мерной случайной величины η_n . Для того чтобы

η_n была асимптотически нормальной с параметрами a, Σ , где $a = (a_1, \dots, a_m)$, $\Sigma = (\sigma_{ij}^2)$ — положительно определенная матрица, необходимо и достаточно, чтобы

$$\varphi_n(t) \xrightarrow{n \rightarrow \infty} e^{it(a) - \frac{1}{2} t' \Sigma t} = \exp \left\{ i \sum_{k=1}^m t_k a_k - \frac{1}{2} \sum_{j,k=1}^m \sigma_{jk}^2 t_j t_k \right\} \quad (6)$$

равномерно по t_1, \dots, t_m на любом ограниченном отрезке их изменения: $|t_1| \leq T, \dots, |t_m| \leq T$. В частности, при $m = 1$ асимптотическая нормальность η_n с параметрами a, σ^2 имеет место в том и только в том случае, если

$$\varphi_n(t) \xrightarrow{n \rightarrow \infty} e^{iat - \sigma^2 t^2 / 2} \quad (7)$$

равномерно по $t \leq |T|$ при любом фиксированном T .

Часто приходится рассматривать вместо последовательностей (η_n) семейства (η_ε) случайных величин, где ε — числовой или векторный параметр, имеющий предельное значение ε_0 (возможно, бесконечное). В таком случае также рассматривается асимптотическая нормальность при $\varepsilon \rightarrow \varepsilon_0$. Все сказанное выше сохраняет смысл, только вместо « $n \rightarrow \infty$ » всюду будет « $\varepsilon \rightarrow \varepsilon_0$ ».

Пример. Пусть s_a — пуассоновская случайная величина с параметром a , $\eta_a = (s_a - a)/\sqrt{a}$. Тогда с вероятностью $e^{-a} a^n/n!$ величина η_a равна $(n - a)/\sqrt{a}$. Следовательно, ее характеристическая функция

$$\begin{aligned} \varphi_a(t) &= \sum_{n=0}^{\infty} e^{-a} \frac{a^n}{n!} e^{it(n-a)/\sqrt{a}} = e^{-a - it\sqrt{a}} \sum_{n=0}^{\infty} (ae^{it/\sqrt{a}})^n / n! = \\ &= \exp \{ a [e^{it/\sqrt{a}} - 1 - it/\sqrt{a}] \}. \end{aligned}$$

По формуле Тейлора выражение в квадратных скобках при $a \rightarrow \infty$ сходится к $(-t^2/2a)$ и притом равномерно по t при $|t| \leq T$. Следовательно, $\varphi_a(t) \xrightarrow{a \rightarrow \infty} e^{-t^2/2}$ равномерно по $t, |t| \leq T$. Тогда

$$\eta_a \sim N(0, 1), \quad a \rightarrow \infty.$$

В очень большом числе практических случаев возникает следующая задача. Одномерная случайная величина s_n есть сумма независимых или слабо зависимых (смысл этого понятия уточнять не будем) случайных величин ξ_k :

$$s_n = \xi_1 + \dots + \xi_n.$$

Обозначим $B_n^2 = Ds_n$ и введем в рассмотрение величину $\eta_n = s_n^0/B_n = (s_n - Ms_n)/B_n$, называемую *нормированным отклонением*. Если

$$\eta_n \sim N(0, 1), \quad n \rightarrow \infty,$$

или, что то же самое,

$$P(s_n < Ms_n + B_n x) = P\left(\frac{s_n^0}{B_n} < x\right) \xrightarrow{n \rightarrow \infty} \Phi(x), \quad (8)$$

говорят, что последовательность (s_n) удовлетворяет *центральной предельной теореме*. Более широкое определение, охватывающее и случай многомерных величин, состоит в следующем. Пусть

$$s_n \sim N(a_n, \Sigma_n), \quad (9)$$

где $a_n = M s_n$, Σ_n — корреляционная матрица s_n . Тогда s_n удовлетворяет центральной предельной теореме (для s_n выполняется *центральная предельная теорема*). Центральная предельная теорема — целая совокупность теорем, указывающих условия, достаточные для асимптотической нормальности сумм случайных величин (в большинстве случаев независимых). Рассмотрим случай одинаково распределенных слагаемых.

Пусть $s_n = \xi_1 + \dots + \xi_n$, где $\xi_k = (\xi_{k1}, \dots, \xi_{km})$ независимы и имеют распределение $F(x)$, не зависящее ни от k , ни от n , причем $\sigma_j^2 = D\xi_{kj} < \infty$ и корреляционная матрица K m -мерной величины ξ_k не вырождена. Тогда

$$s_n \sim N(na, nK), \quad (10)$$

где $a = (a_1, \dots, a_m) = M\xi_k$.

Соотношение (10), очевидно, эквивалентно следующим:

$$s_n^0 \sim N(0, nK), \quad (11)$$

$$s_n^0/\sqrt{n} \sim N(0, K), \quad (12)$$

$$s_n^0/n \sim N\left(0, \frac{1}{n} K\right). \quad (13)$$

Доказательство проведем с помощью исследования характеристической функции $\psi(t)$ случайной величины ξ_k .

Докажем соотношение (12) в одномерном случае. Из § 3.9 $\psi(0) = 1$, $\psi'(0) = iM\xi_k^0 = 0$, $\psi''(0) = -\sigma^2$, а следовательно,

$$\psi(t/\sqrt{n}) = 1 - \sigma^2 t^2/2n + o(1/n) \quad (14)$$

при $|t| \leq T$. Из (14) получаем, что равномерно по t , $|t| \leq T$,

$$\lim_{n \rightarrow \infty} (\psi(t/\sqrt{n}))^n = e^{-\sigma^2 t^2/2}, \quad (15)$$

откуда следует (12). Подробно этот результат формулируется следующим образом.

Пусть ξ_k — независимые одномерные случайные величины с функцией распределения $F(x)$ и положительной конечной дисперсией σ^2 . Тогда нормированное уклонение

$$s_n^0/\sigma\sqrt{n} = (s_n - Ms_n)/\sigma\sqrt{n}$$

асимптотически нормально с параметрами 0, 1, т. е.

$$P(s_n^0/\sigma\sqrt{n} < x) \xrightarrow{n \rightarrow \infty} \Phi(x). \quad (16)$$

Доказательство при произвольном m производится по той же схеме. Примем без доказательства, что $\psi(t)$ дважды дифференци-

руема при $t = 0$. Имеем

$$\psi(0, \dots, 0, t_j, 0, \dots, 0) = M e^{i t_j \xi_j},$$

откуда

$$\left. \frac{\partial \psi}{\partial t_j} \right|_{t=0} = i M \xi_j = 0, \quad \left. \frac{\partial^2 \psi}{\partial t_j^2} \right|_{t=0} = -D \xi_j.$$

Положив в $\psi(t)$ $t_j = t_s = \lambda$, а все остальные переменные равными 0, получим $M e^{i \lambda (\xi_j + \xi_s)}$; вторая производная этой функции по λ при $\lambda = 0$ есть $-D(\xi_j + \xi_s) = -D\xi_j - D\xi_s - 2 \text{cov}(\xi_j, \xi_s)$. В то же время вторая производная по λ равна $\frac{\partial^2 \psi}{\partial t_j^2} + 2 \frac{\partial^2 \psi}{\partial t_j \partial t_s} + \frac{\partial^2 \psi}{\partial t_s^2}$.

Итак, $\left. \frac{\partial^2 \psi}{\partial t_j \partial t_s} \right|_{t=0} = \text{cov}(\xi_j, \xi_s)$. Таким образом, при малых t

$$\psi(t) = 1 - \frac{1}{2} \sum_{j,s=1}^m t_j t_s \sigma_{js}^2 + o(t't), \quad (17)$$

где σ_{js}^2 — элементы матрицы K . Из (17), уменьшив t_j , t_s в \sqrt{n} раз и возведя выражение в n -ю степень, получим в пределе выражение $\exp \left\{ -\frac{1}{2} t' \Sigma t \right\}$, что и доказывает соотношение (12).

Переформулируем доказанное утверждение подробно.

Пусть ξ_n — независимые m -мерные случайные величины с функцией распределения $F(x)$ и конечной положительно определенной корреляционной матрицей $K = (\sigma_{js}^2)$. Тогда случайная величина

$$\eta_n = (\xi_1^0 + \dots + \xi_n^0) / \sqrt{n}$$

асимптотически нормальна с параметрами 0, K , а среднее арифметическое

$$x_n = (\xi_1 + \dots + \xi_n) / n$$

асимптотически нормально с параметрами $M\xi_k$, $\frac{1}{n} K$.

Пример. Пусть имеется последовательность независимых испытаний с возможными исходами 1, 2, ..., r ; вероятность исхода k в отдельном испытании равна $p_k > 0$. Обозначим через s_{nk} число исходов k в n испытаниях, $\eta_{nk} = (s_{nk} - np_k) / \sqrt{np_k}$. Рассмотрим $(r-1)$ -мерные случайные величины $s_n = (s_{n1}, \dots, s_{n,r-1})$, $\eta_n = (\eta_{n1}, \dots, \eta_{n,r-1})$. Имеем $s_{nk} = \sum_{j=1}^n \xi_{jk}$, где ξ_{jk} — индикатор исхода k в j -м испытании. Отсюда находим, что η_n есть поделенная на \sqrt{n} сумма n независимых случайных величин вида

$$\left(\frac{\xi_{j1}^0}{\sqrt{p_1}}, \dots, \frac{\xi_{j,r-1}^0}{\sqrt{p_{r-1}}} \right).$$

Величина ξ_{jk} принимает значение 1 с вероятностью p_k и значение 0 с вероятностью $q_k = 1 - p_k$. Отсюда $D(\xi_{jk} / \sqrt{p_k}) = p_k q_k / p_k = q_k$. Выражение $(\xi_{jk}^0 / \sqrt{p_k}) \times$

$\times (\xi_{js}^0/\sqrt{p_s})$ ($s \neq k$) с вероятностью p_k равно $(q_k/\sqrt{p_k})$ ($-p_s/\sqrt{p_s}) = -q_k\sqrt{p_s/p_k}$, с вероятностью p_s равно $-q_s\sqrt{p_k/p_s}$ и с вероятностью $1 - p_k - p_s$ равно $\sqrt{p_k p_s}$. Взяв математическое ожидание, получим

$$\text{cov}(\xi_{ik}/\sqrt{p_k}, \xi_{js}/\sqrt{p_s}) = -\sqrt{p_k p_s}.$$

Итак, случайная величина η_n асимптотически нормальна с параметрами 0, K , где K — матрица, (k, s) -й элемент которой равен $-\sqrt{p_k p_s}$ при $s \neq k$ и q_k при $s = k$. Нормальное распределение с корреляционной матрицей K , как несложно проверить, не вырождено.

Часто приходится рассматривать преобразования случайных величин с распределением, близким к нормальному. При определенных условиях в результате таких преобразований снова получаются величины с распределением, близким к нормальному. Приведем полезное достаточное условие асимптотической нормальности функции случайного аргумента.

Пусть $\xi_n = (\xi_{n1}, \dots, \xi_{nm})$ — m -мерная случайная величина, $\xi_n \sim N(a_n, \varepsilon_n K)$, где $a_n \rightarrow a$, $\varepsilon_n \rightarrow 0$ при $n \rightarrow \infty$, K — положительно определенная матрица. Рассмотрим случайную величину $\eta_n = f(\xi_n)$, где

$$f(x) = (f_1(x_1, \dots, x_m), \dots, f_s(x_1, \dots, x_m))$$

— функция со значениями в \mathbb{R}^s , $s \leq m$, непрерывно дифференцируема в точке $x = a$, и матрица $C = \left(\frac{\partial f_i}{\partial x_j} \right) \Big|_{x=a}$ имеет ранг s . Тогда

$$\eta_n \sim N(f(a_n), \varepsilon_n C K C'), \quad n \rightarrow \infty. \quad (18)$$

При $m = 1$ данное утверждение формулируется следующим образом. Пусть ξ_n — асимптотически нормальная случайная величина с параметрами $a_n \rightarrow a$, $\sigma_n^2 \rightarrow 0$, $f(x)$ — функция, непрерывно дифференцируемая в точке $x = a$, причем $c = f'(a) \neq 0$. Тогда

$$\eta_n = f(\xi_n) \sim N(f(a_n), c^2 \sigma_n^2). \quad (19)$$

Докажем формулу (19). Поскольку $\xi_n \sim N(a_n, \sigma_n^2)$, то $F_{\xi_n}(x) = \Phi((x - a_n)/\sigma_n) \rightarrow 0$. Значит, если выбрать Δ из условия $\Phi(-\Delta) = 1 - \Phi(\Delta) < \varepsilon/2$, то при достаточно большом n будет $F_{\xi_n}((-\Delta - a_n)/\sigma_n) < \varepsilon/2$, $1 - F_{\xi_n}((\Delta - a_n)/\sigma_n) < \varepsilon/2$, т. е.

$$P(|\xi_n - a_n| > \sigma_n \Delta) < \varepsilon. \quad (20)$$

Если теперь $|\xi_n - a_n| \leq \sigma_n \Delta$, то по теореме Лагранжа

$$\eta_n = f(\xi_n) = f(a_n) + c'(\xi_n - a_n), \quad (21)$$

где c' — значение $f'(x)$ в точке θ , промежуточной между a_n и ξ_n . Поскольку $|\theta - a| \leq |a_n - a| + |\xi_n - a_n| \leq |a_n - a| + \sigma_n \Delta$, то, так как $a_n \rightarrow a$, $\sigma_n \rightarrow 0$, имеем $|c' - c| < \varepsilon$ при достаточно большом n . Теперь можно записать

$$\eta_n = f(a_n) + c(\xi_n - a_n) + \gamma, \quad (22)$$

где $\gamma = (c' - c)(\xi_n - a_n)$, $|\gamma| < \varepsilon \sigma_n \Delta$ при $|\xi_n - a_n| \leq \sigma_n \Delta$. На основании (20), (22) имеем

$$P(\eta_n < x) = P(|\xi_n - a_n| > \sigma_n \Delta, \eta_n < x) + P(|\xi_n - a_n| \leq \sigma_n \Delta, \eta_n < x). \quad (23)$$

Первое слагаемое правой части (23) на основании (20) меньше ε , второе же вследствие (22) больше, чем

$$P(f(a_n) - c(\xi_n - a_n) < x - \varepsilon \sigma_n \Delta) = P(|\xi_n - a_n| > \sigma_n \Delta),$$

и меньше, чем

$$P(f(a_n) + c(\xi_n - a_n) < x + \varepsilon \sigma_n \Delta).$$

Взяв для определенности $c > 0$, получим

$$\begin{aligned} F_{\xi_n}\left(a_n + \frac{x - f(a_n) - \varepsilon \sigma_n \Delta}{c}\right) - \varepsilon &< P(\eta_n < x) < \\ &< F_{\xi_n}\left(a_n + \frac{x - f(a_n) + \varepsilon \sigma_n \Delta}{c}\right) + \varepsilon. \end{aligned} \quad (24)$$

Вследствие асимптотической нормальности ξ_n значения $F_{\xi_n}\left(a_n + \frac{x - f(a_n) \pm \varepsilon \sigma_n \Delta}{c}\right)$ неограниченно приближаются к $\Phi\left(\frac{x - f(a_n)}{c \sigma_n} \pm \frac{\varepsilon \Delta}{c}\right)$. В силу равномерной непрерывности функции $\Phi(x)$ из (24) получаем

$$P(\eta_n < x) - \Phi(x - f(a_n)/c \sigma_n) \rightarrow 0.$$

По ходу доказательства видно, что это соотношение равномерно относительно x . Это и означает, что

$$\eta_n \sim N(f(a_n), c^2 \sigma_n^2).$$

Замечание. Если

$$(a_n - a)/\sigma_n \rightarrow 0, \quad (25)$$

то наряду с (19) выполняется соотношение

$$\eta_n \sim N(f(a), c^2 \sigma_n^2). \quad (26)$$

Это следует из равномерной непрерывности функции $I(x)$.

Пример 1. Нелинейное звено системы автоматического управления преобразует входной сигнал ξ по закону

$$\eta = f(\xi) = \begin{cases} c\xi & \text{при } |\xi| \leq \tau, \\ c\tau & \text{при } \xi > \tau, \\ -c\tau & \text{при } \xi < -\tau, \end{cases}$$

где $\tau > 0$ — постоянная. Если ξ — приблизительно нормальная случайная величина с параметрами $0, \sigma^2$, где σ достаточно мало, то η — приблизительно нормальная величина с параметрами $0, c^2 \sigma^2$. Такой же вывод получим при $f(\xi)$ вида

$$\frac{c}{\alpha} \operatorname{tg}(\alpha \xi), \quad \frac{c}{\alpha} \sin(\alpha \xi), \quad c\xi + c_1 \xi^2, \quad e^{c\xi} - 1$$

и в других подобных примерах,

Пример 2. Два дальномера, расположенных в точках (x_1, y_1) и (x_2, y_2) , измеряют дальность R_1, R_2 до цели, расположенной в точке (x, y) . По показаниям дальномеров и координатам дальномеров из геометрических соображений определяются x, y . Погрешности показаний дальномеров на данной дальности — независимые, приблизительно нормальные случайные величины Δ_1, Δ_2 с математическим ожиданием 0 и малыми дисперсиями σ_1^2, σ_2^2 . Требуется найти приближительный вид распределения погрешностей ξ, η в определении координат цели x, y .

Имеем уравнения

$$(x - x_1)^2 + (y - y_1)^2 = R_1^2, \quad (x - x_2)^2 + (y - y_2)^2 = R_2^2.$$

Продифференцировав их по R_1 , получим

$$(x - x_1) \frac{\partial x}{\partial R_1} + (y - y_1) \frac{\partial y}{\partial R_1} = R_1, \quad (x - x_2) \frac{\partial x}{\partial R_1} + (y - y_2) \frac{\partial y}{\partial R_1} = 0,$$

откуда

$$\frac{\partial x}{\partial R_1} = \frac{1}{\Delta} R_1 (y - y_2), \quad \frac{\partial y}{\partial R_1} = -\frac{1}{\Delta} R_1 (x - x_2),$$

где $\Delta = (x - x_1)(y - y_2) - (x - x_2)(y - y_1)$. Аналогично

$$\frac{\partial x}{\partial R_2} = -\frac{1}{\Delta} R_2 (y - y_1), \quad \frac{\partial y}{\partial R_2} = R_2 (x - x_1).$$

Имеем

$$C = \frac{1}{\Delta} \begin{pmatrix} R_1 (y - y_2) & -R_2 (y - y_1) \\ -R_1 (x - x_2) & R_2 (x - x_1) \end{pmatrix},$$

$$K = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \quad C' = \frac{1}{\Delta} \begin{pmatrix} R_1 (y - y_2) & -R_1 (x - x_2) \\ -R_2 (y - y_1) & R_2 (x - x_1) \end{pmatrix},$$

$$CK = \frac{1}{\Delta} \begin{pmatrix} \sigma_1^2 R_1 (y - y_2) & -\sigma_2^2 R_2 (y - y_1) \\ -\sigma_1^2 R_1 (x - x_2) & \sigma_2^2 R_2 (x - x_1) \end{pmatrix},$$

$$CKC' = \frac{1}{\Delta^2} \times$$

$$\times \begin{pmatrix} \sigma_1^2 R_1^2 (y - y_2)^2 + \sigma_2^2 R_2^2 (y - y_1)^2 \\ (\sigma_1^2 R_1^2 (x - x_2)(y - y_2) + \sigma_2^2 R_2^2 (x - x_1)(y - y_1)) \\ - (\sigma_1^2 R_1^2 (x - x_2)(y - y_2) + \sigma_2^2 R_2^2 (x - x_1)(y - y_1)) \\ \sigma_1^2 R_1^2 (x - x_2)^2 + \sigma_2^2 R_2^2 (x - x_1)^2 \end{pmatrix}.$$

Таким образом, (ξ, η) — асимптотически нормальная случайная величина с параметрами $(0, 0)$, CKC' .

Пример 3. Закон движения свободно падающего тела в вертикальной плоскости (x, y) задается системой уравнений

$$\frac{d^2 x^2}{dt^2} = 0, \quad \frac{d^2 y}{dt^2} = -g.$$

В моменты $t = 0$ и $t = 1$ измерены координаты тела (x_0, y_0) и (x_1, y_1) . По ним найдена продольная скорость $v = x_1 - x_0$, решены уравнения движения $x(t) = x_0 + (x_1 - x_0)t$, $y(t) = -gt^2 + (y_1 - y_0 + g)t + y_0$ и найдены момент T падения тела и координаты $X = x(T)$ точки падения:

$$T = \frac{1}{2g} (y_1 - y_0 + g + \sqrt{(y_1 - y_0 + g)^2 + 4gy_0}),$$

$$X = x_0 + (x_1 - x_0)T.$$

Требуется исследовать погрешности в определении T и X , если погрешности измерения x_i , y_i — независимые, приблизительно нормальные случайные величины с параметрами 0 , σ^2 , где σ — малое число.

Обозначив для краткости $\lambda = 1/\sqrt{(y_1 - y_0 + g^2) + 4gy_0}$, найдем производные T , X по результатам измерений:

$$\begin{aligned}\frac{\partial T}{\partial x_0} &= 0, \quad \frac{\partial T}{\partial x_1} = 0, \quad \frac{\partial T}{\partial y_0} = -\frac{1}{2g} (1 + \lambda (y_0 - 4g)), \\ \frac{\partial T}{\partial y_1} &= \frac{1}{2g} (1 + \lambda y_1); \\ \frac{\partial X}{\partial x_0} &= 1 - T, \quad \frac{\partial X}{\partial x_1} = T, \quad \frac{\partial X}{\partial y_0} = (x_1 - x_0) \frac{\partial T}{\partial y_0}, \\ \frac{\partial X}{\partial x_1} &= (x_1 - x_0) \frac{\partial T}{\partial y_1}.\end{aligned}$$

Если τ , ξ — погрешности в определении T , X , то (τ, ξ) — приблизительно нормальная случайная величина с параметрами $(0, 0)$, CKC^2 , где

$$C = \begin{pmatrix} \frac{\partial T}{\partial x_0} & \frac{\partial T}{\partial x_1} & \frac{\partial T}{\partial y_0} & \frac{\partial T}{\partial y_1} \\ \frac{\partial X}{\partial x_0} & \frac{\partial X}{\partial x_1} & \frac{\partial X}{\partial y_0} & \frac{\partial X}{\partial y_1} \end{pmatrix},$$

$$K = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}.$$

Пусть $\xi_n = (\xi_{n1}, \dots, \xi_{nm}) \sim N(a, K)$, $f(x) = f(x_1, \dots, x_m)$ — непрерывная функция со значениями в \mathbb{R}^s (на этот раз s — любое число). Обозначим $\eta_n = f(\xi_n)$, $\eta = f(\xi)$, где ξ — нормальная случайная величина, $M\xi = a$, $M\xi\xi^T = K$. Тогда

$$F_{\eta_n}(x) \xrightarrow[n \rightarrow \infty]{w} F_{\eta}(x). \quad (27)$$

Эта формула непосредственно следует из определения слабой сходимости (гл. 5).

Пример 4. Пусть имеется n независимых испытаний с равновероятными исходами 1, 2, 3; $\eta_{nk} = (s_{nk} - n/3)/\sqrt{n/3}$, где s_{nk} — число исходов k . Рассмотрим величину $\eta_{n1}^2 + \eta_{n2}^2 + \eta_{n3}^2$.

Имеем $M\eta_{nk} = 0$. Корреляционная матрица K величины $(\eta_{n1}, \eta_{n2}, \eta_{n3})$ имеет вид

$$K = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}.$$

Пусть теперь $\xi = (\xi_1, \xi_2, \xi_3)$ — нормальная величина с математическим ожиданием 0 и корреляционной матрицей K . Возьмем теперь независимые нормальные величины η_1, η_2 ; $M\eta_i = 0$, $D\eta_i = 2/3$, $i = 1, 2$. Положим $\xi'_1 = \eta_1$, $\xi'_2 = -\eta_1/2 + \eta_2\sqrt{3}/2$, $\xi'_3 = -\eta_1/2 - \eta_2\sqrt{3}/2$. Тогда, как это легко проверить, (ξ'_1, ξ'_2, ξ'_3) имеет ту же корреляционную матрицу, что и (ξ_1, ξ_2, ξ_3) ; $M\xi'_i = 0$, $i = 1, 2, 3$. Следовательно, распределение величины (ξ'_1, ξ'_2, ξ'_3) — то же, что и для (ξ_1, ξ_2, ξ_3) . Имеем

$$\xi_1'^2 + \xi_2'^2 + \xi_3'^2 = (\eta_1/\sqrt{2/3})^2 + (\eta_2/\sqrt{2/3})^2,$$

где под знаком квадрата — независимые нормальные случайные величины. Отсюда находим, что величина $\zeta_1'^2 + \zeta_2'^2 + \zeta_3'^2$, а вместе с ней и $\zeta_1^2 + \zeta_2^2 + \zeta_3^2$, имеет распределение χ_2^2 . Итак, распределение $\eta_{n1}^2 + \eta_{n2}^2 + \eta_{n3}^2$ при $n \rightarrow \infty$ слабо сходится к распределению χ_2^2 .

§ 2. Предельная теорема для ограниченных слагаемых

Пусть при любом n имеется n независимых одномерных случайных величин $\xi_{n1}, \dots, \xi_{nn}$; $s_n = \xi_{n1} + \dots + \xi_{nn}$; $a_{nk} = M\xi_{nk}$; $A_n = a_{n1} + \dots + a_{nn}$; $b_{nk}^2 = D\xi_{nk}$; $B_n^2 = b_{n1}^2 + \dots + b_{nn}^2$. Дополнительно предположим, что для некоторой последовательности чисел $\varepsilon_n \rightarrow 0$

$$P(|\xi_{nk}^0| > B_n \varepsilon_n) = 0. \quad (1)$$

Тогда

$$s_n \sim N(A_n, B_n^2). \quad (2)$$

Для доказательства перейдем от величин ξ_{nk} к нормированным величинам $\alpha_{nk} = \xi_{nk}/B_n$ и обозначим $\eta_n = \alpha_{n1} + \dots + \alpha_{nn} = s_n^0/B_n$. Тогда вместо сформулированных выше условий получим следующие:

- 1) α_{nk} независимы;
- 2) $M\alpha_{nk} = 0$;
- 3) $D\alpha_{nk} = c_{nk}^2 = b_{nk}^2/B_n^2$;
- 4) $D\eta_n = 1$;
- 5) $P(|\alpha_{nk}| > \varepsilon_n) = 0$;
- 6) $D\alpha_{nk} \leq \varepsilon_n^2 \rightarrow 0$

(последнее следует из 5), поскольку в данном случае $D\alpha_{nk} = M\alpha_{nk}^2$).

Обозначим через $\psi_{nk}(t)$ характеристическую функцию α_{nk} , через $\varphi_n(t) = \psi_{n1}(t) \dots \psi_{nn}(t)$ — характеристическую функцию η_n . Имеем

$$\begin{aligned} \psi_{nk}(t) &= Me^{it\alpha_{nk}} = M1 + itM\alpha_{nk} - \frac{1}{2}t^2M\alpha_{nk}^2 + \\ &+ M\left(e^{it\alpha_{nk}} - 1 - it\alpha_{nk} + \frac{1}{2}t^2\alpha_{nk}^2\right). \end{aligned} \quad (3)$$

Первое слагаемое правой части (3) равно 1, второе 0, третье $(-t^2c_{nk}^2/2)$. Выражение в скобках в последнем слагаемом по модулю меньше $|t^3\alpha_{nk}^3|/6 \leq |t|^3\alpha_{nk}^2\varepsilon_n/6$; следовательно, $|M(\dots)| \leq |t|^3 \times c_{nk}^2\varepsilon_n/6 = (t^2c_{nk}^2/2)|t\varepsilon_n/6|$. Отсюда находим, что при $n \rightarrow \infty$

$$\psi_{nk}(t) - 1 \sim -t^2c_{nk}^2/2 \quad (4)$$

равномерно по k .

Поскольку при малых z выполняется соотношение $\ln(1+z) \sim z$, то из (4) получаем

$$\ln \psi_{nk}(t) \sim -t^2c_{nk}^2/2, \quad n \rightarrow \infty, \quad (5)$$

равномерно по k , так как вследствие условия $6 \ c_{nk}^2 \leq \varepsilon_n^2 \rightarrow 0$. Из (5)

$$\ln \varphi_n(t) \sim -t^2 \sum_{k=1}^n c_{nk}^2 / 2 = -t^2 / 2, \quad n \rightarrow \infty,$$

откуда

$$\varphi_n(t) \sim e^{-t^2/2}, \quad n \rightarrow \infty. \quad (6)$$

По ходу доказательства видно, что соотношение (6) равномерно по t при $|t| \leq T$. На основании обратной теоремы непрерывности для характеристических функций (§ 5.3) заключаем, что $\eta_n \sim N(0, 1)$, что и требовалось доказать.

§ 3. Теорема Линдеберга

Пусть $\xi_{n1}, \dots, \xi_{nn}$ — независимые случайные величины; $M\xi_{nk} = a_{nk}$; $A_n = a_{n1} + \dots + a_{nn}$; $b_{nk}^2 = D\xi_{nk}$; $B_n^2 = b_{n1}^2 + \dots + b_{nn}^2 > 0$, причем выполняется условие Линдеберга: для любого $\tau > 0$

$$\frac{1}{B_n^2} \sum_{k=1}^n M(\xi_{nk}^0)^2 I(|\xi_{nk}^0| > B_n \tau) \xrightarrow{n \rightarrow \infty} 0, \quad (1)$$

где $I(\dots)$ — индикатор события, указанного в скобках. Тогда, если $s_n = \xi_{n1} + \dots + \xi_{nn}$, то $s_n \sim N(A_n, B_n^2)$.

Данное утверждение называется *теоремой Линдеберга* (более подробно: *центральной предельной теоремой Линдеберга*).

Условие Линдеберга имеет следующий вероятностный смысл. Можно записать

$$\begin{aligned} Ds_n = B_n^2 &= \sum_{k=1}^n b_{nk}^2 = \sum_{k=1}^n \int x^2 dF_{s_{nk}}(x) = \\ &= \sum_{k=1}^n \int_{|x| \leq \tau B_n} x^2 dF_{s_{nk}}(x) + \sum_{k=1}^n \int_{|x| > \tau B_n} x^2 dF_{s_{nk}}(x). \end{aligned}$$

Сумму интегралов по области $|x| > \tau B_n$ естественно назвать вкладом в дисперсию s_n уклонений слагаемых ξ_{nk} от их математических ожиданий, сравнимых со среднеквадратическим отклонением суммы, т. е. B_n . Таким образом, при любом фиксированном $\tau > 0$ основной вклад обеспечивается уклонениями, меньшими τB_n .

Доказательство теоремы Линдеберга состоит в следующем. Как и в предыдущем параграфе, перейдем к нормированным величинам $\alpha_{nk} = \xi_{nk}^0 / B_n$ и их суммам $\eta_n = \alpha_{n1} + \dots + \alpha_{nn}$. Для них условие Линдеберга примет более простой вид

$$\sum_{k=1}^n M\alpha_{nk}^2 I(|\alpha_{nk}| > \tau) \xrightarrow{n \rightarrow \infty} 0. \quad (2)$$

Поскольку (2) имеет место для любого $\tau > 0$, то найдется такая последовательность $\varepsilon_n \rightarrow 0$, что

$$\sum_{k=1}^n M \alpha_{nk}^2 I(|\alpha_{nk}| > \varepsilon_n) \xrightarrow{n \rightarrow \infty} 0. \quad (3)$$

Далее применяется полезный во многих задачах прием усечения, а именно, вводится новая случайная величина β_{nk} , равная α_{nk} при $|\alpha_{nk}| \leq \varepsilon_n$ и равная 0 в противном случае. Для компенсации усечения вводится величина γ_{nk} , равная 0 при $|\alpha_{nk}| \leq \varepsilon_n$ и равная α_{nk} в противном случае. Таким образом, $\alpha_{nk} = \beta_{nk} + \gamma_{nk}$ с вероятностью 1. Величины β_{nk} ограничены, а следовательно, к ним применим результат предыдущего параграфа, т. е. $\sum \beta_{nk}$ — асимптотически нормальная величина. Что касается $\sum \gamma_{nk}$, то сумма стремится к нулю по вероятности: из неравенства Чебышева

$$P\left(\left|\sum_{k=1}^n \gamma_{nk}\right| > \tau\right) \leq \frac{1}{\tau^2} \sum_{k=1}^n D\gamma_{nk} \leq \frac{1}{\tau^2} \sum_{k=1}^n M \alpha_{nk}^2 I(|\alpha_{nk}| > \tau) \xrightarrow{n \rightarrow \infty} 0. \quad (4)$$

Таким образом,

$$\begin{cases} F_{\beta_{n1} + \dots + \beta_{nn}}(x) \xrightarrow{w} \Phi(x), \\ \eta_n - (\beta_{n1} + \dots + \beta_{nn}) \xrightarrow{p} 0, \end{cases}$$

откуда следует, что $\eta_n \sim N(0, 1)$, $n \rightarrow \infty$.

Две технические подробности доказательства были нами опущены: 1) центрирование β_{nk} ; 2) поправка на дисперсию $\beta_{n1} + \dots + \beta_{nn}$, которая отличается от 1 на бесконечно малое слагаемое. В обоих случаях требуемые оценки следуют из условия (3).

§ 4. Теорема Ляпунова

Условие Линдеберга весьма общо, но громоздко при практической проверке. Вместо него проще использовать *условие Ляпунова* при некотором $\delta > 0$

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M |\xi_{nk}^0|^{2+\delta} \xrightarrow{n \rightarrow \infty} 0. \quad (1)$$

Для нормированных величин условие Ляпунова имеет вид

$$\sum_{k=1}^n M |\alpha_{nk}|^{2+\delta} \xrightarrow{n \rightarrow \infty} 0. \quad (2)$$

Из условия Ляпунова следует условие Линдеберга, так как

$$\alpha_{nk}^2 I(|\alpha_{nk}| > \tau) \leq \tau^{-\delta} |\alpha_{nk}|^{2+\delta},$$

откуда

$$\sum_{k=1}^n M \alpha_{nk}^2 I(|\alpha_{nk}| > \tau) \leq \tau^{-\delta} \sum_{k=1}^n M |\alpha_{nk}|^{2+\delta} \xrightarrow{n \rightarrow \infty} 0.$$

Следовательно, справедлива *теорема Ляпунова* (более подробно: *центральная предельная теорема в форме Ляпунова*).

Пусть $s_n = \xi_{n1} + \dots + \xi_{nn}$, где ξ_{nk} — независимые случайные величины, $A_n = M s_n$, $B_n^2 = D s_n$, причем выполняется условие (1). Тогда $s_n \sim N(A_n, B_n^2)$.

В практических случаях чаще всего используется условие Ляпунова при $\delta = 1$:

$$\frac{1}{B_n^3} \sum_{k=1}^n M |\xi_{nk}^0|^3 \xrightarrow{n \rightarrow \infty} 0 \quad (3)$$

и при $\delta = 2$:

$$\frac{1}{B_n^4} \sum_{k=1}^n M (\xi_{nk}^0)^4 \xrightarrow{n \rightarrow \infty} 0. \quad (4)$$

Докажем достаточность (3) непосредственно, а не через условие Линдеберга. Для нормированных величин вместо (3) имеем

$$\sum_{k=1}^n |\alpha_{nk}| \xrightarrow{n \rightarrow \infty} 0. \quad (5)$$

Если $\psi_{nk}(t)$ — характеристическая функция α_{nk} , $F_{nk}(x)$ — ее функция распределения, то

$$\begin{aligned} \psi_{nk}(t) &= \int e^{itx} dF_{nk}(x) = \int (1 + itx - t^2 x^2/2) dF_{nk}(x) + \\ &+ \int [e^{itx} - 1 - itx + t^2 x^2/2] dF_{nk}(x). \end{aligned}$$

Первое слагаемое правой части этого равенства есть $1 - t^2 c_{nk}^2/2$, где $c_{nk}^2 = D \alpha_{nk}$. Выражение в квадратных скобках по модулю не превосходит $|tx|^3/6$, так что интеграл оценивается величиной $|t|^3 M |\alpha_{nk}^3|/6$.

Итак,

$$\psi_{nk}(t) - 1 = -t^2 c_{nk}^2/2 + l_{nk}, \quad (6)$$

где равномерно по t , $|t| \leq T$,

$$|l_{n1}| + \dots + |l_{nn}| \rightarrow 0 \quad (7)$$

вследствие условия Ляпунова. Далее,

$$\begin{aligned} c_{nk}^2 &= \int x^2 dF_{nk}(x) = \int_{|x| \leq \tau} x^2 dF_{nk}(x) + \int_{|x| > \tau} x^2 dF_{nk}(x) \leq \\ &\leq \tau^2 + \frac{1}{\tau} \int |x|^3 dF_{nk}(x). \end{aligned}$$

Поскольку $\tau > 0$ произвольно, то τ^2 можно сделать сколь угодно малым; выражение $\frac{1}{\tau} \int \dots$ сколь угодно мало по условию Ляпунова. Итак, $c_{nk}^2 \rightarrow 0$ при $n \rightarrow \infty$ равномерно по k . В силу (6) это относится и к l_{nk} . Теперь из (6) получаем

$$\ln \psi_{nk}(t) \sim -t^2 c_{nk}/2 + l_{nk},$$

$$\ln \prod_{k=1}^n \psi_{nk}(t) \sim -t^2/2 + \sum_{k=1}^n l_{nk} \sim -t^2/2,$$

а следовательно, равномерно по t , $|t| \leq T$,

$$\prod_{k=1}^n \psi_{nk}(t) \xrightarrow{n \rightarrow \infty} e^{-t^2/2}.$$

В силу обратной теоремы непрерывности из § 5.3 отсюда следует, что $s_n \sim N(A, B_n^2)$.

Пример. Вычисляется суммарное значение некоторого показателя путем сложения n округленных частных показателей. Погрешность округления — равномерно распределенная в интервале $(-\Delta/2; \Delta/2)$ случайная величина, но значение Δ не во всех слагаемых одно и то же: оно равно Δ_1 для n_1 слагаемых и Δ_2 для остальных $n_2 = n - n_1$ слагаемых. Можно ли считать суммарную погрешность приблизительно нормальной?

Обозначим через ξ_k погрешность k -го измерения. Тогда

$$\begin{aligned} M\xi_k &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} x dx = 0, & D\xi_k &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} x^2 dx = \Delta^2/12, \\ M|\xi_k^0|^3 &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} |x|^3 dx = (2/\Delta) \int_0^{\Delta/2} x^3 dx = \Delta^3/32. \end{aligned}$$

Отсюда

$$\begin{aligned} A_n &= 0, & B_n^2 &= (n_1 \Delta_1^2 + n_2 \Delta_2^2)/12, \\ \frac{1}{B_n^3} \sum_{k=1}^n M|\xi_k^0|^3 &= \frac{3\sqrt{3}}{4} \frac{n_1 \Delta_1^3 + n_2 \Delta_2^3}{(n_1 \Delta_1^2 + n_2 \Delta_2^2)^{3/2}}. \end{aligned}$$

Пусть для определенности $\Delta_1 > \Delta_2$, т. е. $\Delta_1 = \rho \Delta_2$, $\rho > 1$. Тогда

$$(n_1 \Delta_1^3 + n_2 \Delta_2^3)/(n_1 \Delta_1^2 + n_2 \Delta_2^2)^{3/2} = (n_1 \rho^3 + n_2)/(n_1 \rho^2 + n_2)^{3/2}.$$

Как ведет себя эта дробь при $n \rightarrow \infty$? Если ρ ограничено, она стремится к 0. Если $\rho \rightarrow \infty$, $n_1 \rightarrow \infty$, эффект тот же. Остается один случай: $\rho \rightarrow 0$, n_1 ограничено. Если n_1 ограничено, $\rho \rightarrow \infty$, но $\rho = o(\sqrt{n})$, то дробь также стремится к нулю. Наконец, при ограниченном n_1 и ρ , растущем как \sqrt{n} , асимптотическая нормальность нарушается. Так, при $n_1 = 1$, $\rho = \sqrt{n}$ значение дроби равно $1/2 \sqrt{2}$. В этом случае единственное грубое измерение искажает нормальность распределения. Поскольку B_n имеет порядок \sqrt{n} , то видим, что нарушается условие Линдеберга: существует отдельное слагаемое, вносящее существенный вклад в дисперсию суммы.

Центральная предельная теорема служит основанием замены распределений случайных величин, образующихся в результате сложения независимых величин, мало влияющих на сумму, нормальным распределением. К подобным величинам относятся погрешности наблюдений, отклонения параметров траекторий объектов от расчетных, характеристики шумов в радиоустройствах и многие другие. Тем не менее использование нормального распределения в любом конкретном приложении требует специального обоснования, включающего обработку опытных данных: случайные факторы могут взаимодействовать нелинейным образом, а отдельные из них оказывают существенное влияние на окончательный эффект.

Пример 1. Число элементов сложной радиотехнической системы, отказавших в течение большого отрезка времени, можно было бы считать нормальной случайной величиной при заданном температурном режиме. При разладке системы регулирования температуры может создаться режим, когда вероятность отказа элемента значительно повышается. В результате получается распределение, отличное от нормального. Особенно опасно применение «трехсигмового интервала» там, где это недопустимо: функция распределения $F(x)$ может действительно быть близкой к $\Phi((x-a)/\sigma)$, скажем, при $|x-a| < 2,5$, но из этого не следует, что 2,5 можно заменить на 3.

Пример 2. Значение уровня шума в канале системы обнаружения — приблизительно нормальная случайная величина с параметрами 0, σ^2 . Требуется рассчитать такой пороговый уровень z , при котором вероятность превышения его шумом (вероятность ложной тревоги) будет меньше 10^{-7} . Исходя из нормального распределения, получаем $z \approx 5,2$. Но в области столь больших отклонений рассматриваемая случайная величина скорее всего не исследовалась, так что данный вывод был бы бесосновательным.

Центральная предельная теорема доказывалась и для *многомерных случайных величин*. Ограничимся формулировкой одного общего критерия при положительно определенной K .

Пусть $s_n = (s_{n1}, \dots, s_{nm})$ — m -мерные случайные величины. Для того чтобы $s_n \sim N(a, K)$, $n \rightarrow \infty$, необходимо и достаточно, чтобы при любом фиксированном $t \in \mathbb{R}^m$ для одномерных величин $(t, s_n) = t_1 s_{n1} + \dots + t_m s_{nm}$ выполнялось соотношение

$$(t, s_n) \sim N((t, a), t' K t), \quad n \rightarrow \infty. \quad (8)$$

Этот критерий позволяет в принципе сводить вопрос об асимптотической нормальности многомерных величин к соответствующему вопросу для одномерных величин.

§ 5. Отклонение от нормального распределения

Центральная предельная теорема не указывает, как далеко может отклониться распределение суммы конечного числа независимых случайных величин от нормального распределения. Имеются многие оценки, дающие ответ на этот вопрос. В работах Берри и Эссеена начала 40-х гг. было доказано, то если ξ_k — независимые одинаково распределенные одномерные случайные величины, $\sigma^2 = D\xi_k > 0$, $\rho = M|\xi_k|^3$, $s_n = \xi_1 + \dots + \xi_n$, то

$$|P(s_n^0 < x\sigma\sqrt{n}) - \Phi(x)| < c\rho/\sigma^3\sqrt{n}, \quad (1)$$

где c — постоянная, не зависящая от вида распределения ξ_k . В. М. Золотарев в 1966 г. установил, что $c \leq 0,9051$. Оценки, подобные (1), называются *точными*. Помимо них, существуют *асимптотические оценки*, указывающие поведение отклонения распределения $s_n^0/\sigma\sqrt{n}$ от $\Phi(x)$.

Пусть характеристическая функция $\varphi(t)$ случайных величин ξ_k удовлетворяет условию

$$\int_{-\infty}^{\infty} |\varphi(t)|^m dt < \infty.$$

Тогда при $n \geq t$ случайная величина $s_n^0/\sigma\sqrt{n}$ имеет плотность $p_n(x)$ и при $n \rightarrow \infty$ разность $p_n(x) - e^{-x^2/2}/\sqrt{2\pi}$ раскладывается в асимптотический ряд по степеням $1/\sqrt{n}$. Наибольшее значение имеет начальный отрезок асимптотического ряда:

$$p_n(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left(1 + \frac{1}{\sqrt{n}} (M(\xi^0)^3/6\sigma^3) (x^3 - 3x) + \right. \\ \left. + \frac{1}{n} (M((\xi^0)^3)^3/72\sigma^6) (x^3 - 3x) + \right. \\ \left. + \frac{1}{n} (M((\xi^0)^4 - 3\sigma^4)/24\sigma^4) (x^4 - 6x^2 + 3) \right) + o\left(\frac{1}{n}\right). \quad (2)$$

Это асимптотическое разложение называется *разложением Эджворта*. О точности аппроксимации формулой (2) дает представление пример, взятый с известной книги Т. Фрая «Теория вероятностей для инженеров» (М.; Л., 1934). Биномиальную вероятность p_{nk} вначале аппроксимируем предельным выражением из формулы Муавра — Лапласа. Тогда при $n = 100$, $p = 0,1$, $k = 20$ имеем $p_{nk} \approx 0,00117$; ошибка указанной аппроксимации равна $-6,6 \times 10^{-4}$; в то же время ошибка аппроксимации (2) равна $3 \cdot 10^{-5}$.

§ 6. Другие предельные законы

Пусть $s_n = \xi_{n1} + \dots + \xi_{nn}$, где $\xi_{n1}, \dots, \xi_{nn}$ — независимые одномерные случайные величины. Предположим, что для любого $\varepsilon > 0$

$$\max_{1 \leq k \leq n} P(|\xi_{nk}| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0. \quad (1)$$

В таком случае случайные величины ξ_{nk} называются *бесконечно малыми*. Если

$$F_{s_n}(x) \xrightarrow{n \rightarrow \infty} G(x), \quad (2)$$

где $G(x)$ — распределение вероятностей, то $G(x)$ называется *предельным распределением* суммы независимых бесконечно малых случайных величин. Три типа предельных распределений нам уже знакомы. Это:

1) вырожденное распределение

$$E(x - a) = \begin{cases} 0 & \text{при } x \leq a, \\ 1 & \text{при } x > a, \end{cases}$$

так как по закону больших чисел $\bar{x}_n \xrightarrow{p} a = M\xi_k$ при $n \rightarrow \infty$, так что для выполнения (2) достаточно взять $\xi_{nk} = \xi_k/n$, $1 \leq k \leq n$, где ξ_k независимы, одинаково распределены и имеют математическое ожидание a ;

2) распределение Пуассона

$$p(m) = e^{-a} a^m / m!,$$

поскольку (2) выполняется при независимых ξ_{nk} , равных 1 с вероятностью a/n и 0 с вероятностью $1 - a/n$ ($n \geq a$);

3) нормальное распределение $\Phi((x - a)/\sigma)$: достаточно взять α_{nk} , удовлетворяющие условию Линдеберга, а затем $\xi_{nk} = \sigma\alpha_{nk} + a/n$.

Кроме этих трех, существует бесконечное множество типов предельных распределений. Исследование таких распределений и условий сходимости к ним долгое время составляло основную задачу теории вероятностей. К середине 50-х гг. была создана исчерпывающая теория предельных распределений. Все они оказались *безгранично делимыми*, т. е. случайная величина s с распределением $G(x)$ при любом $n \geq 2$ может быть представлена в виде суммы n независимых одинаково распределенных случайных величин: $s = \xi_{n1} + \dots + \xi_{nn}$.

Приведем полезную в приложениях схему сходимости к предельному закону. Пусть ξ_{nk} , $1 \leq k \leq n$, — независимые бесконечно малые величины с общей функцией распределения $F_n(x)$. Из бесконечной малости ξ_{nk} следует, что «почти вся» вероятность сосредоточена в интервале $|x| \leq \varepsilon_n$, где $\varepsilon_n \rightarrow 0$, $n \rightarrow \infty$. При условии Линдеберга уклонения, выходящие за пределы этого интервала, не оказывают заметного влияния ни на распределение суммы, ни на ее дисперсию, и в пределе получается нормальный закон. В более общем случае это не так. Пусть, например, ξ_{nk} с вероятностью $1 - a/n$ равна величине α_{nk} , $|\alpha_{nk}| \leq \varepsilon_n$, и с вероятностью a/n — не бесконечно малой величине β_{nk} с функцией распределения $B(x)$. Сумма величин α_{nk} при соответствующих условиях асимптотически нормальна, в то время как $\beta_n = \sum_k \beta_{nk}$ имеет предельное распределение

$$G_1(x) = \lim_{n \rightarrow \infty} P(\beta_n < x) = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} B^{*(k)}(x),$$

где $B^{*(k)}(x)$ — k -кратная свертка распределения $B(x)$. Переходя к характеристическим функциям, получим следующий результат. Пусть

$$\psi(t) = \int e^{itx} dB(x), \quad \varphi_1(t) = \int e^{itx} dG_1(x).$$

Тогда

$$\varphi_1(t) = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} \psi^k(t) = e^{-a(1-\psi(t))}.$$

Учитывая нормальную величину, найдем, что

$$\varphi(t) = \int e^{itx} dG(x) = \exp\{ibt - \sigma^2 t^2/2 - a(1 - \psi(t))\},$$

или, что то же самое,

$$\ln \varphi(t) = ibt - \sigma^2 t^2/2 + a \int_{-\infty}^{\infty} (e^{itx} - 1) dB(x). \quad (3)$$

В частности, при $a = 0$ из (3) получаем нормальное распределение, при $b = \sigma = 0$, $a > 0$ — распределение Пуассона с параметром a .

Пример 1. Выполнение программы ЭВМ требует 10^8 операций, случайных по длительности (например, из-за случайности времени записи и считывания). Среднее время операции $1,5 \cdot 10^{-8}$ с, среднее квадратическое отклонение $2,5 \cdot 10^{-7}$ с. Возможны случайные срывы программы с вероятностью 10^{-9} на каждой операции; каждый срыв затягивает ее выполнение на $1,2 \cdot 10^3$ с. Какое время T следует отнести на решение задачи, чтобы вероятность решения за это время была не менее 0,999?

Среднее время решения без учета срывов равно 1,5 с; $\sigma = 2,5 \cdot 10^{-3}$ с, т. е. трехсигмовый запас времени менее 10^{-2} с. Число срывов — приблизительно пуассоновская случайная величина с параметром $a = 0,1$. Вероятность трех или большего числа срывов приблизительно равна $2 \cdot 10^{-4}$. Таким образом, нужно запасное время 2,4 с в расчете на возможные срывы. Итак, $T \approx 3,91$ с.

Пример 2. При отказе технического устройства требуется восстановление в течение случайного времени η , $F_\eta(x) = B(x)$. Отказы образуют простейший поток однородных событий с параметром λ (см. § 4.5). Найти математическое ожидание и дисперсию суммарного времени γ восстановления устройств, отказавших в интервале длины T .

Как и в предыдущем примере,

$$F_\gamma(x) = e^{-\lambda T} \sum_{k=0}^{\infty} (\lambda^k T^k / k!) P(\eta_1 + \dots + \eta_k < x),$$

где η_i — независимые величины с распределением $B(x)$.

Обозначим $\varphi(t) = Me^{it\gamma}$, $\psi(t) = Me^{it\eta}$. Тогда

$$\varphi(t) = \exp\{-\lambda T(1 - \psi(t))\}; \quad \varphi'(t) = \lambda T \psi'(t) \varphi(t);$$

$$\varphi''(t) = \lambda T(\psi''(t) + \lambda T(\psi'(t))^2) \varphi(t).$$

Пусть $\tau = M\eta$, $\sigma^2 = D\eta$. Тогда $\psi'(0) = i\tau$, $\psi''(0) = iM\gamma$, $\varphi(0) = 1$, откуда $M\gamma = \lambda T\tau$. Далее $\varphi''(0) = -M\gamma^2$, $\psi''(0) = -M\eta^2$.

После несложных преобразований получаем $D\eta = \lambda T(\tau^2 + \sigma^2)$. Пусть $\varphi(t)$ — характеристическая функция произвольного безграничного делимого распределения. Тогда справедлива *формула Леви — Хинчина*

$$\ln \varphi(t) = ita - \sigma^2 t^2 / 2 + \int_{-\infty}^{\infty} (e^{itx} - 1 - itx) \frac{1+x^2}{x^2} dG(x), \quad (4)$$

где $G(x)$ — неубывающая ограниченная функция.

Задачи

1. Случайная величина s_n равна сумме независимых случайных величин:

$$s_n = \sum_{k=1}^m \sum_{j=1}^{n_k} \xi_{kj},$$

где ξ_{kj} имеют функцию распределения $F_k(x)$ и дисперсию $\sigma_k^2 > 0$. Методом характеристических функций доказать, что если $n_k \rightarrow \infty$, $1 \leq k \leq m$, то s_n асимптотически нормальна при $n \rightarrow \infty$.

2. Пусть $s_n = \xi_{n1} + \dots + \xi_{nn}$, где ξ_{nk} — независимые пуассоновские случайные величины с параметром $a = 1/n$. Тогда $s_n = \text{Poiss}(1)$ при любом n , т. е. центральная предельная теорема не выполняется. Доказать, что условие Ляпунова для последовательности (s_n) нарушено.

3. Решить задачу, аналогичную предыдущей, в предположении, что ξ_{nk} имеют гамма-распределение с параметрами $\lambda, 1/n$.

4. Пусть $s_n = \xi_{n1} + \dots + \xi_{nn}$, $n \geq 2$, где ξ_{nk} независимы и имеют распределение

$$P(\xi_{nk} = \pm 1) = \frac{1}{2} \left(1 - \frac{1}{n}\right),$$

$$P(\xi_{nk} = \pm n^\alpha) = \frac{1}{2n}.$$

При каких α для последовательности выполняется центральная предельная теорема?

5. Пусть ξ — стандартная нормальная случайная величина,

$$s_n = n \arcsin(\xi/n).$$

Найти разложение плотности распределения случайной величины s_n в ряд по степеням $1/n$. Выписать формулы для коэффициентов при n^{-k} , $k \leq 3$.

6. Пусть ξ, η — одномерные нормальные случайные величины с нулевыми математическими ожиданиями, единичными дисперсиями и коэффициентом корреляции ρ . Найти приближенные формулы при малых ρ для вероятностей следующих событий:

- а) $\{\xi^2 + \eta^2 < R^2\}$; б) $\{|\xi| < u, |\eta| < v\}$;
 в) $\{\max\{\xi, \eta\} > z\}$; г) $\{\max\{|\xi|, |\eta|\} > z\}$.

7. На плоскости задан треугольник Δ , не содержащий начало координат. Обозначим через $P(\sigma, \Delta)$ вероятность того, что двумерная нормальная случайная величина (ξ_1, ξ_2) с нулевым математическим ожиданием и корреляционной матрицей $\sigma^2 I$ попадет в треугольник Δ . Найти приближенную формулу для $P(\sigma, \Delta)$ при малых σ . Отдельно рассмотреть случаи, когда минимум функции $x^2 + y^2$ при $(x, y) \in \Delta$ достигается на одной из вершин и на внутренней точке стороны треугольника.

8. Пусть s_n — биномиальная случайная величина с параметрами $n, 1/2$. Доказать, что момент распределения величины $(2s_n - n)/\sqrt{n}$ любого фиксированного порядка сходится при $n \rightarrow \infty$ к соответствующему моменту стандартного нормального распределения.

9. Дать пример такой последовательности независимых случайных величин ξ_{nk} , для которых выполняются следующие условия:

1) ξ_{nk} одинаково распределены при любом фиксированном n ;

2) $M\xi_{nk} = 0$; 3) $D\xi_{nk} = 1/n$;

4) $s_n = \xi_{n1} + \dots + \xi_{nk}$ асимптотически нормальна с математическим ожиданием 0 и дисперсией $\sigma^2 < 1$.

У к а з а н и е. Вероятностную «массу», определяющую распределения ξ_{nk} , снабдить такими «грузиками», которые практически не влияют на вероятности событий $\{s_n < x\}$, но вносят существенный вклад в дисперсию s_n .

10. Пусть ω_k — независимые случайные величины, равномерно распределенные в интервале $(-1; 1)$. Найти предельное распределение при $n \rightarrow \infty$ произведения $\omega_1 \dots \omega_n$ при соответствующей нормировке.

У к а з а н и е. Догадаться, как свести задачу к произведению положительных величин, чтобы перейти к логарифмам, а следовательно, к сумме независимых случайных величин.

11. Имеется N «ячеек», в которые последовательно независимо бросаются «дробинки». Вероятность данной дробинке попасть в определенную ячейку равна $1/N$. Обозначим s_n число брошенных дробин, при которых впервые число ячеек, содержащих хотя бы одну дробинку, становится равным n . (Так, $s_1 = 1$ в любом

случае; $s_2 = 4$ в случае, если первые три дробинки попали в одну и ту же ячейку, четвертая — в некоторую другую). Доказать асимптотическую нормальность s_n при $N \rightarrow \infty$, $n = [\alpha N]$, где $\alpha \in (0; 1)$ — фиксированное число.

У к а з а н и е. Воспользоваться представлением $s_n = \tau_1 + \dots + \tau_n$, где τ_i — «время ожидания» попадания в новую ячейку в состоянии, когда заполнены $i - 1$ ячеек. Доказать и использовать тот факт, что

$$P(\tau_i = k) = \left(\frac{i-1}{N} \right)^{k-1} \left(1 - \frac{i-1}{N} \right), \quad k \geq 0.$$

Глава 7

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 1. Случайный процесс, заданный характеристиками второго порядка

Пусть Δ — некоторое числовое множество. Семейство случайных величин $(\xi(t), t \in \Delta)$, заданных на вероятностном пространстве (Ω, \mathcal{F}, P) , называется *случайным процессом*. При любом заданном $t \in \Delta$ случайная величина $\xi(t)$ называется *значением* случайного процесса в момент t . Случайный процесс называется *одномерным* или *многомерным*, если случайные величины $\xi(t)$ — соответственно одномерные или многомерные. Если эти величины — комплексные, случайный процесс называется *комплексным*, если действительные, — *действительным*. Поскольку $\xi(t) = \xi(t, \omega)$, $\omega \in \Omega$, то случайный процесс задается функцией двух переменных: t и ω . При фиксированном ω получаем функцию $\xi(t)$ одной переменной t , называемую *траекторией* (*реализацией*) случайного процесса.

Для любого выбора $t_i \in \Delta$, $1 \leq i \leq n$, вектор $(\xi(t_1), \dots, \xi(t_n))$ есть многомерная случайная величина, имеющая некоторую функцию распределения. Преобладающее большинство приложений не требует знания этих распределений и основано на *характеристиках второго порядка* — математическом ожидании случайного процесса и корреляционной функции.

Ограничимся рассмотрением одномерных, действительных или комплексных, случайных процессов $(\xi(t), t \in \Delta)$, для которых $M|\xi(t)|^2 < \infty$, $t \in \Delta$. Функция $a(t) = M\xi(t)$, $t \in \Delta$, называется *математическим ожиданием* случайного процесса. Случайный процесс $(\xi^0(t), t \in \Delta)$, где $\xi^0(t) = \xi(t) - a(t)$, называется *центрированным случайным процессом*. Функция двух переменных $K_\xi(t, \tau) = K(t, \tau) = M\xi^0(t)\xi^0(\tau)$, $t \in \Delta$, $\tau \in \Delta$, называется *корреляционной функцией* случайного процесса $(\xi(t), t \in \Delta)$. Для действительного процесса $K(t, \tau) = M\xi^0(t)\xi^0(\tau)$. Функция $D(t) = M|\xi^0(t)|^2 = K(t, t)$ называется *дисперсией* случайного процесса $(\xi(t), t \in \Delta)$. Пусть $(\xi(t), t \in \Delta)$, $(\eta(t), t \in \Delta)$ — случайные процессы. Функция $K_{\xi\eta}(t, \tau) = M\xi^0(t)\eta^0(\tau)$ называется *взаимной корреляционной функцией* этих процессов. На основании неравен-

ства Коши — Буняковского

$$|K(t, \tau)|^2 \leq D(t) D(\tau). \quad (1)$$

Основное свойство корреляционной функции — *неотрицательная определенность*; для любых n , $t_k \in \Delta$, и любых комплексных z_1, \dots, z_n выполняется неравенство

$$\sum_{k,j=1}^n K(t_k, t_j) z_k \bar{z}_j \geq 0. \quad (2)$$

В самом деле,

$$\begin{aligned} \sum_{k,j=1}^n K(t_k, t_j) z_k \bar{z}_j &= M \sum_{k=1}^n \xi^0(t_k) z_k \sum_{j=1}^n \overline{\xi^0(t_j)} \bar{z}_j = \\ &= M \left| \sum_{k=1}^n \xi^0(t_k) z_k \right|^2 \geq 0. \end{aligned}$$

Множество Δ чаще всего бывает отрезком (конечным или бесконечным) или бесконечной в одну или обе стороны последовательностью равноотстоящих моментов времени. В первом случае $\xi(t)$ называется *случайным процессом с непрерывным временем*, в последнем — *с дискретным временем (случайной последовательностью)*. Случайный процесс с дискретным временем часто получается в результате периодического измерения значений процесса с непрерывным временем (например, температуры в данной точке).

Пример. Пусть $X(t)$, $t \geq 0$ — процесс Пуассона с параметром λ (см. § 4.5). Тогда $a(t) = M X(t) = \lambda t$. Возьмем t и τ , для которых $0 \leq t \leq \tau$. Имеем

$$X^0(t) X^0(\tau) = X^0(t) (X^0(t) + \eta^0),$$

где $\eta^0 = X^0(\tau) - X^0(t)$. Отсюда

$$M X^0(t) X^0(\tau) = M (X^0(t))^2 + M X^0(t) \eta^0.$$

По свойству закона Пуассона первое слагаемое равно λt . Поскольку по свойству процесса Пуассона $X(t)$ и η независимы, то второе слагаемое равно 0. Итак, $K(t, \tau) = \lambda t$ при $0 \leq t \leq \tau$. Аналогично $K(t, \tau) = \lambda \tau$ при $0 \leq \tau \leq t$.

Обозначим $\eta = \xi(t) - \xi(\tau)$, $b = a(t) - a(\tau)$. Тогда $\eta = b + \eta^0$, $M |\eta|^2 = M (b + \eta^0) (\bar{b} + \bar{\eta}^0) = |b|^2 + M |\eta^0|^2$. Далее, $M |\eta^0|^2 = M (\xi^0(t) - \xi^0(\tau)) (\bar{\xi}^0(t) - \bar{\xi}^0(\tau)) = K(t, t) - K(t, \tau) - K(\tau, t) + K(\tau, \tau)$. Следовательно, для того чтобы $M |\eta|^2 \rightarrow 0$ при $t \rightarrow \tau$, достаточно выполнение двух условий: непрерывности $a(t)$ в точке $t = \tau$ и непрерывности $K(t_1, t_2)$ в точке (τ, τ) .

Случайный процесс $\xi(t)$ с непрерывным временем называется *непрерывным в среднем квадратическом*, если $M |\xi(t) - \xi(\tau)|^2 \rightarrow 0$ при $t \rightarrow \tau$, $\tau \in \Delta$, или, что то же самое, $\xi(t)^2 \rightarrow \xi(\tau)^2$ при $t \rightarrow \tau$. Видим, что достаточные условия непрерывности в среднем квадратическом состоят в непрерывности $a(t)$ и $K(t, \tau)$.

Эти условия не только достаточны, но и необходимы; предлагаем читателю доказать это самостоятельно.

Пусть функция $K(t, \tau)$ интегрируема при $a \leq t \leq b$, $a \leq \tau \leq b$. Для любой непрерывной комплексной функции $z(t)$ имеем

$$\int_a^b \int_a^b K(t, \tau) z(t) \overline{z(\tau)} dt d\tau \geq 0, \quad (3)$$

так как левая часть (2) есть интегральная сумма этого интеграла.

§ 2. Операции над процессами

Пусть $\xi(t)$ — случайный процесс с непрерывным временем, $f(t)$ — функция, непрерывная на отрезке $[a; b]$. Рассмотрим интегральную сумму

$$s_n = \sum_{k=1}^n f(\tau_k) \xi(\tau_k) \Delta t_k,$$

где Δt_k — длины интервалов разбиения отрезка $[a; b]$, n — число интервалов, τ_k — произвольная точка k -го интервала. Если при $\max \Delta t_k \rightarrow 0$, $n \rightarrow \infty$ существует предел s_n в среднем квадратическом, не зависящий от разбиения и точек τ_k , то этот предел

$$s = \lim_{n \rightarrow \infty} s_n \text{ (с. к.)}$$

называется *стохастическим интегралом* и обозначается

$$s = \int_a^b f(t) \xi(t) dt.$$

Выясним условия существования стохастического интеграла. Пусть s_n , s_m — две интегральные суммы. Можно записать

$$s_n - s_m = \sum_{k=1}^N (f(\tau_k) \xi(\tau_k) - f(\tau'_k) \xi(\tau'_k)) \Delta t_k, \quad (4)$$

где Δt_k — интервалы, образованные наложением n -го и m -го разбиений. Очевидно,

$$|s_n - s_m|^2 = (s_n - s'_m)(\bar{s}_n - \bar{s}'_m).$$

Умножив левую и правую части (4) на сопряженную величину и взяв математическое ожидание, получим четыре суммы: две со знаком «+» и две со знаком «—». Например, одна из них имеет вид

$$\begin{aligned} M \sum_{k=1}^N \sum_{j=1}^N f(\tau_k) \xi(\tau_k) \overline{f(\tau_j) \xi(\tau_j)} \Delta t_k \Delta t_j = \\ = \sum_{k=1}^N \sum_{j=1}^N f(\tau_k) \overline{f(\tau_j)} M \xi(\tau_k) \overline{\xi(\tau_j)} \Delta t_k \Delta t_j. \end{aligned}$$

Заметив, что $M\xi(\tau_k)\overline{\xi(\tau_j)} = K_\xi(\tau_k, \tau_j) + a(\tau_k)a(\tau_j)$, получим интегральную сумму интеграла:

$$\int_a^b \int_a^b (K_\xi(t, \tau) + a(t)a(\tau)) f(t) \overline{f(\tau)} dt d\tau. \quad (5)$$

Это же свойство выполняется и для трех остальных сумм. Таким образом, если интеграл (5) существует, то в пределе произойдет взаимное уничтожение четырех сумм, откуда $|s_n - s_m| \rightarrow 0$ (с. к.). Существование предела s следует из теоремы о полноте (см. § 5.1).

Полученное условие существования стохастического интеграла легко преобразуется к условию существования следующих двух интегралов:

$$\int_a^b f(t) a(t) dt, \quad \int_a^b \int_a^b K_\xi(t, \tau) f(t) \overline{f(\tau)} dt d\tau.$$

Если существуют пределы этих интегралов при $a \rightarrow -\infty$, $b \rightarrow \infty$, то стохастический интеграл по интервалу $(a; b)$ сходится в среднем квадратическом к случайной величине, которую обозначают

$$s = \int_{-\infty}^{\infty} f(t) \xi(t) dt$$

и называют *стохастическим интегралом (в бесконечных пределах)*. Можно записать

$$s = s^0 + \int_{-\infty}^{\infty} f(t) a(t) dt.$$

Для дисперсии s имеем формулу

$$M|s^0|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_\xi(t, \tau) f(t) \overline{f(\tau)} dt d\tau. \quad (6)$$

Если функция f , участвующая в определении стохастического интеграла, зависит от параметра, этот интеграл определяет случайный процесс по этому параметру в качестве времени. Так, можно рассмотреть случайный процесс

$$\eta(t) = \int_{-\infty}^{\infty} f(t, \tau) \xi(\tau) d\tau. \quad (7)$$

Его характеристики второго порядка

$$M\eta(t) = \int_{-\infty}^{\infty} f(t, \tau) a(\tau) d\tau, \quad (8)$$

$$\begin{aligned} K_\eta(t, t') &= M \int_{-\infty}^{\infty} f(t, \tau) \xi^0(\tau) d\tau \int_{-\infty}^{\infty} \overline{f(t', \tau')} \overline{\xi^0(\tau')} d\tau' = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(\tau, \tau') f(t, \tau) \overline{f(t', \tau')} d\tau d\tau'. \end{aligned} \quad (9)$$

Приведенные формулы обосновываются предельным переходом от интегральных сумм.

Пусть $\xi(t)$ — случайный процесс с непрерывным временем. Если существует $\eta(t) = \lim \Delta\xi/\Delta t$ (с. к.) при $\Delta t \rightarrow 0$, то этот предел называется *среднеквадратической производной* $\xi(t)$ и обозначается $\xi'(t)$. По теореме о полноте § 5.1 для существования $\xi'(t)$ достаточно, чтобы выражение

$$\zeta = \frac{1}{u} (\xi(t+u) - \xi(t)) - \frac{1}{v} (\xi(t+v) - \xi(t)) \quad (10)$$

сходилась к 0 в среднем квадратическом при $u \rightarrow 0, v \rightarrow 0$. Из (10) получаем аналогичное выражение для ξ^0 , заменив всюду ξ на ξ^0 . Умножив (10) на сопряженное ему равенство, получим $M|\zeta^0|^2$ в виде суммы четырех слагаемых, из которых два со знаком «+», два со знаком «-». Все они оцениваются одинаково, поэтому рассмотрим одно из них:

$$\begin{aligned} & M \frac{1}{uv} (\xi^0(t+u) - \xi^0(t)) \overline{(\xi^0(t+v) - \xi^0(t))} = \\ & = \frac{1}{uv} [K_{\xi}(t+u, t+v) - K_{\xi}(t+u, t) - K_{\xi}(t, t+v) + K_{\xi}(t, t)]. \end{aligned}$$

В квадратных скобках — интеграл функции $\partial^2 K_{\xi}(x, y)/\partial x \partial y$ по области $t < x < t+u, t < y < t+v$. Если указанная вторая производная непрерывна в точке (t, t) , то все выражение при $u \rightarrow 0, v \rightarrow 0$ эквивалентно $\partial^2 K_{\xi}(x, y)/\partial x \partial y$ при $x = y = t$. В таком случае $M|\zeta^0|^2 \rightarrow 0$ при $u \rightarrow 0, v \rightarrow 0$. Поскольку

$$M|\zeta|^2 = \left| \frac{1}{u} (a(t+u) - a(t)) - \frac{1}{v} (a(t+v) - a(t)) \right|^2 + M|\zeta^0|^2,$$

то для существования $\xi'(t)$ достаточны следующие два условия:

- 1) $\partial^2 K(x, y)/\partial x \partial y$ существует и непрерывна при $x = y = t$;
- 2) существует $a'(t)$.

Пусть $\xi(t), \eta(t)$ — случайные процессы с непрерывным временем. Справедливы формулы

$$K_{\xi'\eta}(t, \tau) = M(\xi'(t))^0 \overline{\eta^0(\tau)} = \frac{\partial K_{\xi\eta}(t, \tau)}{\partial t}, \quad (11)$$

$$K_{\xi\eta'}(t, \tau) = M\xi^0(t) \overline{(\eta'(\tau))^0} = \frac{\partial K_{\xi\eta}(t, \tau)}{\partial \tau}, \quad (12)$$

$$K_{\xi'\eta'}(t, \tau) = \frac{\partial^2 K_{\xi\eta}(t, \tau)}{\partial t \partial \tau}, \quad (13)$$

$$K_{\xi'}(t, \tau) = \frac{\partial^2 K_{\xi}(t, \tau)}{\partial t \partial \tau}. \quad (14)$$

Формула (11) следует из того, что

$$\begin{aligned} M(\xi'(t))^0 \overline{\eta^0(\tau)} &= \lim_{u \rightarrow 0} M \frac{1}{u} (\xi^0(t+u) - \xi^0(t)) \overline{\eta^0(\tau)} = \\ &= \lim_{u \rightarrow 0} \frac{1}{u} (K_{\xi\eta}(t+u, \tau) - K_{\xi\eta}(t, \tau)) = \frac{\partial K_{\xi\eta}(t, \tau)}{\partial t}. \end{aligned}$$

Формула (12) симметрична (11), (13) сводится к последовательному использованию двух предыдущих, а (14) — частный случай (13) при $\eta = \xi$.

§ 3. Случайный процесс, стационарный в широком смысле

Пусть $(\xi(t), -\infty < t < \infty)$ — комплексный случайный процесс, у которого $M\xi(t) = a = \text{const}$, а корреляционная функция зависит лишь от разности аргументов: $M\xi^0(t)\overline{\xi^0(\tau)} = K_\xi(t - \tau)$. Такой процесс называется *стационарным в широком смысле*.

Если $\xi(t), \eta(t)$ — два комплексных случайных процесса, для которых $M\xi^0(t)\eta^0(\tau) = K_{\xi\eta}(t - \tau)$, то эти процессы называются *стационарно связанными*.

Из определенной стационарности и стационарной связанности комплексных процессов следует как частные случаи определения для действительных процессов: следует лишь убрать знак комплексного сопряжения.

Пример 1. Пусть α_k, β_k — действительные случайные величины, удовлетворяющие условиям $M\alpha_k = M\beta_k = 0$; $M\alpha_k\beta_j = 0$ для всех k, j ; $M\alpha_k\alpha_j = M\beta_k\beta_j = 0$ для $k \neq j$; $M\alpha_k^2 = M\beta_k^2 = \sigma_k^2$. Рассмотрим действительный случайный процесс $\xi(t)$ вида

$$\xi(t) = \sum_{k=1}^n \alpha_k \cos \omega_k t + \beta_k \sin \omega_k t, \quad (1)$$

где ω_k — различные положительные числа. Имеем

$$M\xi(t) = \sum_{k=1}^n (M\alpha_k) \cos \omega_k t + (M\beta_k) \sin \omega_k t = 0,$$

$$\begin{aligned} M\xi^0(t)\xi^0(\tau) &= M\xi(t)\xi(\tau) = M \sum_{k=1}^n (\alpha_k \cos \omega_k t + \dots) \sum_{j=1}^n (\alpha_j \cos \omega_j \tau + \dots) = \\ &= \sum_{k,j=1}^n (M\alpha_k\alpha_j) \cos \omega_k t \cos \omega_j \tau + \dots \end{aligned}$$

В силу некоррелированности α_k, β_j в данной сумме слагаемые с $M\alpha_k\beta_j$ при всех k, j , а также с $M\alpha_k\alpha_j$ и $M\beta_k\beta_j$ при $k \neq j$ равны нулю; остается лишь следующее выражение:

$$\begin{aligned} M\xi^0(t)\xi^0(\tau) &= \sum_{k=1}^n (M\alpha_k^2 \cos \omega_k t \cos \omega_k \tau + M\beta_k^2 \sin \omega_k t \sin \omega_k \tau) = \\ &= \sum_{k=1}^n \sigma_k^2 \cos \omega_k (t - \tau). \end{aligned}$$

Имеем действительный процесс, стационарный в широком смысле; его корреляционная функция

$$K_\xi(t) = \sum_{k=1}^n \sigma_k^2 \cos \omega_k t. \quad (2)$$

Пример 2. Пусть ζ_k — комплексные случайные величины; $M\zeta_k = 0$, $M\zeta_k \bar{\zeta}_j = 0$ при $k \neq j$; $M|\zeta_k|^2 = \sigma_k^2$. Рассмотрим комплексный случайный процесс

$$X(t) = \sum_{k=1}^n \zeta_k e^{i\omega_k t}, \quad (3)$$

где ω_k — различные действительные числа. Как и в примере 1, доказывается, что $MX(t) = 0$. Поскольку $\overline{\exp\{i\omega t\}} = \exp\{-i\omega t\}$, то

$$\begin{aligned} MX(t) \overline{X(\tau)} &= \sum_{k,j=1}^n M\zeta_k \bar{\zeta}_j e^{i(\omega_k t - \omega_j \tau)} = \\ &= \sum_{k=1}^n M\zeta_k \bar{\zeta}_k e^{i\omega_k(t-\tau)} = \sum_{k=1}^n \sigma_k^2 e^{i\omega_k(t-\tau)}. \end{aligned}$$

Следовательно, $X(t)$ — комплексный процесс, стационарный в широком смысле; его корреляционная функция

$$K_X(t) = \sum_{k=1}^n \sigma_k^2 e^{i\omega_k t}. \quad (4)$$

Пример 3. Пусть σ — постоянная, φ и ω — независимые случайные величины, причем φ равномерна в интервале $(-\pi; \pi)$, ω имеет функцию распределения $F(x)$. Рассмотрим случайный процесс

$$\zeta(t) = \sigma e^{i(\omega t + \varphi)}. \quad (5)$$

В данном случае

$$M\zeta(t) = \sigma M e^{i\omega t} M e^{i\varphi} = 0,$$

так как $M e^{i\varphi} = \frac{1}{2\pi} \int_0^{2\pi} e^{iz} dz = 0;$

$$M\xi(t) \overline{\xi(\tau)} = M\sigma^2 e^{i\omega(t-\tau)} = \sigma^2 \int e^{i\omega(t-\tau)} dF(x).$$

Корреляционная функция этого стационарного в широком смысле процесса равна $\sigma^2 \varphi(t)$, где $\varphi(t)$ — характеристическая функция ω .

Пример 4. Пусть $\Delta > 0$, $X(t)$ — процесс Пуассона с параметром λ , $\xi(t) = 1$, если $X(t + \Delta) - X(t) > 0$, $\xi(t) = 0$ в противном случае. Поскольку (см. § 4.5) $X(t + \Delta) - X(t)$ — пуассоновская случайная величина с параметром $\lambda\Delta$, то $M\xi(t) = 1 - e^{-\lambda\Delta}$. Если $|t - \tau| \geq \Delta$, то события $\{\xi(t) = 1\}$ и $\{\xi(\tau) = 1\}$ независимы, откуда $\text{cov}(\xi(t), \xi(\tau)) = 0$. Если $|t - \tau| < \Delta$, то событие $\{\xi(t) = 0, \xi(\tau) = 0\}$ означает отсутствие событий потока в интервале длины $s = |t - \tau| + \Delta$, а следовательно, имеет вероятность $e^{-\lambda s}$. Отсюда

$$\begin{aligned} e^{-\lambda s} &= M(1 - \xi(t))(1 - \xi(\tau)) = 1 - M\xi(t) - M\xi(\tau) + M\xi(t)\xi(\tau), \\ \text{cov}(\xi(t), \xi(\tau)) &= M\xi(t)\xi(\tau) - M\xi(t)M\xi(\tau) = e^{-\lambda s} - 1 + M\xi(t) + \\ &+ M\xi(\tau) - M\xi(t)M\xi(\tau) = e^{-\lambda s} - 1 + 2(1 - e^{-\lambda\Delta}) - (1 - e^{-\lambda\Delta})^2 = \\ &= e^{-\lambda s} - e^{-2\lambda\Delta} = e^{-\lambda\Delta}(e^{-\lambda|t-\tau|} - e^{-\lambda\Delta}). \end{aligned}$$

В данном примере $K_\xi(t) = e^{-\lambda\Delta}(e^{-\lambda|t|} - e^{-\lambda\Delta})$ при $|t| < \Delta$, $K_\xi(t) = 0$ при $|t| \geq \Delta$.

Выясним связь между понятиями стационарности (в широком смысле) для действительных и комплексных процессов.

Пусть $\xi(t)$, $\eta(t)$ — стационарные, стационарно связанные действительные процессы. Тогда $\zeta(t) = \xi(t) + i\eta(t)$ — стационарный процесс. Действительно,

$$\begin{aligned} M\zeta^0(t)\overline{\zeta^0(\tau)} &= M\xi^0(t)\xi^0(\tau) + M\eta^0(t)\eta^0(\tau) - iM\xi^0(t)\eta^0(\tau) + \\ &+ iM\eta^0(t)\xi^0(\tau) = K_\xi(t-\tau) + K_\eta(t-\tau) - iK_{\xi\eta}(t-\tau) + \\ &+ iK_{\eta\xi}(t-\tau), \end{aligned}$$

т. е. данная функция зависит только от $t - \tau$. Поскольку многие выкладки с комплексными процессами проще, чем с действительными, часто переходят от действительного процесса к комплексному и после получения окончательной формулы возвращаются к исходному действительному.

§ 4. Анализ корреляционной функции. Эргодичность

Пусть $K(t)$ — корреляционная функция стационарного в широком смысле комплексного случайного процесса $\xi(t)$. Если $\xi(t)$ непрерывен в среднем квадратическом, то $K(t)$ непрерывна. Действительно,

$$K(t + \Delta) - K(t) = M(\xi^0(t + \Delta) - \xi^0(t))\overline{\xi^0(0)},$$

откуда по неравенству Коши — Буняковского

$$|K(t + \Delta) - K(t)|^2 \leq M|\xi^0(t + \Delta) - \xi^0(t)|^2 M|\xi^0(0)|^2 \xrightarrow{\Delta \rightarrow 0} 0.$$

Предположим, что $\int_{-\infty}^{\infty} |K(t)| dt < \infty$. Возьмем любое действительное ω и рассмотрим интеграл

$$f_T(\omega) = \frac{1}{2\pi T} \int_0^T \int_0^T K(t, \tau) e^{-i(t-\tau)\omega} dt d\tau.$$

По формуле (1.3) $f_T(\omega) \geq 0$ (принимая $z(t) = e^{-it\omega}$). Преобразуем данный интеграл, введя новую переменную $x = t - \tau$. При данном значении x переменная t изменяется в интервале длины $T - |x|$; следовательно,

$$f_T(\omega) = \frac{1}{2\pi} \int_{-T}^T K(x) \left(1 - \frac{|x|}{T}\right) e^{-ix\omega} dx.$$

Поскольку по условию $K(x)$ абсолютно интегрируема, то данный интеграл при $T \rightarrow \infty$ сходится к пределу $f(\omega)$, а так как $f_T(\omega) \geq 0$, то и $f(\omega) \geq 0$. Итак,

$$0 \leq f(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(x) e^{-ix\omega} dx. \quad (1)$$

Видим, что $f(\omega)$ — преобразование Фурье функции $K(t)$. По теореме обращения для преобразований Фурье

$$K(x) = \int_{-\infty}^{\infty} f(\omega) e^{i\omega x} d\omega. \quad (2)$$

Формулу (2) называют *спектральным разложением корреляционной функции*; $f(\omega)$ называется *спектральной плотностью* случайного процесса $\xi(t)$. Положив в (2) $x = 0$, получим

$$M|\xi(t)|^2 = K(0) = \int_{-\infty}^{\infty} f(\omega) d\omega; \quad (3)$$

стало быть, $f(\omega)$ интегрируема в $(-\infty; \infty)$.

Если $K(t)$ не является абсолютно интегрируемой в интервале $(-\infty; \infty)$, приведенное рассуждение незаконно, и приходится использовать «обходной путь». Пусть $K(t)$ непрерывна. Рассмотрим функцию

$$K^{(c)}(t) = \begin{cases} K(t) \left(1 - \frac{|t|}{c}\right) & \text{при } |t| \leq c, \\ 0 & \text{при } |t| > c. \end{cases}$$

Эта функция, как можно показать, неотрицательно определена; кроме того, она абсолютно интегрируема в бесконечных пределах. Поэтому вследствие (2)

$$K^{(c)}(t) = \int_{-\infty}^{\infty} f^{(c)}(\omega) e^{i\omega t} d\omega. \quad (4)$$

Обозначив $F^{(c)}(\omega) = \int_{-\infty}^{\omega} f^{(c)}(\omega') d\omega'$, можем переписать (4) в виде

$$K^{(c)}(t) = \int_{-\infty}^{\infty} e^{i\omega t} dF^{(c)}(\omega). \quad (5)$$

При $c \rightarrow \infty$ левая часть (5) сходится к $K(t)$ равномерно в любом конечном интервале. Доказывается (в общих чертах так же, как обратная теорема непрерывности § 5.3), что $F^{(c)}(\omega) \xrightarrow{w} F(\omega)$, где $F(\omega)$ — неубывающая функция, $F(-\infty) = 0$, $F(\infty) = K(0)$, и выполняется соотношение

$$K(t) = \int_{-\infty}^{\infty} e^{i\omega t} dF(\omega). \quad (6)$$

Формула (6) называется *формулой Хинчина*. Она представляет собой *спектральное разложение* интегрируемой корреляционной функции.

Заметим, что при выводе формулы (6) использовались лишь следующие свойства функции $K(t)$: непрерывность и неотрицательная определенность. Теорема, согласно которой для функции с указан-

ными свойствами выполняется спектральное разложение (6) с неубывающей неограниченной функцией $F(\omega)$, называется *теоремой Хинчина*.

Возвратимся теперь к случаю, когда $K(t)$ — корреляционная функция стационарного процесса.

Функция $F(\omega)$ называется *спектральной функцией* случайного процесса $\xi(t)$. Если существует *спектральная плотность* $f(\omega)$, то $F(\omega)$ называется *абсолютно непрерывной спектральной функцией*; в этом случае $F(\omega) = \int_{-\infty}^{\omega} f(\omega') d\omega'$, а в интеграле (6) можно положить $dF(\omega) = f(\omega) d\omega$. Случайный процесс со спектральной плотностью называется *процессом с непрерывным спектром*. Если $F(\omega)$ возрастает лишь в конечном или счетном множестве точек ω_k , причем $F(\omega_k + 0) - F(\omega_k - 0) = \sigma_k^2$, то в соответствии с (6)

$$K(t) = \sum_k \sigma_k^2 e^{i\omega_k t}; \quad (7)$$

такой процесс называется *процессом с дискретным спектром*. Заметим, что $K(t)$ вида (7) нам уже встречалась в примере 2 § 3: σ_k^2 равна дисперсии амплитуды комплексной гармоники с частотой ω_k . Для процессов с дискретным спектром

$$F(\omega) = \sum_{\omega_k < \omega} \sigma_k^2. \quad (8)$$

В различных приложениях (к механике, электротехнике, радиотехнике) квадрат амплитуды комплексной гармоники измеряет мощность рассеяния энергии на данной частоте. Поэтому $dF(\omega)$ измеряет среднюю мощность колебаний с частотами интервала $d\omega$. Этим объясняется используемое в прикладных дисциплинах название «*энергетический спектр*» для спектральной плотности. Суммарная мощность всех колебаний есть $K(0) = \int_{-\infty}^{\infty} dF(\omega) = F(\infty)$.

Точный смысл этого понятия будет раскрыт в § 5.

Для действительного стационарного в широком смысле процесса формулу (2) можно записать в виде

$$K(x) = \int_{-\infty}^{\infty} f(\omega) \cos \omega x d\omega + i \int_{-\infty}^{\infty} f(\omega) \sin \omega x d\omega.$$

Поскольку $K(x)$ действительна, второе слагаемое равно 0, а так как косинус — четная функция, то

$$K(x) = \int_0^{\infty} (f(\omega) + f(-\omega)) \cos \omega x d\omega.$$

Видим, что положительные и отрицательные частоты различать нет смысла. Положив

$$f(\omega) + f(-\omega) = g(\omega), \quad (9)$$

получим

$$K(x) = \int_0^{\infty} g(\omega) \cos \omega x d\omega. \quad (10)$$

Функция $g(\omega)$ называется *спектральной плотностью* действительного случайного процесса. Подобным же образом можно преобразовать и *формулу Хинчина* (6):

$$K(t) = \int_0^{\infty} \cos \omega t dG(\omega), \quad (11)$$

где

$$G(\omega) = F(\omega) - F(-\omega) \quad (12)$$

называется *спектральной функцией* действительного случайного процесса. Для уточнения, к какому процессу относятся введенные характеристики, пишут $f(\omega) = f_{\xi}(\omega)$, $F(\omega) = F_{\xi}(\omega)$ и т. п., ξ — символ случайного процесса $\xi(t)$.

Из формулы (11) видно, что для действительного процесса $K(t)$ — четная функция. Обратное косинус-преобразование, примененное к (10), приводит к формуле

$$g(\omega) = \frac{2}{\pi} \int_0^{\infty} K(t) \cos \omega t dt. \quad (13)$$

Для вычисления $f(\omega)$, $g(\omega)$ по $K(t)$ и наоборот пользуются таблицей преобразований Фурье, а при «неберущихся» интегралах — численными методами, использующими различные варианты сглаживания функций.

Пример 1. Шумовой процесс, прошедший через частотный фильтр радиоприемного устройства, имеет энергетический спектр $g(\omega) = c$ при $a < \omega < b$; $g(\omega) = 0$ при $\omega \leq a$ и при $\omega \geq b$.

По формуле (10)

$$K(t) = c \int_a^b \cos \omega t d\omega = c (\sin bt - \sin at)/t.$$

В частности, при $a = -\Delta/2$, $b = \Delta/2$

$$K(t) = c\Delta \sin(\Delta t/2)/(\Delta t/2).$$

Предел $K(t)$ при $t \rightarrow 0$ равен $c(b-a)$, т. е. суммарная мощность равна площади под кривой $g(\omega)$.

Пример 2. Задана корреляционная функция

$$K(t) = ce^{-\alpha|t|} \cos \omega_0 t.$$

Имеем

$$g(\omega) = \frac{c\alpha}{\pi} \left(\frac{1}{\alpha^2 + (\omega - \omega_0)^2} + \frac{1}{\alpha^2 + (\omega + \omega_0)^2} \right).$$

В частности, при $\omega_0 = 0$

$$g(\omega) = \frac{2c\alpha}{\pi} \frac{1}{\alpha^2 + \omega^2}, \quad f(\omega) = \frac{c\alpha}{\pi} \frac{1}{\alpha^2 + \omega^2}.$$

Этот результат легко получаем с помощью вычетов: например, при $\omega_0 = 0$

$$K(t) = \frac{c\alpha}{\pi} \int_{-\infty}^{\infty} e^{it\omega} (\alpha^2 + \omega^2)^{-1} d\omega$$

есть интеграл комплексной функции $\omega(z) = (c\alpha/\pi) e^{itz}/(\alpha^2 + z^2)$ по действительной оси комплексной плоскости. При $t > 0$ экспонента $e^{izt} = e^{it(x+iy)} = e^{-ty} \times (\cos tx + i \sin tx)$ убывает к нулю при $y \rightarrow \infty$. Функция $\omega(z)$ имеет в верхней полуплоскости полюс в точке $z = i\alpha$; ее вычет в данной точке $(c\alpha/\pi) e^{itz}/2z = ce^{-\alpha^2/2\pi i}$.

Запишем тождество

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-\sigma^2 t^2/2} dt,$$

выражающее связь между плотностью нормального распределения с параметрами 0, σ^2 и его характеристической функцией. Заменяв x на $(-t)$, t на ω , получим тождество

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2} = \int_{-\infty}^{\infty} e^{it\omega} \frac{1}{2\pi} e^{-\sigma^2 \omega^2/2} d\omega, \quad (14)$$

означающее, что корреляционной функции

$$K_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2}$$

соответствует спектральная плотность

$$f_\sigma(\omega) = \frac{1}{2\pi} e^{-\sigma^2 \omega^2/2}.$$

При $\sigma \rightarrow 0$ функция $K_\sigma(t)$ является δ -образной, т. е.

- 1) $K_\sigma(t) \geq 0$ при всех t ;
- 2) $K_\sigma(t) \rightarrow 0$ при всех $t \neq 0$;
- 3) $\int_{-\varepsilon}^{\varepsilon} K_\sigma(t) dt \rightarrow 0$, $\int_{\varepsilon}^{\infty} K_\sigma(t) dt \rightarrow 0$ при любом $\varepsilon > 0$;
- 4) $\int_{-\varepsilon}^{\varepsilon} K_\sigma(t) dt \rightarrow 1$ при любом $\varepsilon > 0$.

Эти свойства нестрого выражают соотношением

$$K_\sigma(t) \xrightarrow{\sigma \rightarrow 0} \delta(t),$$

где $\delta(t)$ — дельта-функция Дирака.

В то же время при $\sigma \rightarrow 0$ для любого ω

$$f_\sigma(\omega) \xrightarrow{\sigma \rightarrow 0} \frac{1}{2\pi}. \quad (15)$$

Полученные предельные соотношения можно формально записать в виде равенств

$$\delta(t) = \int_{-\infty}^{\infty} e^{i\omega t} \frac{1}{2\pi} d\omega, \quad (16)$$

$$\frac{1}{2\pi} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \delta(t) dt, \quad (17)$$

считая, что некоей идеальной корреляционной функции $\delta(t)$ соответствует постоянная на всех частотах ω спектральная плотность $1/2\pi$. Подобных процессов в природе не существует, так как мощность $\int_{-\infty}^{\infty} f(\omega) d\omega = \int_{-\infty}^{\infty} \frac{1}{2\pi} d\omega = \infty$. Однако введение такого иде-

ального (обобщенного) процесса, называемого *белым шумом*, упрощает многие выкладки и интерпретации формул. Вполне корректное понимание белого шума состоит в следующем.

Пусть $\eta_{\sigma}(t)$ — случайный процесс с δ -образной корреляционной функцией, $\xi_{\sigma}(t) = (L\eta_{\sigma})(t)$ — некоторое преобразование данного процесса в процесс с корреляционной функцией $R_{\sigma}(t)$, причем $R_{\sigma}(t) \xrightarrow{\sigma \rightarrow 0} R(t)$, где $R(t)$ — корреляционная функция стационарного процесса. Тогда вместо того, чтобы находить $R_{\sigma}(t)$ и переходить к пределу при $\sigma \rightarrow 0$, можно сразу, пользуясь формальными правилами, применить оператор L к белому шуму с корреляционной функцией $\delta(t)$. Использование белого шума как бы вбирает в себя трудность предельного перехода.

Применим спектральное разложение $K(t)$ к получению признака эргодичности процесса.

Стационарный в широком смысле случайный процесс $\xi(t)$ называется *эргодическим*, если

$$\frac{1}{T} \int_0^T \xi(t) dt \xrightarrow{T \rightarrow \infty} a \text{ (с. к.)},$$

где $a = M\xi(0)$. Очевидно, данное свойство эквивалентно соотношению

$$\frac{1}{T} \int_0^T \xi^0(t) dt \xrightarrow{T \rightarrow \infty} 0 \text{ (с. к.)}.$$

Предполагая, что $K(t)$ непрерывна, получаем

$$\begin{aligned} M \left| \frac{1}{T} \int_0^T \xi^0(t) dt \right|^2 &= \frac{1}{T^2} \int_0^T \int_0^T M \xi^0(t) \overline{\xi^0(\tau)} dt d\tau = \\ &= \frac{1}{T^2} \int_0^T \int_0^T K(t - \tau) dt d\tau = \frac{1}{T} \int_{-T}^T \left(1 - \frac{|x|}{T}\right) K(x) dx. \end{aligned}$$

Стремление последнего выражения к нулю при $T \rightarrow \infty$, очевидно, необходимо и достаточно для эргодичности процесса. (В частности, процесс эргодичен при $K(t) \xrightarrow{t \rightarrow \infty} 0$.) Заменяв в этом интеграле $K(x)$ ее спектральным разложением, а затем изменив порядок интегрирования, получим

$$\frac{1}{T} \int_{-T}^T \left(1 - \frac{|x|}{T}\right) K(x) dx = \int_{-\infty}^{\infty} \left[\frac{1}{T} \int_{-T}^T e^{i\omega x} \left(1 - \frac{|x|}{T}\right) dx \right] dF(\omega).$$

Выражение в квадратных скобках по модулю не превосходит 2 при любом ω и в то же время при $|\omega| \geq \delta$ равномерно стремится к 0 при $T \rightarrow \infty$. Интегрированием по частям оно сводится к выражению $\frac{2}{\omega T} \sin \omega T - \frac{2}{\omega^2 T^2} (1 - \cos \omega T)$. Значит, в пределе при $T \rightarrow \infty$ от всего первоначального выражения остается величина, сравнимая с $\int_{-\delta}^{\delta} dF(\omega)$. Если $F(\omega)$ непрерывна в нуле, то эта величина

сколь угодно мала. Итак, достаточным признаком эргодичности процесса является непрерывность $F(\omega)$ в точке $\omega = 0$. В частности, если процесс имеет спектральную плотность, то он эргодичен. Сумма гармонических колебаний эргодична, если частоты всех этих колебаний ненулевые.

§ 5. Спектральное разложение стационарного процесса

Начнем с простейшего примера. Пусть комплексный случайный процесс, стационарный в широком смысле, имеет вид

$$\xi(t) = \sum_{k=0}^n \gamma_k e^{it\omega_k}, \quad (1)$$

где γ_k — комплексные случайные величины, ω_k — различные действительные числа, причем $\omega_0 = 0$.

Имеем

$$M\xi(t) = \sum_{k=0}^n (M\gamma_k) e^{it\omega_k}. \quad (2)$$

Если хотя бы одно из чисел $M\gamma_1, \dots, M\gamma_n$ отлично от нуля, то функция $M\xi(t)$ не может быть постоянной. Итак,

$$M\gamma_k = 0, \quad 1 \leq k \leq n, \quad (3)$$

а следовательно,

$$a = M\xi(t) = M\gamma_0.$$

Отсюда

$$\xi^0(t) = \sum_{k=1}^n M\gamma_k e^{it\omega_k}. \quad (4)$$

Корреляционная функция

$$K(\tau) = \sum_{k,j=1}^n M\gamma_k \bar{\gamma}_j e^{it(\omega_k - \omega_j) + i\tau\omega_k}. \quad (5)$$

Это выражение не зависит от t ; стало быть, не зависят от t коэффициенты правой части (5) при $e^{it\omega_k}$, т. е.

$$\sum_{k=1}^n M\gamma_k \bar{\gamma}_j e^{it(\omega_k - \omega_j)}.$$

Очевидно, это возможно лишь в случае, если

$$M\gamma_k \bar{\gamma}_j = 0, \quad k \neq j. \quad (6)$$

Итак, если случайный процесс разложен по комплексным гармоникам в виде (1), то коэффициенты разложения — ортогональные случайные величины (условие ортогональности и есть (6)) с нулевым при частотах $\omega_k \neq 0$ математическим ожиданием. Из (5) получаем

$$K(\tau) = \sum_{k=1}^n \sigma_k^2 e^{it\omega_k}, \quad (7)$$

что совпадает с формулой (3.4).

Построим случайный процесс $Z(\omega)$, переменная которого есть не время t , а частота ω , по формуле

$$Z(\omega) = \sum_{\omega_k < \omega} \gamma_k \quad (8)$$

и отметим следующие его свойства:

$$1) MZ(\omega) = 0; \quad (9)$$

$$2) M|Z(\omega') - Z(\omega)|^2 = F(\omega') - F(\omega) \quad (10)$$

при любых $\omega \leq \omega'$;

$$3) M(Z(\omega') - Z(\omega))(\overline{Z(\omega'') - Z(\omega')}) = 0 \quad (11)$$

для любых непересекающихся интервалов $(\omega; \omega')$ и $(\omega''; \omega''')$;

$$4) M|Z(\omega)|^2 \xrightarrow{\omega \rightarrow -\infty} 0. \quad (12)$$

Равенство (10) доказывается так. Поскольку

$$Z(\omega') - Z(\omega) = \sum_{\omega \leq \omega_k < \omega'} \gamma_k,$$

то квадрат модуля этого выражения есть сумма $\gamma_k \bar{\gamma}_j$ по значениям ω_k из полуинтервала $[\omega; \omega')$. Переходя к математическим ожиданиям, получим $M\gamma_k \bar{\gamma}_j = 0$ для различных k и j , $M|\gamma_k|^2 = \sigma_k^2$. Отсюда левая часть (10) равна

$$\sum_{\omega \leq \omega_k < \omega'} \sigma_k^2 = F(\omega') - F(\omega),$$

что и требовалось доказать. Аналогично доказывается (11). С помощью процесса $Z(\omega)$ из формулы (11) выводим представление

$$\xi(t) = a + \int_{-\infty}^{\infty} e^{it\omega} dZ(\omega). \quad (13)$$

На этом рассмотрение примера заканчивается.

Пусть теперь $F(\omega)$ — произвольная неубывающая, непрерывная слева, ограниченная функция, стремящаяся к нулю при $\omega \rightarrow -\infty$, и $(Z(\omega), \omega \in \mathbb{R})$ — набор случайных величин, обладающих свойствами (9) — (12). Тогда $Z(\omega)$ как функция ω есть случайный процесс, называемый *процессом с ортогональными приращениями*.

Дадим определение *стохастического интеграла*

$$I_u = \int u(\omega) dZ(\omega).$$

Если функция $u(\omega)$ равна c при $\omega' \leq \omega < \omega''$ и равна 0 в противном случае, полагаем $I_u = c(Z(\omega'') - Z(\omega'))$. Если $u_k(\omega)$, $1 \leq k \leq n$, таким образом определены и $u(\omega) = u_1(\omega) + \dots + u_n(\omega)$, полагаем $I_u = I_{u_1} + \dots + I_{u_n}$. Если, наконец, дана любая функция $u(\omega)$, удовлетворяющая условию

$$\int_{-\infty}^{\infty} |u(\omega)|^2 dF(\omega) < \infty, \quad (14)$$

и I_{v_n} определены, причем

$$\int_{-\infty}^{\infty} |u(\omega) - v_n(\omega)|^2 dF(\omega) \xrightarrow{n \rightarrow \infty} 0, \quad (15)$$

то полагаем

$$I_u = \lim_{n \rightarrow \infty} I_{v_n} \text{ (с. к.)}. \quad (16)$$

Оказывается, что таким образом I_u можно определить для любой функции $u(\omega)$, удовлетворяющей условию (14). При этом

$$MI_u = 0, \quad M|I_u|^2 = \int_{-\infty}^{\infty} |u(\omega)|^2 dF(\omega), \quad (17)$$

$$MI_u \bar{I}_v = \int_{-\infty}^{\infty} u(\omega) \overline{v(\omega)} dF(\omega). \quad (18)$$

Положив $u(\omega) = e^{it\omega}$, получим случайную величину

$$\xi(t) = \int_{-\infty}^{\infty} e^{it\omega} dZ(\omega). \quad (19)$$

По формуле (18) находим $M\xi(t) = 0$;

$$M\xi(t + \tau) \overline{\xi(t)} = \int_{-\infty}^{\infty} e^{i(t+\tau)\omega} e^{-it\omega} dF(\omega) = \int_{-\infty}^{\infty} e^{i\tau\omega} dF(\omega). \quad (20)$$

Итак, $\xi(t)$ — стационарный в широком смысле случайный процесс. Из (20) следует, что его корреляционная функция

$$K(\tau) = \int_{-\infty}^{\infty} e^{i\tau\omega} dF(\omega), \quad (21)$$

т. е. $F(\omega)$ — спектральная функция данного процесса. Данная функция непрерывна. Полагая $\tau = 0$, получаем равенство

$$D\xi(t) = K(0) = F(\infty). \quad (22)$$

Методом гильбертова пространства доказывается, что любой стационарный в широком смысле случайный процесс, удовлетворяющий условию непрерывности в среднем квадратическом

$$\xi(t) \xrightarrow{t \rightarrow \tau} \xi(\tau) \text{ (с. к.)},$$

допускает представление

$$\xi(t) = a + \int_{-\infty}^{\infty} e^{it\omega} dZ(\omega), \quad (23)$$

где $a = M\xi(t)$, $Z(\omega)$ — процесс с ортогональными приращениями, удовлетворяющий условиям (9) — (12).

Опуская подробности доказательства, отметим лишь главные его моменты. Значения процесса $\xi(t)$, их линейные комбинации и среднеквадратические пределы таких комбинаций можно представить в виде точек гильбертова пространства H со скалярным произведением $(\alpha, \beta) = M\alpha\bar{\beta}$ для любых двух элементов α, β этого пространства. Норма $\|\alpha\|$ точки α определяется как $(\alpha, \alpha)^{1/2}$. Сходимость $\alpha_n \rightarrow \beta$ означает, что $\|\alpha_n - \beta\| \rightarrow 0$. Поставим в соответствие величине $\xi(t)$ функцию $e^{it\omega}$ переменной ω , величине $c_1\xi(t_1) + \dots + c_n\xi(t_n)$ — функцию $c_1e^{it_1\omega} + \dots + c_ne^{it_n\omega}$. В пространстве функций $u(\omega)$, удовлетворяющих условию (14), введем скалярное произведение

$$(u, v) = \int_{-\infty}^{\infty} u(\omega) \overline{v(\omega)} dF(\omega). \quad (24)$$

Заметим, что если величинам α и β соответствуют функции $u(\omega)$ и $v(\omega)$, то $(\alpha, \beta) = (u, v)$. Этим устанавливается взаимно однозначное изометрическое (сохраняющее расстояния и «углы») соответствие между пространством H и гильбертовым пространством H^* функций, для которых выполняется условие (14).

Искомый процесс $Z(\omega)$ определяется тем свойством, что при любом ω_0 случайная величина $Z(\omega_0)$ определяется как точка H , соответствующая функции $u(\omega)$, равной единице при $\omega < \omega_0$ и нулю при $\omega \geq \omega_0$.

Широкое применение находит тот случай, когда спектральное разложение случайного процесса имеет более простой вид по сравнению с (23), а именно:

$$\xi(t) = a + \int_{-\infty}^{\infty} e^{it\omega} z(\omega) d\omega, \quad (25)$$

где $z(\omega)$ — случайный процесс, зависящий от переменной ω . Полагаем, что с вероятностью 1

$$Z(\omega) = \int_{-\infty}^{\omega} z(\omega') d\omega', \quad \omega \in \mathbb{R}. \quad (26)$$

Обозначим

$$z(\omega) = c(\omega) e^{-i\psi(\omega)}, \quad (27)$$

где $c(\omega) \geq 0$, $\psi(\omega) \in \mathbb{R}$. Тогда по формуле (23) находим

$$\xi(t) = a + \int_{-\infty}^{\infty} c(\omega) e^{it\omega - i\psi(\omega)} d\omega. \quad (28)$$

Величина $c^2(\omega) d\omega$ есть случайная мощность комплексных гармоник в элементарном интервале частот $(\omega; \omega + d\omega)$, в то время как $\psi(\omega)$ есть случайная начальная фаза гармоники частоты ω . Имеем

$$Mc^2(\omega) = f(\omega). \quad (29)$$

Для действительного случайного процесса $\xi(t)$ из (28) находим

$$\xi(t) = a + \int_0^{\infty} c_0(\omega) \cos(t\omega - \psi_0(\omega)) d\omega, \quad (30)$$

где $c_0(\omega) \geq 0$, $\psi_0(\omega) \in \mathbb{R}$.

§ 6. Анализ узкополосного случайного процесса

Стационарный в широком смысле случайный процесс называется *узкополосным*, если его спектр сосредоточен в полосе частот $[\omega_0 - \Delta; \omega_0 + \Delta]$, т. е. вне этого отрезка $dF(\omega) = 0$. Имеем

$$\xi(t) = a + \int_{\omega_0 - \Delta}^{\omega_0 + \Delta} e^{it\omega} dZ(\omega).$$

Без ограничения общности полагаем, что $\omega_0 = 0$, так как в противном случае от процесса $\xi(t)$ детерминированным преобразованием $\eta(t) = a + e^{-it\omega_0} \xi(t)$ можно перейти к процессу $\eta(t)$ с этим свойством. Таким образом, полагаем

$$\xi(t) = a + \int_{-\Delta}^{\Delta} e^{it\omega} dZ(\omega). \quad (1)$$

Определим коэффициенты Фурье

$$\gamma_n = \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} e^{-in\pi\omega/\Delta} dZ(\omega). \quad (2)$$

Из (1) находим

$$\gamma_n = \frac{1}{2\Delta} \xi^0\left(-\frac{n\pi}{\Delta}\right). \quad (3)$$

Тогда, как можно показать,

$$\xi^0(t) = \lim_{N \rightarrow \infty} \int_{-\Delta}^{\Delta} e^{it\omega} \sum_{n=-N}^N \gamma_n e^{in\pi\omega/\Delta} d\omega =$$

$$\begin{aligned}
&= \lim_{N \rightarrow \infty} \frac{1}{2\Delta} \sum_{n=-N}^N \xi^0 \left(-\frac{n\pi}{\Delta} \right) \int_{-\Delta}^{\Delta} e^{i\omega(t+n\pi/\Delta)} d\omega = \\
&= \lim_{N \rightarrow \infty} \sum_{n=-N}^N \xi^0 \left(\frac{n\pi}{\Delta} \right) \frac{(-1)^n}{t\Delta - n\pi} \sin t\Delta, \quad (4)
\end{aligned}$$

где сходимость понимается в среднеквадратическом смысле; при $t = n\pi/\Delta$ выражение $\frac{(-1)^n}{t\Delta - n\pi} \sin t\Delta$ заменяется единицей.

Отметим основные моменты доказательства важной формулы (4). Допредельное выражение в правой части этой формулы можно преобразовать следующим образом:

$$\begin{aligned}
&\sum_{n=-N}^N \xi^0 \left(\frac{n\pi}{\Delta} \right) \frac{(-1)^n}{t\Delta - n\pi} \sin t\Delta = \\
&= \frac{1}{2\Delta} \sum_{n=-N}^N \int_{-\Delta}^{\Delta} e^{-in\pi\omega/\Delta} dZ(\omega) \int_{-\Delta}^{\Delta} e^{i\omega'(t+n\pi/\Delta)} d\omega'. \quad (5)
\end{aligned}$$

В подынтегральной функции выделим сомножители, зависящие от n , и просуммируем как геометрическую прогрессию:

$$\sum_{n=-N}^N e^{-in\pi(\omega' - \omega)/\Delta} = \frac{e^{i(N+1)\pi(\omega' - \omega)/\Delta} - e^{-iN\pi(\omega' - \omega)/\Delta}}{e^{i\pi(\omega' - \omega)/\Delta} - 1}.$$

Разделив числитель и знаменатель на $2ie^{i\pi(\omega' - \omega)/(2\Delta)}$, преобразуем это выражение к виду

$$\frac{\sin \left(N + \frac{1}{2} \right) \pi (\omega' - \omega)/\Delta}{\sin \frac{1}{2} \pi (\omega' - \omega)/\Delta}.$$

Подставляя его в равенство (5), находим

$$\begin{aligned}
&\sum_{n=-N}^N \xi^0 \left(\frac{n\pi}{\Delta} \right) \frac{(-1)^n}{t\Delta - n\pi} \sin t\Delta = \\
&= \int_{-\Delta}^{\Delta} dZ(\omega) \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} e^{i\omega't} \frac{\sin \left(N + \frac{1}{2} \right) \pi (\omega' - \omega)/\Delta}{\sin \frac{1}{2} \pi (\omega' - \omega)/\Delta} d\omega'.
\end{aligned}$$

Внутренний интеграл, известный в теории рядов Фурье, при $N \rightarrow \infty$ сходится к $e^{i\omega't}$, а отсюда все выражение сходится к $\int_{-\Delta}^{\Delta} e^{i\omega't} dZ(t) = \xi^0(t)$, что и требовалось доказать (нами лишь опущены подробности предельного перехода).

Формула (4) выражает важный математический результат, известный как *теорема Шеннона — Котельникова*: с вероятностью 1 траектория стационарного случайного процесса, спектр которого

сосредоточен в полосе ширины 2Δ , полностью определяется значениями этого процесса в моменты $n\pi/\Delta$, $n \in \mathbb{Z}$. Эта теорема служит основой импульсной модуляции при передаче непрерывных сигналов.

§ 7. Линейные преобразования случайных процессов

Пусть X — множество комплексных функций $x(t)$, $t \in \mathbb{R}$, удовлетворяющее свойству линейности, а именно, $c_1 x_1(t) + \dots + c_n x_n(t)$ принадлежит X в любом случае, когда $x_1(t), \dots, x_n(t)$ принадлежат X . *Линейным функционалом* на множестве X называется соответствие $x(t) \rightarrow y$, $y \in \mathbb{R}$, сохраняющее линейность. Можно записать $y = A(x(t), t \in \mathbb{R})$, где A — символ функционала.

Пусть, например: X — множество интегрируемых функций, $y = \int_a^b x(t) dt$; X — множество дифференцируемых функций, $y = x'(a)$; X — множество всех комплексных функций, $y = x(0) - 2x(1) + x(2)$. *Оператором* на множестве X назовем соответствие

$$y(t) = A_t(x(\tau), \tau \in \mathbb{R}), \quad t \in \mathbb{R}, \quad (1)$$

при котором A_t , $t \in \mathbb{R}$, — линейные функционалы. Таким образом, линейный оператор ставит в соответствие любой функции $x(t)$ из X определенную комплексную функцию $y(t)$, причем выполняется свойство линейности этого соответствия. Например,

$$y(t) = \frac{1}{2\Delta} \int_{t-\Delta}^{t+\Delta} x(\tau) d\tau; \quad y(t) = x''(0) - ax'(t).$$

Линейным преобразованием случайного процесса $(\xi(t), t \in \mathbb{R})$ назовем преобразование траекторий этого процесса в траектории другого случайного процесса $(\eta(t), t \in \mathbb{R})$, осуществляемое некоторым линейным оператором.

С прикладной точки зрения имеют наибольшее значение линейные преобразования случайного процесса, удовлетворяющие некоторому дополнительному условию. Именно, ниже будем рассматривать лишь линейные преобразования вида

$$\eta(t) = \sum_{k=1}^n c_k \xi(t - \tau_k), \quad (2)$$

где c_k — любые комплексные числа, τ_k — любые действительные числа, и их пределы в среднем квадратическом. Таким образом, если $\eta(t)$ — линейное преобразование $\xi(t)$, то постулируется существование для любого $\varepsilon > 0$ таких n , c_1, \dots, c_n , τ_1, \dots, τ_n , что

$$M \left| \eta(t) - \sum_{k=1}^n c_k \xi(t - \tau_k) \right|^2 < \varepsilon. \quad (3)$$

Вначале рассмотрим простейшее преобразование вида (2).
Имеем

$$\begin{aligned}\eta(t) &= a \sum_{k=1}^n c_k + \sum_{k=1}^n c_k \int_{-\infty}^{\infty} e^{i(t-\tau_k)\omega} dZ(\omega) = \\ &= a \sum_{k=1}^n c_k + \int_{-\infty}^{\infty} e^{it\omega} \sum_{k=1}^n c_k e^{-i\tau_k\omega} dZ(\omega).\end{aligned}\quad (4)$$

Обозначим

$$\Phi(i\omega) = \sum_{k=1}^n c_k e^{-i\tau_k\omega} \quad (5)$$

и назовем эту функцию *частотной характеристикой* данного линейного преобразования.

Таким образом,

$$M\eta(t) = a \sum_{k=1}^n c_k, \quad (6)$$

$$\eta^0(t) = \int_{-\infty}^{\infty} e^{it\omega} \Phi(i\omega) dZ(\omega). \quad (7)$$

По формуле (5.24) из (7) получаем

$$\begin{aligned}M\eta^0(t+\tau)\overline{\eta^0(t)} &= \int_{-\infty}^{\infty} (e^{i(t+\tau)\omega} \Phi(i\omega)) \overline{(e^{it\omega} \Phi(i\omega))} dF_{\xi}(\omega) = \\ &= \int_{-\infty}^{\infty} e^{i\tau\omega} |\Phi(i\omega)|^2 dF_{\xi}(\omega),\end{aligned}\quad (8)$$

откуда заключаем, что $\eta(t)$ — стационарный в широком смысле случайный процесс. Его спектральная функция выражается через спектральную функцию $\xi(t)$ соотношением

$$F_{\eta}(\omega) = \int_{-\infty}^{\omega} |\Phi(i\omega')|^2 dF_{\xi}(\omega') \quad (9)$$

или в дифференциальном виде

$$dF_{\eta}(\omega) = |\Phi(i\omega)|^2 dF_{\xi}(\omega). \quad (10)$$

В предыдущем параграфе было отмечено, что среднеквадратической сходимости случайных величин соответствует сходимость функций переменной ω . Если $\eta(t)$ — некоторое линейное преобразование случайного процесса $\xi(t)$, возьмем последовательность

$$\eta^{(n)}(t) = \sum_{k=1}^n c_k^{(n)} \xi(t - \tau_k^{(n)}),$$

для которой

$$M|\eta(t) - \eta^{(n)}(t)|^2 \xrightarrow{n \rightarrow \infty} 0.$$

Тогда, очевидно,

$$M |\eta^{(n)}(t) - \eta^{(m)}(t)|^2 \xrightarrow{n, m \rightarrow \infty} 0. \quad (11)$$

Левая часть (11) равна

$$\int_{-\infty}^{\infty} \left| \sum_{k=1}^n c_k^{(n)} e^{-i\tau_k^{(n)}\omega} - \sum_{k=1}^m c_k^{(m)} e^{-i\tau_k^{(m)}\omega} \right|^2 dF_{\xi}(\omega).$$

Из того что это выражение бесконечно мало при $n, m \rightarrow \infty$, следует (теорема о полноте пространства функций, интегрируемых с квадратом с весом) существование предельной функции, которую обозначим через $\Phi(i\omega)$ и для которой

$$\int_{-\infty}^{\infty} \left| \Phi(i\omega) - \sum_{k=1}^n c_k^{(n)} e^{-i\tau_k^{(n)}\omega} \right|^2 dF_{\xi}(\omega) \xrightarrow{n \rightarrow \infty} 0. \quad (12)$$

Отсюда, ввиду изоморфизма между пространством случайных величин и пространством функций, имеем

$$M\eta(t) = a \lim_{n \rightarrow \infty} \sum_{k=1}^n c_k^{(n)} = a\Phi(0), \quad (13)$$

$$\eta^0(t) = \int_{-\infty}^{\infty} e^{it\omega} \Phi(i\omega) dZ(\omega), \quad (14)$$

что совпадает с формулой (7), относящейся к простейшим линейным преобразованиям. Функцию $\Phi(i\omega)$, как и в этом частном случае, назовем *частотной характеристикой* линейного преобразования процесса $\xi(t)$ в процесс $\eta(t)$. Формулы (9), (10) также справедливы в общем случае.

Рассмотрим важнейшие частные случаи.

1. Разностный оператор

$$\eta_h(t) = \frac{1}{h} (\xi(t+h) - \xi(t))$$

характеризуется частотной характеристикой

$$\Phi_h(t) = \frac{1}{h} (e^{ih\omega} - 1).$$

Очевидно, $\Phi_h(i\omega) \rightarrow i\omega$ при $h \rightarrow 0$. Предел $\eta_h(t)$ в среднеквадратическом, т. е. среднеквадратическая производная $\xi'(t)$ случайного процесса $\xi(t)$, существует при условии

$$\int_{-\infty}^{\infty} \omega^2 dF_{\xi}(\omega) < \infty. \quad (15)$$

Имеем

$$dF_{\xi'}(\omega) = \omega^2 dF_{\xi}(\omega). \quad (16)$$

Подобным же образом может быть определен оператор вида $a_0 \xi(t) + a_1 \xi'(t) + \dots + a_m \xi^{(m)}(t)$. Частотная характеристика этого оператора равна $a_0 + a_1 \omega^2 + \dots + a_m \omega^{2m}$.

2. Если $c(t)$ — комплексная функция, для которой $\int_{-\infty}^{\infty} |c(t)|^2 dt < \infty$, то оператор скользящего среднего

$$\eta(t) = \int_{-\infty}^{\infty} c(t - \tau) \xi(\tau) d\tau, \quad (17)$$

где правая часть принимается как стохастический интеграл, определяет случайный процесс, стационарный в широком смысле, с характеристиками

$$\begin{aligned} M\eta(t) &= \int_{-\infty}^{\infty} c(\tau) d\tau M\xi(t), \\ dF_{\eta}(\omega) &= |\Phi(i\omega)|^2 dF_{\xi}(\omega), \end{aligned} \quad (18)$$

где

$$\Phi(i\omega) = \int_{-\infty}^{\infty} e^{-i\tau\omega} c(\tau) d\tau. \quad (19)$$

Таким образом, частотная характеристика оператора скользящего среднего есть преобразование Фурье функции $c(\tau)$. Так, например, оператору

$$\eta(t) = b \int_{-\infty}^t e^{-\alpha(t-\tau)} \xi(\tau) d\tau \quad (20)$$

соответствуют характеристики

$$\begin{aligned} M\eta(t) &= \frac{b}{\alpha} M\xi(t), \\ \Phi(i\omega) &= \frac{b}{\alpha + i\omega}. \end{aligned} \quad (21)$$

Поскольку в данном случае

$$|\Phi(i\omega)|^2 = b^2/(\alpha^2 + \omega^2),$$

то процесс вида (20) всегда существует и стационарен в широком смысле.

Оператор вида (17) при условии $\int_{-\infty}^{\infty} |c(\tau)| d\tau < \infty$ может быть применен к комплексным гармоникам $x(t) = Ae^{i\omega t}$, где A — любая постоянная, ω — действительное число. Имеем

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} c(t - \tau) Ae^{i\omega\tau} d\tau = \\ &= Ae^{i\omega t} \int_{-\infty}^{\infty} c(t - \tau) e^{i\omega(t-\tau)} d\tau = Ae^{i\omega t} \Phi(i\omega), \end{aligned}$$

или

$$y(t) = \Phi(i\omega) x(t).$$

Следовательно, $\Phi(i\omega)$ можно интерпретировать как отношение $y(t)$ к $x(t)$ при указанном виде $x(t)$.

Пример. Можно ли преобразованием $y(t) = \int_{-\infty}^{\infty} c(t-\tau) x(\tau) d\tau$ сделать так называемый идеальный прямоугольный фильтр частот, а именно, удовлетворить условиям $y(t) = 0$ при $x(t) = e^{i\omega t}$, $|\omega| > \omega_0$; $y(t) = x(t)$ при $x(t) = e^{i\omega t}$, $|\omega| < \omega_0$?

Для гипотетического оператора имеем

$$\Phi(i\omega) = \begin{cases} 1 & \text{при } |\omega| < \omega_0, \\ 0 & \text{при } |\omega| > \omega_0. \end{cases}$$

Применив к (19) обратное преобразование Фурье, получим

$$c(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\tau\omega} \Phi(i\omega) d\omega = \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} e^{i\tau\omega} d\omega = \sin \tau\omega_0 / (\pi\tau). \quad (23)$$

Видим, что решение задачи существует, однако в данном случае

$$\int_{-\infty}^{\infty} |c(\tau)| d\tau = \infty.$$

Задачи

1. Случайные процессы $\xi(t)$, $\eta(t)$ называются *некоррелированными*, если их взаимная корреляционная функция равна нулю. Вывести формулы для корреляционной функции, спектральной функции и спектральной плотности линейной комбинации попарно некоррелированных комплексных случайных процессов. Сделать то же самое для действительных процессов.

2. Пусть $\xi(t)$, $\eta(t)$ — стационарные, стационарно связанные комплексные случайные процессы. При каком условии существует такое ортогональное преобразование A , что для всех t случайные процессы $X(t)$, $Y(t)$, задаваемые равенством

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = A \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix},$$

некоррелированы?

3. Пусть $K(t, \tau)$ — корреляционная функция случайного процесса, определенная при $a \leq t \leq b$, $a \leq \tau \leq b$. Доказать, что из непрерывности $K(t, \tau)$ в точках (t, t) , $a \leq t \leq b$, следует ее непрерывность во всех точках (t, τ) , $a \leq t \leq b$, $a \leq \tau \leq b$.

4. Пусть $\xi(t)$ — комплексный стационарный в широком смысле процесс с корреляционной функцией $K(t)$. Найти корреляционную функцию процесса

$$\eta(t) = \int_0^1 x^m \xi(t-x) dx.$$

5. Пусть действительный процесс $\xi(t)$ имеет корреляционную функцию вида $\sigma^2 e^{-\alpha\tau}$. Вывести выражение для дисперсии разности

$$\xi(h) - (1-\rho) \sum_{n=0}^{\infty} \rho^n \xi(-nh),$$

интерпретируемой как дисперсия ошибки экспоненциального экстраполятора значения процесса $\xi(t)$ в момент h по наблюдениям в моменты $0, -h, -2h, \dots$. Указать значение ρ , при котором дисперсия ошибки минимальна.

6. Вывести формулу для корреляционной функции случайного процесса

$$\xi_N(t) = A \sum_{k=0}^{N-1} (u_k \cos k\Delta t + v_k \sin k\Delta t),$$

где u_k, v_k — некоррелированные случайные величины с нулевым средним; $Du_k = Dv_k = N - k$. Можно ли подобрать параметры $A = A(N)$, $\Delta = \Delta(N)$ таким образом, чтобы при $N \rightarrow \infty$ корреляционная функция случайного процесса

$$\eta_N(t) = \int_0^t \xi_N(\tau) d\tau \text{ сходилась к функции } K(t, \tau) = \min\{t, \tau\} \quad (t \geq 0, \tau \geq 0)?$$

7. Пусть $X(t)$ — процесс Пуассона с параметром λ . а) Найти его корреляционную функцию. б) Найти корреляционную функцию случайного процесса $Y(t) = X(t+h) - X(t)$. Является ли этот процесс стационарным в широком смысле; непрерывным в среднем квадратическом?

8. На отрезок $(a; b)$ независимо бросается n точек $\omega_1, \dots, \omega_n$ по равномерному закону. Найти корреляционную функцию и спектральную плотность комплексного случайного процесса

$$\xi(t) = \sum_{k=1}^n \gamma_k e^{i\omega_k t},$$

где $M\gamma_k = 0$, $M\gamma_k \bar{\gamma}_j = 0$, $k \neq j$, $M|\gamma_k|^2 = 1$.

9. Обосновать формулу (6.4).

Глава 8

ДРУГИЕ ВАЖНЫЕ КЛАССЫ СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 1. Винеровский процесс

Винеровским процессом (процессом броуновского движения) с коэффициентом сноса a и коэффициентом диффузии σ^2 называется одномерный случайный процесс $(\xi(t), t \geq 0)$ со следующими свойствами.

1. $\xi(0) = 0$ с вероятностью 1.

2. При любом n и любых $0 < t_1 < t_2 < \dots < t_n$ случайные величины $\xi(t_1)$, $\xi(t_2) - \xi(t_1)$, ..., $\xi(t_n) - \xi(t_{n-1})$ независимы.

3. При любых $0 \leq t_0 < t$ случайная величина $\xi(t) - \xi(t_0)$ имеет нормальное распределение с математическим ожиданием $a(t - t_0)$ и дисперсией $\sigma^2(t - t_0)$.

4. Траектория $\xi(t)$ непрерывна с вероятностью 1.

Возникает вопрос: существует ли процесс с заданными свойствами? Иными словами: существует ли вероятностное пространство (Ω, \mathcal{F}, P) и такая случайная функция $\xi(t) = \xi(t, \omega)$, $t \geq 0$, $\omega \in \Omega$, что для этой функции выполняются свойства 1—4? Для доказательства существования такой случайной функции можно применить прием конструктивного определения винеровского процесса, т. е. выражения его через последовательность независимых случайных величин. Это делается следующим образом. Непосредственным

выписыванием условных плотностей устанавливается следующий факт. Пусть известно, что $\xi\left(\frac{k+1}{2^n}\right) + \xi\left(\frac{k-1}{2^n}\right) = \Delta$. Тогда, независимо от значений процесса в точках $(k \pm 1)/2^n$, случайная величина $\xi\left(\frac{k}{2^n}\right)$ распределена нормально с математическим ожиданием $\Delta/2$ и дисперсией $\sigma^2/2^{n+1}$. Вначале возьмем последовательность (w_{nj}) независимых стандартных нормальных случайных величин и положим

$$\xi(n) = an + w_{10} + \dots + w_{n0}, \quad n \geq 0.$$

Затем положим

$$\xi\left(n + \frac{1}{2}\right) = \frac{1}{2} (\xi(n) + \xi(n+1)) + \frac{\sigma}{2} w_{n+1,1}.$$

Подобным же образом по известным $\xi\left(\frac{k}{2^m-1}\right)$, $k \geq 0$, и последовательности w_{nm} строятся значения $\xi(t)$ в промежуточных точках вида $t = k/2^m$. Таким образом мы построим $\xi(t)$ во всех двоично-рациональных точках. В иррациональных точках значение $\xi(t)$ определяется предельным переходом, подробности которого (доказательство непрерывности $\xi(t)$ с вероятностью 1) мы опускаем.

Обозначим $u(x, t)$ плотность вероятности случайной величины $\xi(t)$. Тогда, очевидно,

$$u(x, t) = \frac{1}{\sqrt{2\pi t} \sigma} e^{-(x-at)^2/(2\sigma^2 t)}.$$

Непосредственной подстановкой можно показать, что $u(x, t)$ есть решение уравнения теплопроводности

$$\frac{\partial u}{\partial t} = -a \frac{\partial u}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} \quad (1)$$

при начальном условии

$$u(x, 0) = \delta(x). \quad (2)$$

Стандартным винеровским процессом называется винеровский процесс $w(t)$ с параметрами $a = 0$, $\sigma^2 = 1$.

Легко видеть, что если $w(t)$ — стандартный винеровский процесс, то случайный процесс $\xi(t) = at + \sigma w(t)$ будет винеровским процессом с коэффициентом сноса a и коэффициентом диффузии σ^2 .

Винеровский процесс является математической идеализацией движения «безынерционной» частицы. Любая реальная частица за конечное время проходит конечный путь. Для винеровского процесса (для определенности возьмем стандартный) путь, проходимый за положительное время t ,

$$s(t) \geq \left| w\left(\frac{t}{n}\right) \right| + \left| w\left(\frac{2t}{n}\right) - w\left(\frac{t}{n}\right) \right| + \dots + \\ + \left| w(t) - w\left(\frac{n-t}{n} t\right) \right|.$$

Слагаемые этой суммы — модули независимых случайных величин с математическим ожиданием 0 и дисперсией t/n . Можно записать

$$\left| w\left(\frac{t}{n}\right) \right| = \sqrt{\frac{t}{n}} \alpha_1, \quad \left| w\left(\frac{2t}{n}\right) - w\left(\frac{t}{n}\right) \right| = \sqrt{\frac{t}{n}} \alpha_2, \dots,$$

где $\alpha_1, \alpha_2, \dots$ — модули независимых стандартных нормальных величин. Отсюда

$$s(t) \geq \sqrt{\frac{t}{n}} (\alpha_1 + \dots + \alpha_n) = \sqrt{nt} \bar{\alpha}_n,$$

где $\bar{\alpha}_n = \frac{1}{n} (\alpha_1 + \dots + \alpha_n)$. Поскольку по закону больших чисел $\bar{\alpha}_n \xrightarrow{P} M\alpha_k > 0$, то при любом $t > 0$ величина $s(t)$ сходится по вероятности к $+\infty$ при $n \rightarrow \infty$. Поскольку же эта величина от n не зависит, то $s(t) \rightarrow +\infty$ по вероятности.

§ 2. Сходимость к винеровскому процессу

Пусть $(\xi_n(t), n \geq 1)$ — последовательность случайных процессов. По определению, эта последовательность *слабо сходится* к случайному процессу $\eta(t)$, если все конечномерные распределения $\xi_n(t)$ слабо сходятся при $n \rightarrow \infty$ к соответствующим распределениям $\eta(t)$. Формулой это можно выразить так: для любого $m \geq 1$ и любых t_1, \dots, t_m

$$P(\xi_n(t_k) < x_k, 1 \leq k \leq m) \xrightarrow{n \rightarrow \infty} P(\eta(t_k) < x_k, 1 \leq k \leq m)$$

для любой точки (x_1, \dots, x_m) , в которой правая часть этого соотношения непрерывна. (Если $\xi_n(t), \eta(t)$ — многомерные процессы, то x_k — векторы, и тогда, как обычно, неравенство для векторов понимается как система неравенств для их соответствующих компонент.) Если $\xi_n(t)$ слабо сходится к $\eta(t)$, говорят, что $\eta(t)$ — предельный процесс последовательности процессов $\xi_n(t)$, и записывают

$$\xi_n(t) \xrightarrow{w} \eta(t).$$

(Строго говоря, нужно говорить о слабой сходимости не самих процессов, а их конечномерных распределений, но такое уточнение для сокращения терминологии обычно опускается.)

Винеровский процесс служит предельным процессом во многих постановках задач теории вероятностей. Приведем лишь простейшие из них.

1. Пусть $(\gamma_k, k \geq 1)$ — последовательность независимых одинаково распределенных случайных величин; $M\gamma_k = a$, $D\gamma_k = \sigma^2$. Рассмотрим последовательность случайных процессов

$$\xi_n(t) = \frac{1}{\sqrt{n}} \sum_{k \leq nt} \gamma_k^0.$$

Если $0 = t_0 < t_1 < \dots < t_m$, то можно записать

$$\xi_n(t_1) = \zeta_1, \quad \xi_n(t_2) = \zeta_1 + \zeta_2, \quad \dots, \quad \xi_n(t_m) = \zeta_1 + \dots + \zeta_m,$$

где

$$\zeta_k = \frac{1}{\sqrt{n}} \sum_{nt_{k-1} \leq k < nt_k} \gamma_k^0.$$

Отсюда следует, что случайная величина $(\zeta_1, \dots, \zeta_m)$ асимптотически нормальна с нулевым математическим ожиданием и корреляционной матрицей

$$\begin{pmatrix} \sigma^2 t_1 & & & 0 \\ & \sigma^2(t_2 - t_1) & & \\ & & \ddots & \\ 0 & & & \sigma^2(t_m - t_{m-1}) \end{pmatrix}.$$

Далее можно записать

$$\xi_n = \begin{pmatrix} \xi_n(t_1) \\ \vdots \\ \xi_n(t_m) \end{pmatrix} = B \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} = B\zeta,$$

где

$$B = \begin{pmatrix} 1 & & 0 \\ 1 & 1 & \\ \cdot & \cdot & \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 1 & 1 \end{pmatrix}.$$

Отсюда характеристическая функция $\varphi_n(t)$ случайной величины $(\xi_n(t_k))$ имеет вид

$$\mathbf{M} \exp \{it' \xi_n\} = \varphi_n(Bt),$$

где $\varphi_n(t)$ — характеристическая функция случайной величины ζ . Поскольку по доказанному $\varphi_n(Bt)$ сходится к предельной характеристической функции, то сходится и $\mathbf{M} \exp \{it' \xi_n\}$. По обратной теореме непрерывности для характеристических функций величина $(\xi_n(t_1), \dots, \xi_n(t_m))$ имеет предельное распределение, совпадающее с m -мерным распределением винеровского процесса $\eta(t)$ с нулевым коэффициентом сноса и коэффициентом диффузии σ^2 . Итак, $\xi_n(t) \xrightarrow[n \rightarrow \infty]{w} \eta(t)$.

2. Пусть имеется простейший поток с параметром λ ; t_k — момент k -го события этого потока; ξ_k — случайная величина, связанная с k -м событием (например, амплитуда порождаемого им импульса). Обозначим через $s(t)$ сумму случайных величин, связанных с событиями, происшедшими в интервале времени $(0; t)$:

$$s(t) = \sum_{t_k < t} \xi_k.$$

Допустим теперь, что ξ_k — независимые случайные величины с общей характеристической функцией $\psi(t)$, математическим ожида-

нием a и конечной дисперсией σ^2 , причем последовательности (ξ_k) и (t_k) независимы. По построению $s(t)$ видно, что при $0 < T_1 < \dots < T_m$ случайные величины $s(T_1)$, $s(T_2) - s(T_1)$, ..., $s(T_m) - s(T_{m-1})$ независимы; при $T > 0$, $\tau \geq 0$ величина $s(T + \tau) - s(\tau)$ распределена так же, как и $s(T)$. Характеристическая функция $\varphi_T(t)$ случайной величины $s(T)$ вычисляется следующим образом:

$$\begin{aligned}\varphi_T(t) &= M \exp \{its(T)\} = \sum_{n=0}^{\infty} e^{-\lambda T} \frac{(\lambda T)^n}{n!} M e^{it(\xi_1 + \dots + \xi_n)} = \\ &= e^{-\lambda T} \sum_{n=0}^{\infty} \frac{(\lambda T \psi(t))^n}{n!} = \exp \{\lambda T (\psi(t) - 1)\}.\end{aligned}$$

Дифференцированием находим

$$\varphi'_T(0) = i\lambda Ta, \quad \varphi''_T(0) = -\lambda T(a^2 + \sigma^2) - (\lambda Ta)^2.$$

Отсюда

$$Ms(T) = -i\varphi'_T(0) = \lambda Ta, \quad (1)$$

$$Ds(T) = -\varphi''_T(0) - (Ms(T))^2 = \lambda T(a^2 + \sigma^2). \quad (2)$$

Рассмотрим центрированный случайный процесс

$$s^0(T) = s(T) - Ms(T) = s(T) - \lambda Ta.$$

Имеем

$$\varphi^0_T(t) = M \exp \{its^0(T)\} = \exp \{\lambda T (\psi(t) - 1 - ita)\}.$$

Наконец, рассмотрим случайный процесс

$$\xi(T) = \frac{1}{\sqrt{\lambda}} s^0(T).$$

Характеристическая функция случайной величины $\xi(T)$ равна

$$\begin{aligned}\varphi^0_T(t/\sqrt{\lambda}) &= \exp \{\lambda T (\psi(t/\sqrt{\lambda}) - 1 - iat/\sqrt{\lambda})\} \xrightarrow{\lambda \rightarrow \infty} \\ &\xrightarrow{\lambda \rightarrow \infty} \exp \left\{ -\frac{t^2}{2} T(a^2 + \sigma^2) \right\}.\end{aligned} \quad (3)$$

При выводе этого соотношения использованы только формула Тейлора и равенство (2). Полученное соотношение интерпретируется так: случайный процесс $\xi(T)$ при $\lambda \rightarrow \infty$ слабо сходится к винеровскому процессу с нулевым коэффициентом сноса и коэффициентом диффузии $a^2 + \sigma^2$.

Заметим, что параметр λ зависит от выбранного масштаба времени, поэтому всегда можно подобрать масштаб так, чтобы $\lambda \rightarrow \infty$, но это означает, что «винеровское» поведение наблюдается на достаточно больших интервалах реального времени таких, в течение которых происходит в среднем достаточно много событий простейшего потока.

Выше было замечено, что винеровский процесс отражает закон движения «безынерционной», т. е. физически не существующей

частицы. Польза, получаемая от него, заключается в том, что он служит приближенной моделью многих реальных явлений. (Выше были приведены два примера такого рода.) Но одного этого обстоятельства еще мало. Нужно, чтобы: 1) интересующая нас характеристика (скажем, распределение момента выхода траектории процесса из интервала, математическое ожидание времени пребывания в интервале и т. п.) легко находилась для винеровского процесса и 2) значение характеристики для реального (допредельного) процесса хорошо приближалось характеристикой винеровского (предельного) процесса. Оказывается, оба эти вопроса решаются положительно во многих случаях. Остановимся на первом вопросе: какие задачи можно решать для винеровского процесса?

Простейшая из таких задач нами уже отмечена: имеем точную формулу и уравнение (1.1) с условием (1.2) и точную формулу для плотности $u(x, t)$ значения процесса $\xi(t)$. Рассмотрим еще некоторые задачи.

1. Пусть $x = x_1(t)$, $x = x_2(t)$ — две гладкие кривые в плоскости (x, t) , причем $x_1(0) < 0 < x_2(0)$, $x_1(t) < x_2(t)$, $t \geq 0$. Обозначим через $v(\tau, x)$ вероятность события, состоящего в том, что траектория случайного процесса $x + \xi(t - \tau)$, $t \geq \tau$, пересечет кривую $x = x_2(t)$ раньше, чем кривую $x = x_1(t)$. В частности, $v(0, 0)$ есть вероятность того, что винеровский процесс $\xi(t)$, $t \geq 0$, пересечет раньше кривую $x = x_2(t)$, чем кривую $x = x_1(t)$. Оказывается, что $v(\tau, x)$ есть решение обратного уравнения теплопроводности

$$\frac{\partial v}{\partial \tau} + \frac{\partial v}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2} = 0 \quad (4)$$

в области $x_1(\tau) < x < x_2(\tau)$, $\tau \neq 0$, при краевых условиях

$$v(x_1(\tau), \tau) = 0, \quad v(x_2(\tau), \tau) = 1. \quad (5)$$

Пусть, например, при $\sigma^2 = 1$ в случае прямолинейных границ $x_1(\tau) = -\Delta$, $x_2(\tau) = \Delta$ имеем уравнение $\frac{\partial v}{\partial \tau} + \frac{\partial^2 v}{\partial x^2} = 0$ при условиях $v(-\Delta, \tau) = 0$, $v(\Delta, \tau) = 1$.

Легко видеть, что решение имеет вид

$$v(x, \tau) = (x + \Delta)/(2\Delta), \quad -\Delta \leq x \leq \Delta.$$

2. Пусть s — момент первого достижения винеровским процессом границы. Если это нижняя граница, то получается доход $f_1(s)$ если верхняя — доход $f_2(s)$; функции $f_1(s)$ и $f_2(s)$ считаем заданными ограниченными функциями с ограниченными производными. Допустим, что до момента τ выход на какую-либо границу не произошел, причем $\xi(\tau) = x$. Тогда $v(\tau, x)$ пусть обозначает средний будущий доход.

Можно показать, что функция $v(\tau, x)$ удовлетворяет уравнению (4) в области $x_1(\tau) < x < x_2(\tau)$, $\tau \geq 0$, при граничных условиях

$$v(x_i(\tau), \tau) = f_i(\tau), \quad \tau \geq 0, \quad i = 1, 2. \quad (5')$$

Пусть, например, $a = 0$, $\sigma^2 = 1$, $x_1(\tau) = -\Delta$, $x_2(\tau) = \Delta$, $f_1(\tau) = e^{-\tau}$, $f_2(\tau) = 0$.

Решив уравнение (4) при условиях (5), получим

$$v(x, \tau) = e^{-\tau} (e^{3\Delta-x} - e^{\Delta+x}) / (e^{\Delta} - 1).$$

Заметим, что результаты приведенных примеров можно было частично предвидеть и не решая уравнение (4). Именно, в первом примере сразу ясно, что $v(x, \tau)$ не зависит от τ ; во втором — ситуация в момент τ отличается от ситуации в момент 0 только тем, что произошло «снижение ставки», а следовательно, $v(x, \tau) = e^{-\tau} v(x, 0)$.

Перейдем ко второму из поставленных выше вопросов. Прежде всего уточним постановку задачи.

Пусть $(\xi_n(t))$ — последовательность случайных процессов, $\xi_n(t)$ — случайный процесс. Предположим, что $\xi_n(t) \xrightarrow[n \rightarrow \infty]{w} \xi(t)$. Рассмотрим функционал f от траектории процесса и обозначим

$$f_n = f(\xi_n(\dots)), \quad f = f(\xi(\dots)).$$

По определению, для последовательности случайных процессов $(\xi_n(\dots))$ выполняется *инвариантность предельного перехода* по отношению к функционалу f , если распределение величины f_n при $n \rightarrow \infty$ слабо сходится к распределению величины f . Теоремы, обосновывающие инвариантность предельного перехода по отношению к тем или иным классам функционалов f , называются *принципами инвариантности*. Каждый принцип инвариантности формулируется так: *при некоторых дополнительных условиях* (указывается, каких именно) *из слабой сходимости случайных процессов следует слабая сходимость распределения функционала f_n к распределению f при $n \rightarrow \infty$* .

Кратко охарактеризуем один из принципов инвариантности, а именно, инвариантности Ю. В. Прохорова.

Пусть f — функционал, заданный на траекториях процесса на отрезке $[a; b]$, обладающий следующим свойством. Для любой непрерывной функции $x(t)$ существует такая функция $\varepsilon(\delta) \rightarrow 0$, $\delta \downarrow 0$, что для любой другой непрерывной функции $y(t)$, для которой $|y(t) - x(t)| \leq \delta$, $a \leq t \leq b$, выполняется условие $|f(y(\dots)) - f(x(\dots))| \leq \varepsilon(\delta)$. Этому условию удовлетворяют, в частности, функционалы $f(x(\dots))$ вида

$$\int_a^b f^m(t) dt, \quad \max_{a \leq t \leq b} f(t), \quad \min_{a \leq t \leq b} f(t), \quad \max_{a \leq t \leq b} |f(t)|.$$

Обозначим, далее, через $\omega_n(\Delta)$ максимальное колебание случайного процесса $\xi_n(t)$ в интервалах длины Δ :

$$\omega_n(\Delta) = \sup_{\substack{[t_1; t_2] \subset [a; b] \\ t_2 - t_1 \leq \Delta}} |\xi_n(t_2) - \xi_n(t_1)|.$$

Тогда для любого $\varepsilon > 0$

$$\sup_n P(\omega_n(\Delta) > \varepsilon) \xrightarrow[\Delta \rightarrow 0]{} 0. \quad (6)$$

Данные условия достаточны для того, чтобы из слабой сходимости $\xi_n(t)$ и $\xi(t)$ следовала слабая сходимость распределения $f_n = f(\xi_n(\dots))$ к распределению $f = f(\xi(\dots))$ при $n \rightarrow \infty$.

Идея (но не техника) доказательства теоремы Ю. В. Прохорова достаточно проста. Рассмотрим многомерные случайные величины

$$\bar{\xi}_n = (\xi_n(a), \xi_n(a + \Delta), \dots, \xi_n(a + N\Delta)), \\ \bar{\xi} = (\xi(a), \xi(a + \Delta), \dots, \xi(a + N\Delta)),$$

где $a + N\Delta \leq b < a + (N + 1)\Delta$. По условию, распределение $\bar{\xi}_n$ при $n \rightarrow \infty$ слабо сходится к распределению $\bar{\xi}$. Используем *метод единого вероятностного пространства*, определив $\bar{\xi}_n$ и $\bar{\xi}$ на одном и том же вероятностном пространстве: $\bar{\xi}_n = \bar{\xi}_n(\omega)$, $\bar{\xi} = \bar{\xi}(\omega)$, $\omega \in \Omega$, причем так, чтобы отклонение вектора $\bar{\xi}_n$ от вектора $\bar{\xi}$ стремилось к нулю по вероятности при $n \rightarrow \infty$. (Например, при $N = 0$ можно положить: ω — равномерная в интервале $(0; 1)$ случайная величина; $\xi_n(a) = F_n^{-1}(\omega)$, $\xi(a) = F^{-1}(\omega)$, где F_n , F — функции распределения случайных величин $\xi_n(a)$, $\xi(a)$. При $N > 0$ тот же прием применяется к последовательному построению пар $(\xi_n(a), \xi(a))$, $(\xi_n(a + \Delta), \xi(a + \Delta))$ и т. д. с использованием условных распределений; см. § 3.12.)

Итак, $\xi_n(t)$ и $\xi(t)$ будут иметь близкие значения в «узловых» точках $t = a + k\Delta$. Что касается промежуточных точек, то условие (6) не позволит функции $\xi_n(t)$ сколь-либо сильно отклониться от значения в ближайшей узловой точке, а условие непрерывности траектории $\xi(t)$ обеспечит то же самое для функции $\xi(t)$. Тогда получим, что максимальное на отрезке $[a; b]$ отклонение $\xi_n(t)$ от $\xi(t)$ — бесконечно малая при $n \rightarrow \infty$ случайная величина; из этого следует бесконечная малость отклонения f_n от f . Наконец, воспользуемся тем, что из сходимости $f_n \xrightarrow{p} f$ следует слабая сходимость их распределений (§ 5.1).

Разумеется, схема доказательства изложена весьма конспективно.

Заметим, что условие (6) трудно проверяемое на практике, можно заменить более жестким, но более практичным, а именно: существуют такие $c > 0$, $\delta > 0$ и $\rho > 0$, что для всех n и всех t_1, t_2 из отрезка $[a; b]$

$$M |\xi_n(t_2) - \xi_n(t_1)|^\delta \leq c |t_2 - t_1|^{1+\rho}.$$

Известны также принципы инвариантности Донскера, И. И. Гихмана, А. В. Скорохода [8].

§ 3. Процессы с независимыми приращениями¹

Определение. *Процессом с независимыми приращениями называется такой случайный процесс $\xi(t)$, $t \geq 0$, что для любого $n \geq 2$ и любых $0 < t_1 < \dots < t_n$ случайные величины $\xi(t_1)$, $\xi(t_2) - \xi(t_1)$, \dots , $\xi(t_n) - \xi(t_{n-1})$ независимы в совокупности.*

¹ Материал данного параграфа заимствован из книги [24].

Процесс с независимыми приращениями называется *процессом с независимыми однородными приращениями*, если распределение случайной величины $\xi(t) - \xi(s)$ при $0 \leq s < t < \infty$ зависит только от $t - s$.

Мы уже рассмотрели два важнейших примера процессов с независимыми приращениями — винеровский процесс и процесс Пуассона.

Назовем *обобщенным пуассоновским процессом* случайный процесс

$$\xi(t) = \sum_{k=1}^{v(t)} \eta_k,$$

где $v(t)$ — процесс Пуассона с параметром λ ; η_k — независимые между собой и не зависящие от процесса $v(t)$ случайные величины с функцией распределения $H(x)$. (Процесс Пуассона — частный случай такого процесса; он получается, если предположить, что все η_k равны 1.) Поскольку $v(0) = 0$, то $\xi(0) = 0$. Найдем функцию распределения случайной величины $\xi(t)$. Для этого заметим, что $v(t) = n$ с вероятностью $\frac{(\lambda t)^n}{n!} e^{-\lambda t}$, $n \geq 0$. Следовательно, если $F(t, x) = P(\xi(t) < x)$, то

$$F(t, x) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} P(\eta_1 + \dots + \eta_n < x). \quad (1)$$

В случае $n = 0$ под $P(\eta_1 + \dots + \eta_n < x)$ понимается

$$P(0 < x) = E(x) = \begin{cases} 1 & \text{при } x > 0, \\ 0 & \text{при } x \leq 0. \end{cases}$$

Обозначим через $\psi(u)$ характеристическую функцию случайных величин η_k , $k \geq 1$, через $\varphi(t, u)$ — характеристическую функцию $\xi(t)$, т. е. $\varphi(t, u) = M e^{i u \xi(t)}$. Тогда из (1) находим

$$\varphi(t, u) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \psi^n(u) = e^{\lambda t(\psi(u) - 1)}$$

(эта формула уже встречалась нам в § 2).

Логарифм характеристической функции значения процесса в момент t

$$\begin{aligned} \ln \varphi(t, u) &= \lambda t (\psi(u) - 1) = \lambda t \left(\int e^{i u x} dH(x) - 1 \right) = \\ &= t \int (e^{i u x} - 1) \mu(dx), \\ \mu(dx) &= \lambda (H(x + dx) - H(x)). \end{aligned} \quad (2)$$

Выражение $\mu(dx) dt$ есть вероятность того, что в интервале длины dt процесс $\xi(t)$ испытает скачок, величина которого лежит в преде-

лах от x до $x + dx$. Очевидно,

$$\int_{-\infty}^{\infty} \mu(dx) = \lambda,$$

т. е. мера μ конечна. Возникает вопрос: нельзя ли определить аналогичный процесс с бесконечной мерой μ ?

В случае обобщенного пуассоновского процесса обозначим

$$\mu(dx) = \mu_0(dx) + \mu_1(dx),$$

где $\mu_0(dx)$ — мера, сосредоточенная на отрезке $[-1; 1]$; $\mu_1(dx)$ — мера на множестве точек $\{|x| > 1\}$. Для этого следует положить в случае, когда dx символически обозначает интервал $(x; x + dx)$, $\mu_0(dx) = \mu(x) dx$ при $|x| \leq 1$; $\mu_0(dx) = 0$ при $|x| > 1$; $\mu_1(dx) = 0$ при $|x| \leq 1$; $\mu_1(dx) = \mu(x) dx$ при $|x| > 1$. Как следует из формулы (2), в этом случае

$$\ln \varphi(t, u) = t \int (e^{iux} - 1) \mu_0(dx) + t \int (e^{iux} - 1) \mu_1(dx),$$

т. е. случайная величина $\xi(t)$ распадается в сумму независимых случайных величин $\xi_0(t)$ и $\xi_1(t)$, где $\xi_0(t)$ — сумма скачков процесса $\xi(t)$ в интервале $(0; t)$, величина каждого из которых по абсолютной величине не больше 1; $\xi_1(t)$ — сумма всех остальных скачков. Таким образом,

$$\ln \varphi(t, u) = \ln \varphi_0(t, u) + \ln \varphi_1(t, u),$$

где

$$\ln \varphi_j(t, u) = \ln \mathbf{M} e^{iu\xi_j(t)} = t \int (e^{iux} - 1) \mu_j(dx), \quad j = 0, 1. \quad (3)$$

Рассмотрим процесс $\xi_0(t)$. По известной формуле

$$\mathbf{M}\xi_0(t) = -i \frac{\partial \varphi_0(t, u)}{\partial u} \Big|_{u=0} = -i \frac{\partial \ln \varphi_0(t, u)}{\partial u} \Big|_{u=0}$$

(второе равенство выполняется на основании того, что $\frac{\partial \ln \varphi_0}{\partial u} = \frac{1}{\varphi_0} \frac{\partial \varphi_0}{\partial u}$ и $\varphi_0 = 1$ при $u = 0$). Из (3) находим

$$\mathbf{M}\xi_0(t) = -it \frac{d}{du} \int (e^{iux} - 1) \mu_0(dx) \Big|_{u=0} = t \int x \mu_0(dx).$$

Рассмотрим случайный процесс

$$\eta_0(t) = \xi_0(t) - t\mathbf{M}\xi_0(1).$$

Тогда

$$\begin{aligned} \ln \mathbf{M} e^{iu\eta_0(t)} &= \ln \mathbf{M} e^{iu\xi_0(t) - iut\mathbf{M}\xi_0(1)} = \ln \mathbf{M} e^{iu\xi_0(t)} - iut\mathbf{M}\xi_0(1) = \\ &= t \int (e^{iux} - 1) \mu_0(dx) - iut\mathbf{M}\xi_0(1) = t \int (e^{iux} - 1 - iux) \mu_0(dx). \end{aligned} \quad (4)$$

Оказывается, что для сходимости интеграла, в формуле (4) μ_0 уже не обязательно должна быть конечной мерой.

Рассмотрим теперь три независимых процесса с независимыми приращениями $\eta_0(t)$, $\xi_1(t)$ и винеровский процесс $\zeta(t)$ с коэффициентом сноса a и коэффициентом диффузии σ^2 . Поскольку для винеровского процесса

$$Me^{iu\zeta(t)} = e^{iuta - \frac{\sigma^2 t u^2}{2}},$$

то, обозначив

$$\xi(t) = \zeta(t) + \eta_0(t) + \xi_1(t), \quad f(t, u) = Me^{iu\xi(t)},$$

найдем

$$\begin{aligned} \ln f(t, u) = t \left\{ iua - \frac{\sigma^2 u^2}{2} + \int_{|x| \leq 1} (e^{iux} - 1 - iux) \mu(dx) + \right. \\ \left. + \int_{|x| > 1} (e^{iux} - 1) \mu(dx) \right\}. \end{aligned} \quad (5)$$

Введем одно важное понятие.

Случайный процесс $\xi(t)$ называется *стохастически непрерывным*, если

$$\xi(t+h) \xrightarrow[h \rightarrow 0]{P} \xi(t).$$

(В частности, если процесс непрерывен в среднем квадратическом, то он стохастически непрерывен.)

Справедлива следующая теорема А. Я. Хинчина.

Теорема. Если $\xi(t)$ — одномерный стохастический непрерывный процесс с независимыми однородными приращениями, причем $\xi(0) = 0$, то логарифм характеристической функции случайной величины $\xi(t)$ имеет вид (5), причем $\mu(dx)$, a , σ^2 определяются этой формулой однозначно.

Доказательство данной теоремы или эквивалентных ей утверждений можно найти во многих курсах теории вероятностей и теории случайных процессов. Изложение наиболее важных аспектов теории процессов с независимыми приращениями можно найти в книге [8].

§ 4. Гауссовские процессы

Гауссовским процессом называется одномерный или многомерный (т. е. со значениями в \mathbb{R}^1 или \mathbb{R}^m) случайный процесс $\xi(t)$, для которого при любом $n \geq 1$ и любых моментах времени t_1, \dots, t_n случайная величина $(\xi(t_1), \dots, \xi(t_n))$ имеет нормальное распределение (невырожденное или вырожденное). Гауссовский процесс в очень многих случаях служит хорошей моделью реального явления в различных областях науки и техники. Кроме того, аппарат гауссовских процессов удобен тем, что множество конечномерных распределений полностью характеризуется заданием двух функций: математического ожидания и корреляционной функции. Пусть $\xi(t) = (\xi_1(t), \dots, \xi_m(t))$, где $\xi_1(t), \dots, \xi_m(t)$ — одномерные случайные процессы. Тогда математическое ожидание $\xi(t)$ — вектор-

ная функция

$$\mathbf{M}\xi(t) = (M\xi_1(t), \dots, M\xi_m(t)),$$

корреляционная функция — матричная функция двух аргументов t, τ :

$$K_\xi(t, \tau) = (K_{i/j}(t, \tau)),$$

где

$$K_{i/j}(t, \tau) = M\xi_i^0(t) \xi_j^0(\tau)$$

и, в свою очередь,

$$\xi_i^0(t) = \xi_i(t) - M\xi_i(t).$$

Ниже будет рассмотрен лишь одномерный гауссовский процесс. Множество конечномерных распределений такого процесса $\xi(t)$ полностью характеризуется математическим ожиданием $M\xi(t)$ и корреляционной функцией $K_\xi(t, \tau) = M\xi^0(t) \xi^0(\tau)$. Вместо этого можно задать характеристическую функцию

$$\varphi_{t_1, t_2}(z_1, z_2) = M \exp \{i(z_1 \xi(t_1) + z_2 \xi(t_2))\},$$

которая в данном случае имеет вид

$$\begin{aligned} \varphi_{t_1, t_2}(z_1, z_2) = \exp \left\{ i(z_1 M\xi(t_1) + z_2 M\xi(t_2)) - \frac{1}{2}(z_1^2 K_\xi(t_1, t_1) + \right. \\ \left. + 2z_1 z_2 K_\xi(t_1, t_2) + z_2^2 K_\xi(t_2, t_2)) \right\}. \end{aligned}$$

В отличие от общих стационарных процессов для гауссовских процессов можно исследовать, помимо линейных преобразований, также широкий класс нелинейных преобразований. Особенно простой вид имеют формулы для математических ожиданий некоторых нелинейных функций гауссовского процесса. Так, пусть $P(x_1, \dots, x_m)$ — полином. Требуется найти значение

$$\delta = MP(\xi(t_1), \dots, \xi(t_m)),$$

где $\xi(t)$ — гауссовский процесс с математическим ожиданием $m(t)$ и корреляционной функцией $K(t_1, t_2)$. Введем ортогональное преобразование

$$\xi = (\xi(t_1), \dots, \xi(t_m))' = C(\eta_1, \dots, \eta_m)',$$

(штрих обозначает транспонирование), при котором новая случайная величина $\eta = (\eta_1, \dots, \eta_m)'$ будет иметь диагональную корреляционную матрицу

$$K_\eta = C' K_\xi C,$$

где K_ξ — корреляционная матрица ξ .

Тогда полином от $\xi(t_1), \dots, \xi(t_m)$ преобразуется в некоторый полином от η , а именно, $P(\xi(t_1), \dots, \xi(t_m)) = Q(\eta_1, \dots, \eta_m)$. Если $M\eta_k = a_k$, $D\eta_k = \sigma_k^2$ и Q есть линейная комбинация выражений $x_1^{i_1} \dots x_m^{i_m}$, то δ будет линейной комбинацией с теми же коэффи-

$$\prod_{k=1}^m \frac{1}{\sqrt{2\pi}\sigma_k} \int_{-\infty}^{\infty} x^r \exp \left\{ -\frac{1}{2\sigma_k^2} (x - a_k)^2 \right\} dx =$$

$$= \prod_{k=1}^m \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (a + \sigma t)^r e^{-t^2/2} dt.$$

В этом произведении сомножители (стоит лишь раскрыть степень суммы) выражаются через моменты стандартного нормального распределения.

Изучая гауссовские процессы, интересуются не только свойствами, связанными с конечномерными распределениями, но и свойствами, связанными с локальным и глобальным поведением траектории процесса. Так, может возникнуть вопрос о дифференцируемости $\xi(t)$ с вероятностью 1, об ограниченности траектории и т. п. Подобные вопросы освещены в сборнике статей «Случайные процессы. Выборочные функции и пересечения» (М., 1978).

Для иллюстрации методов теории гауссовских процессов рассмотрим вывод известной *формулы Райса* для среднего числа пересечений уровня. Пусть $\xi(t)$ — стационарный гауссовский процесс, $M\xi(t) = 0$, $\xi'(t)$ существует с вероятностью 1 и представляет собой стационарный гауссовский процесс. Имеем

$$k_{\xi}(\tau) = \int_0^{\infty} \cos \tau \omega dG(\omega),$$

$$k_{\xi'}(\tau) = \int_0^{\infty} \omega^2 \cos \tau \omega dG(\omega),$$

где $G(\omega)$ — спектральная функция случайного процесса $\xi(t)$. Отсюда видим, что $D\xi(t) = \sigma^2 = G(\infty)$, $D\xi'(t) = \sigma_1^2 = \int_0^{\infty} \omega^2 dG(\omega)$. Очевидно, также, что $M\xi'(t) = 0$. Далее,

$$M\xi(t)\xi'(t) = \lim_{h \rightarrow 0} \frac{1}{h} (M\xi(t)\xi(t+h) - M\xi^2(t)) =$$

$$= \lim_{h \rightarrow 0} \frac{1}{h} [k_{\xi}(h) - k_{\xi}(0)] = k_{\xi}'(0) = - \int_0^{\infty} \omega \sin \tau \omega dG(\omega) |_{\tau=0} = 0.$$

Следовательно, $\xi(t)$ и $\xi'(t)$ — независимые при любом данном t нормальные случайные величины.

Пусть z — любое действительное число. Если $\xi(t) = z$, $\xi'(t) > 0$, то говорят, что в момент t произошло пересечение процессом $\xi(t)$ уровня z снизу вверх. Аналогично при $\xi(t) = z$, $\xi'(t) < 0$ говорят, что в момент t произошло пересечение указанного уровня сверху вниз. Моменты пересечения процессом уровня z снизу вверх образуют ординарный стационарный поток однородных событий.

Обозначим λ его интенсивность, т. е. математическое ожидание числа пересечений в единицу времени. По теореме Королюка λ равна параметру потока, т. е. вероятность пересечения уровня в интервале $(t; t + dt)$ есть λdt . Чему же равна эта вероятность?

Допустим, что $\xi'(t) = y$. Тогда пересечение уровня z в интервале $(t; t + dt)$ произойдет в том и только том случае, если $z - ydt < \xi(t) < z$. Учитывая независимость $\xi(t)$ и $\xi'(t)$, получаем, что при условии $\xi'(t) = y$ условная вероятность пересечения равна $y p_{\xi(t)}(z) dt$. Усреднив по всем $y > 0$, найдем

$$\begin{aligned} \lambda &= p_{\xi(t)}(z) \int_0^{\infty} y p_{\xi'(t)}(y) dy = \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/(2\sigma^2)} \frac{1}{\sqrt{2\pi}\sigma_1} \int_0^{\infty} y e^{-y^2/(2\sigma_1^2)} dy = \\ &= \frac{1}{2\pi} \frac{\sigma_1}{\sigma} e^{-z^2/(2\sigma^2)}. \end{aligned} \quad (1)$$

Это и есть формула Райса. Легко видеть, что интенсивность потока пересечений уровня z сверху вниз также задается выражением (1).

Из формулы Райса следует оценка вероятности события $\{v_T \geq 1\}$, где v_T — число пересечений (например, снизу вверх) в интервале времени $(t_0; t_0 + T)$. Именно, так как $P(v_T \geq 1) \leq M v_T$, то

$$P(v_T \geq 1) \leq \frac{T}{2\pi} \frac{\sigma_1}{\sigma} e^{-z^2/(2\sigma^2)}. \quad (2)$$

§ 5. Марковские процессы

Пусть $\xi(t)$ — многомерный случайный процесс¹. Положим для определенности, что $\xi(t)$ принимает значения из \mathbb{R}^m и что для любых различных $t_1 < t_2 < \dots < t_n$ существует плотность вероятности $p_{t_1, \dots, t_n}(x_1, \dots, x_n)$ случайной величины $(\xi(t_1), \dots, \xi(t_n))$. Тогда можно записать

$$\begin{aligned} &p_{t_1, \dots, t_n}(x_1, \dots, x_n) = \\ &= p_{t_1}(x_1) p_{t_1, t_2}(x_2 | x_1) \dots p_{t_1, \dots, t_n}(x_n | x_1, \dots, x_{n-1}); \end{aligned}$$

в правой части этой формулы фигурируют условные плотности: так, $p_{t_1, t_2}(x_2 | x_1)$ есть условная плотность $\xi(t_2)$ при условии, что $\xi(t_1)$ приняла значение x_1 , $p_{t_1, t_2, t_3}(x_3 | x_1, x_2)$ — условная плотность $\xi(t_3)$ при условии, что $\xi(t_1)$, $\xi(t_2)$ приняли значения x_1 , x_2 , и т. д. Таким образом, в общем случае каждая условная плотность зависит от значений процесса во все предыдущие моменты времени. Свойство, состоящее в том, что условная плотность зависит только от значения процесса в последний момент времени, т. е. для всех $n \geq 2$

$$p_{t_1, \dots, t_n}(x_n | x_1, \dots, x_{n-1}) = f_{t_1, t_2}(x_{n-1}, x_n), \quad (1)$$

¹ Следующее изложение охватывает также процессы $(v(t), \xi(t))$, где $v(t)$ — «дискретная» компонента с конечным или счетным множеством возможных значений, $\xi(t)$ — «непрерывная» компонента со значениями из \mathbb{R}^m .

называется *марковским свойством*. Случайный процесс, обладающий марковским свойством, называется *марковским процессом*. Для такого процесса имеем

$$p_{t_1, \dots, t_n}(x_1, \dots, x_n) = p_{t_1}(x_1) f_{t_1, t_2}(x_1, x_2) f_{t_2, t_3}(x_2, x_3) \times \dots \times f_{t_{n-1}, t_n}(x_{n-1}, x_n). \quad (2)$$

Проинтегрировав по x_1, \dots, x_n по множествам A_1, \dots, A_n , получим

$$P(\xi(t_k) \in A_k, 1 \leq k \leq n) = \int_{A_1} p_{t_1}(x_1) dx_1 \int_{A_2} f_{t_1, t_2}(x_1, x_2) dx_2 \times \dots \times \int_{A_n} f_{t_{n-1}, t_n}(x_{n-1}, x_n) dx_n. \quad (3)$$

Функция $f_{t, \tau}(x, y)$, определенная при $t < \tau$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^m$, называется *плотностью перехода* марковского процесса $\xi(t)$. Функция $P_{t_1, t_2}(x, A) = \int_A f_{t_1, t_2}(x, y) dy$ называется *марковской переходной функцией* данного процесса.

Заметим, что существование плотности фигурирует в нашем определении лишь для упрощения символики. Марковское свойство можно выразить формулой

$$P(\xi(t_k) \in A_k, 1 \leq k \leq n) = \int_{A_1} dF_{\xi(t_1)}(x_1) \int_{A_2} P_{t_1, t_2}(x_1, dx_2) \dots \int_{A_n} P_{t_{n-1}, t_n}(x_{n-1}, dx_n), \quad (4)$$

где $t_1 < \dots < t_n$, A_k — борелевские множества \mathbb{R}^m . Для справедливости этой формулы уже не обязательно, чтобы существовала плотность.

Марковский процесс называется *однородным*, если

$$P_{t_1, t_2}(x, A) = P(t_2 - t_1, x, A). \quad (5)$$

Для этого достаточно, чтобы

$$f_{t_1, t_2}(x, y) = f_{t_2 - t_1}(x, y). \quad (6)$$

Поскольку существует прием, сводящий произвольные марковские процессы к однородным, то обычно в теоретических работах ограничиваются рассмотрением однородных марковских процессов.

Примеры однородного марковского процесса — процесс Пуассона, винеровский процесс. Обозначим через $u_t(x)$ плотность вероятности случайной величины $\xi(t)$ при некоторой заданной плотности $u_0(x)$. Для любого $s < t$

$$u_t(x) = \int u_s(z) f_{s, t}(z, x) dz. \quad (7)$$

Это уравнение, как и его более общая форма

$$U_t(A) = \int U_s(dz) P_{s, t}(z, A), \quad (8)$$

где $U_t(A) = P(\xi(t) \in A)$, называется *уравнением Чепмена — Колмогорова*.

Предположим, что одномерный марковский процесс $\xi(t)$ обладает следующим свойством. При малом $\Delta > 0$

$$M(\xi(t + \Delta) | \xi(t) = x) = x + a(x, t) \Delta + o(\Delta), \quad (9)$$

$$D(\xi(t + \Delta) - \xi(t) | \xi(t) = x) = b(x, t) \Delta + o(\Delta), \quad (10)$$

где $a(x, t)$, $b(x, t)$ — достаточно гладкие функции. Такой процесс называется *диффузионным*. Функции $a(x, t)$ и $b(x, t)$ называются соответственно *локальными коэффициентами сноса и диффузии* данного процесса.

При некоторых дополнительных условиях функция $u_t(x)$ удовлетворяет *уравнению Колмогорова*

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (a(x, t) u_t(x)) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (b(x, t) u_t(x)) \quad (11)$$

и определяется этим уравнением однозначно при $t > t_0$, если задано $u_{t_0}(x)$.

Задачи

1. Пусть $X(t)$ — процесс Пуассона с параметром λ . В моменты t_n скачков $X(t)$ возникают прямоугольные импульсы случайной величины α_n и случайной длительности τ_n . Предполагается, что α_n , τ_n имеют конечные дисперсии и что α_n , τ_n и моменты скачков процесса $X(t)$ независимы в совокупности. Рассмотрим случайный процесс

$$Y(t) = \sum_{t_n < t} \alpha_n \min\{\tau_n, t - t_n\}.$$

Доказать, что последовательность случайных процессов

$$\xi_N(t) = \frac{1}{\sqrt{N}} Y(Nt), \quad N \geq 1,$$

слабо сходится к винеровскому процессу. Найти выражения коэффициентов сноса и диффузии.

2. Пусть $w_1(t), \dots, w_m(t)$ — независимые стандартные винеровские процессы, C — матрица порядка $m \times m$. Найти корреляционные функции случайных процессов

$$\eta(t) = C \begin{pmatrix} w_1(t) \\ \vdots \\ w_m(t) \end{pmatrix},$$

$$\xi(t) = \eta(t + h) - \eta(t).$$

3. Пусть $w(\omega)$ — стандартный винеровский процесс с переменной ω в качестве времени. Найти одномерную плотность случайного процесса

$$\xi(t) = \int_0^T e^{\omega(t-1)} dw(t).$$

4. Пусть $X(t)$ — процесс Пуассона с параметром λ . Обозначим $\tau(n)$ момент n -го скачка $X(t)$. Доказать, что $(\tau(n), n = 0, 1, \dots)$ — процесс с независимыми приращениями (с дискретным временем); найти плотность вероятности приращения этого процесса на отрезке $[m; n]$.

5. В междугородную телефонную станцию поступают заявки на телефонные разговоры. Моменты заявок образуют простейший поток с параметром λ . В заявке указывается длительность разговора в минутах; это случайная величина с распределением f_k ($k = 1, 2, 3, \dots$). Найти формулы для моментов распределения порядка 1, 2, 3, 4 суммарной длительности разговоров, заказанных за время t . Какой вид в данном примере примет формула (3.5)?

6. Обобщить формулу Райса на среднее число пересечений снизу вверх в интервале времени $(0; 2\pi)$ стационарным гауссовским процессом синусоидально изменяющегося уровня $z = c_0 + c \sin t$.

Глава 9

ТЕОРИЯ ВОССТАНОВЛЕНИЯ И ПОЛУМАРКОВСКИЕ ПРОЦЕССЫ

§ 1. Процесс восстановления

Пусть $(\xi_n, n \geq 1)$ — последовательность независимых положительных случайных величин. Обозначим $s_n = \xi_1 + \dots + \xi_n, n \geq 1$. Если все величины ξ_1, ξ_2, \dots имеют общую функцию распределения $P(x) = \mathbf{P}(\xi_n < x)$, то последовательность $(s_n, n \geq 1)$ называется *процессом восстановления*. Случайная величина s_n называется *моментом n -го восстановления*. Термин «процесс восстановления» объясняется очень просто. Допустим, что в момент $t = 0$ включается в работу некоторое устройство. Обозначим ξ_1 время его безотказной работы. В момент $s_1 = \xi_1$ устройство отказывает и заменяется новым, время безотказной работы которого ξ_2 . Таким образом, $s_2 = \xi_1 + \xi_2$ — время второго отказа (и восстановления) и т. д.

Обозначим N_t число восстановлений в интервале $(0; t)$. Справедливо тождество

$$N_t = \max \{n \geq 1 : s_n < t\}. \quad (1)$$

Обозначим через $I_n(t)$ индикатор события $\{s_n < t\}$. Таким образом, $I_n(t) = 1$ в случае $s_n < t$, $I_n(t) = 0$ в противном случае. Тогда, как легко видеть,

$$N_t = \sum_{n=1}^{\infty} I_n(t). \quad (2)$$

Важнейшей характеристикой процесса восстановления является *функция восстановления*

$$H(t) = \mathbf{M}N_t. \quad (3)$$

По теореме Лебега (см. «Дополнение») математическое ожидание суммы ряда из неотрицательных случайных величин равно сумме их математических ожиданий. Применив этот факт к равенству (2), найдем

$$H(t) = \sum_{n=1}^{\infty} \mathbf{P}(s_n < t) = \sum_{n=1}^{\infty} P^{*(n)}(t). \quad (4)$$

Докажем, что ряд в правой части (4) имеет конечную сумму при любом $t \geq 0$, т. е. $H(t)$ всюду конечна. Обозначим

$$\psi(z) = M e^{-z \xi_n}$$

и возьмем любое $z > 0$. Тогда

$$0 < \psi(z) < 1. \quad (5)$$

В самом деле, так как $e^{-zx} < 1$ при $z \geq 0$, $e^{-zx} \leq e^{-z\delta}$ при $x \geq \delta$, то

$$\begin{aligned} \psi(z) &\leq \int_0^\delta dP(x) + e^{-z\delta} \int_\delta^\infty dP(x) = P(\delta) + e^{-z\delta} (1 - P(\delta)) = \\ &= 1 - (1 - e^{-z\delta}) (1 - P(\delta)). \end{aligned}$$

Поскольку $\xi_n > 0$ с вероятностью 1, то $1 - P(\delta) > 0$ для некоторого $\delta > 0$, что и доказывает (5). Итак, при данном z можно положить $\psi(z) = \rho$, $0 < \rho < 1$. В силу независимости ξ_1, \dots, ξ_n

$$M e^{-z s_n} = M e^{-z \xi_1} \dots e^{-z \xi_n} = M e^{-z \xi_1} \dots M e^{-z \xi_n} = \psi^n(z) = \rho^n.$$

В то же время

$$P(s_n < t) \leq \int e^{-z(x-t)} dP^{*(n)}(x) = e^{zt} \rho^n.$$

Отсюда по формуле (4) находим

$$H(t) \leq e^{zt} \sum_{n=1}^{\infty} \rho^n < \infty, \quad (6)$$

что и требовалось доказать.

Поскольку $P(s_n < t)$ — неубывающие функции, то из (4) следует, что $H(t)$ не убывает. Из (6) следует также, что $H(t)$ растет медленнее любой экспоненты (ведь $z > 0$ произвольно).

Введем преобразование Лапласа — Стильтеса

$$\varphi(z) = \int_0^\infty e^{-zt} dH(t)^1. \quad (7)$$

Умножив обе части (4) на e^{-zt} , где $\operatorname{Re} z > 0$, и проинтегрировав в пределах от 0 до ∞ , получим

$$\varphi(z) = \sum_{n=1}^{\infty} \psi^n(z), \quad (8)$$

так как, очевидно,

$$\int_0^\infty e^{-zt} dP^{*(n)}(t) = M e^{-z s_n} = \psi^n(z).$$

¹ Если $F(x)$ — функция распределения неотрицательной случайной величины ξ , то преобразование Лапласа — Стильтеса данного распределения (данной величины) определяется как $M e^{-z \xi}$, т. е. это характеристическая функция с измененным аргументом: $t = iz$. Поэтому на преобразования Лапласа — Стильтеса переносятся основные свойства характеристических функций. Однако интеграл (7) не существует при чисто мнимом z . При анализе положительных случайных величин часто предпочтительнее использовать преобразования Лапласа — Стильтеса, чем характеристические функции.

Пусть $z = u + iv$, $u > 0$, $v \in \mathbb{R}$. Имеем

$$|e^{-zt}| \leq e^{-ut},$$

откуда $|\psi(z)| \leq \psi(u)$. Из (5) заключаем, что ряд (8) сходится, откуда

$$\varphi(z) = \psi(z)/(1 - \psi(z)), \quad \operatorname{Re} z > 0. \quad (9)$$

Пусть, например, ξ_n распределены по экспоненциальному закону с параметром λ . Имеем

$$\psi(z) = \int_0^\infty e^{-zt} \lambda e^{-\lambda t} dt = \lambda/(\lambda + z).$$

Отсюда

$$\varphi(z) = \lambda/((\lambda + z)(1 - \lambda/(\lambda + z))) = \lambda/z.$$

Итак,

$$\frac{\lambda}{z} = \int_0^\infty e^{-zt} d(\lambda t) = \int_0^\infty e^{-zt} dH(t),$$

откуда $H(t) = \lambda t$. Поскольку в данном случае (см. § 4.5) N_t есть пуассоновская случайная величина с параметром λt , то это равенство непосредственно очевидно.

Формулу (9) можно записать еще и так:

$$\varphi(z) = \frac{\psi(z) - \psi^2(z)}{1 - \psi(z)} + \frac{\psi^2(z)}{1 - \psi(z)},$$

или

$$\varphi(z) = \psi(z) + \psi(z)\varphi(z). \quad (10)$$

Произведению преобразований соответствует свертка оригиналов, поэтому из (10) находим

$$dH(t) = dP(t) + \int_0^t dP(\tau) dH(t - \tau), \quad (11)$$

или после интегрирования

$$H(t) = P(t) + \int_0^t H(t - \tau) dP(\tau). \quad (12)$$

Выведенное уравнение (12) называется *уравнением восстановления*.

Если

$$H(t) = \int_0^t h(\tau) d\tau, \quad t \geq 0, \quad (13)$$

то функция $h(t)$ называется *плотностью восстановления*. Если

$P(t) = \int_0^t p(\tau) d\tau$, то обращение равенства (10) имеет вид

$$h(t) = p(t) + \int_0^t h(t - \tau) p(\tau) d\tau. \quad (14)$$

Это уравнение также называется *уравнением восстановления*. Для существования решения уравнения (14) необходимо существование $p(t)$. Если $p(t) = \lambda e^{-\lambda t}$, $t \geq 0$, то $h(t) = \lambda$, $t \geq 0$, что проверяется непосредственной подстановкой в уравнение (14).

Во многих случаях уравнение восстановления в форме (12) или (14) можно решить аналитически или численно.

Разъясним важнейший способ использования функций $H(t)$, $h(t)$.

Пусть в каждый момент s_n восстановления возникает импульс; в момент $t > s_n$ этот импульс производит эффект $g(t - s_n)$. Тогда суммарный эффект, производимый в момент t импульсами, возникшими до этого момента времени,

$$\eta(t) = \sum_{s_n < t} g(t - s_n) = \sum_{n=1}^{\infty} I(s_n < t) g(t - s_n). \quad (15)$$

Средний суммарный эффект

$$M\eta(t) = \sum_{n=1}^{\infty} M I(s_n < t) g(t - s_n), \quad (16)$$

если данный ряд сходится абсолютно. Имеем

$$M I(s_n < t) g(t - s_n) = \int_0^t g(t - \tau) dP^{*(n)}(\tau). \quad (17)$$

Если $g(t)$ — ступенчатая функция, то правая часть (17) представляет собой линейную комбинацию конечного числа значений функции $P^{*(n)}(t)$ в некоторых точках. Суммирование по n в соответствии с формулами (16), (17) приводит к равенству

$$M\eta(t) = \int_0^t g(t - \tau) dH(\tau) \quad (18)$$

сначала для ступенчатых функций $g(t)$, а затем и для произвольных непрерывных $g(t)$, так как любую непрерывную функцию можно аппроксимировать ступенчатой с точностью до любого заданного $\varepsilon > 0$, а для погрешности $r(t)$ этой аппроксимации

$$\left| \int_0^t r(t - \tau) dH(\tau) \right| \leq \varepsilon \int_0^t dH(\tau) = \varepsilon H(t).$$

Итак, средний суммарный эффект задается формулой (18) или ее аналогом

$$M\eta(t) = \int_0^t g(t - \tau) h(\tau) d\tau. \quad (19)$$

Пусть, например, рассматривается функционирование системы, сменяющей попеременно исправное состояние состоянием отказа. Именно, система исправна в интервале длины α_1 , затем неисправна в интервале длины β_1 , после чего идет интервал исправности

длины α_2 , интервал неисправности длины β_2 и т. д. до бесконечности. Предположим, что $\alpha_n, \beta_n, n \geq 1$, независимы в совокупности и имеют распределения $F_{\alpha_n}(t) = A(t)$, $F_{\beta_n}(t) = B(t)$. Требуется найти выражение для вероятности $f(t)$ исправного состояния системы в момент t .

Рассмотрим функции

$$b_n(\tau) = \begin{cases} 1 & \text{при } 0 \leq \tau < \alpha_n, \\ 0 & \text{в противном случае.} \end{cases}$$

Обозначив через $I(t)$ индикатор исправного состояния системы в момент t , получим

$$I(t) = \sum_{n=0}^{\infty} b_n(t - s_n), \quad (20)$$

где $s_0 = 0$. (Проверку формулы (20) предоставляем читателю в качестве упражнения). Далее, средний эффект n -го импульса через время τ после его начала

$$Mb_n(\tau) = 1 - F(\tau).$$

Отсюда

$$f(t) = MI(t) = 1 - F(t) + \int_0^t (1 - F(t - \tau)) dH(\tau). \quad (21)$$

В этой формуле $H(t)$ — функция восстановления, соответствующая процессу восстановления ($s_n, n \geq 1$), где

$$s_n = \sum_{k=1}^n (\alpha_k + \beta_k).$$

Рассмотрим еще следующую задачу. Предположим, что эксплуатируется некоторое устройство; $u(\tau)$ — производительность устройства через время τ после начала его эксплуатации, причем $u(\tau)$ — случайная величина, монотонно убывающая по τ с вероятностью 1. Если производительность падает до уровня u_0 , устройство заменяется новым; поведение $u(\tau)$ для различных «копий» устройства происходит независимым образом. Пусть $f(t)$ — средняя производительность в момент t . Тогда

$$f(t) = \Phi(t) + \int_0^t \Phi(t - \tau) dH(\tau),$$

где $H(t)$ — функция восстановления, соответствующая процессу замены системы, $\Phi(\tau) = M(u(\tau); u(\tau) > u_0)$.

В заключение отметим одно полезное тождество для событий, связанных с процессом восстановления:

$$\{s_n < t\} = \{N_t \geq n\}. \quad (22)$$

Перейдя к противоположным событиям, можно записать

$$\{s_n \geq t\} = \{N_t < n\}. \quad (23)$$

§ 2. Теоремы восстановления

Большое теоретическое и практическое значение имеют предельные теоремы, связанные с процессом восстановления. Простейшая из них называется *элементарной теоремой восстановления* и состоит в том, что

$$\frac{1}{t} H(t) \xrightarrow{t \rightarrow \infty} \frac{1}{M\xi_n} \quad (1)$$

(не исключается в случай $M\xi_n = \infty$).

Ограничимся доказательством соотношения (1) для частного случая, когда $M\xi_n^4 < \infty$. В этом случае, обозначив $\tau = M\xi_n$, $\xi_n^0 = \xi_n - \tau$, $s_n^0 = \xi_n^0 + \dots + \xi_n^0$, найдем

$$M(s_n^0)^4 = M(\xi_n^0 + \dots + \xi_n^0)^4 = \sum_{i,j,k,l} M\xi_i^0 \xi_j^0 \xi_k^0 \xi_l^0. \quad (2)$$

Если хотя бы один из индексов i, j, k, l не совпадает со всеми остальными индексами, то в правой части (2), вследствие независимости ξ_1, \dots, ξ_n , выделяется множитель (скажем, $M\xi_i^0$), равный нулю. Остается лишь n слагаемых, для которых $i = j = k = l$, и $3n(n-1)$ слагаемых, для которых либо $i = j \neq k = l$, либо $i = k \neq j = l$, либо $i = l \neq j = k$. Отсюда

$$M(s_n^0)^4 = nM(\xi_1^0)^4 + 3n(n-1)(D\xi_1)^2 \leq cn^2. \quad (3)$$

Имеем

$$P(|s_n - n\tau| \geq \rho n) \leq M(s_n^0)^4 / (\rho n)^4 \leq c / (\rho^4 n^2). \quad (4)$$

Отсюда

$$H(t) = \sum_{n=1}^{\infty} P(s_n < t) \leq N + \sum_{n=N+1}^{\infty} P(s_n < t).$$

В этом соотношении возьмем N из условия

$$\frac{t(1+\varepsilon)}{\tau} - 1 \leq N < \frac{t(1+\varepsilon)}{\tau}.$$

Тогда для любого слагаемого суммы в правой части (5) имеем

$$\begin{aligned} P(s_n < t) &= \\ &= P(s_n^0 < -(n\tau - t)) \leq P(|s_n^0| \geq n\tau - t) \leq c / \left(\left(\tau - \frac{t}{n} \right)^4 n^2 \right), \end{aligned}$$

а так как $n \geq t(1+\varepsilon)/\tau$, то $\tau - \frac{t}{n} \geq \tau\varepsilon/(1+\varepsilon)$, а следовательно,

$$P(s_n < t) \leq c \left(\frac{1+\varepsilon}{\tau\varepsilon} \right)^4 n^{-2}. \quad (6)$$

Подстановка (6) в (5) с учетом того, что ряд $\sum n^{-2}$ сходится, приводит к оценке

$$H(t) \leq \frac{t(1+\varepsilon)}{\tau} + c_1, \quad (7)$$

откуда вследствие произвольности $\varepsilon > 0$ получаем

$$\overline{\lim}_{t \rightarrow \infty} \frac{1}{t} H(t) \leq \frac{1}{\tau}. \quad (8)$$

Аналогично

$$H(t) \geq \sum_{n=1}^N P(s_n < t) = N - \sum_{n=1}^N P(s_n \geq t). \quad (9)$$

На этот раз возьмем N из условия

$$\frac{t(1-\varepsilon)}{\tau} - 1 \leq N < \frac{t(1+\varepsilon)}{\tau}.$$

Тогда для любого $n \leq N$

$$\begin{aligned} P(s_n \geq t) &= P(s_n^0 \geq t - n\tau) \leq P(|s_n^0| \geq n\left(\frac{t}{n} - \tau\right)) \leq P(|s_n^0| \geq n\tau\varepsilon) \leq \\ &\leq c/((\tau\varepsilon)^4 n^2). \end{aligned} \quad (10)$$

Теперь из (9) заключаем, что

$$H(t) \geq \frac{t(1-\varepsilon)}{\tau} - c', \quad (11)$$

где c' — некоторая положительная постоянная. Оценка (11) означает, что

$$\lim_{t \rightarrow \infty} \frac{1}{t} H(t) \geq \frac{1}{\tau}. \quad (12)$$

Из (8) и (12) вытекает (1).

Узловая теорема восстановления. Пусть ξ_n имеют нерешетчатое распределение, т. е. не принадлежат множеству $\{n\Delta, n \in \mathbb{Z}\}$ с вероятностью 1, и $Q(t), t \geq 0$, — монотонная функция, интегрируемая в интервале $(0; \infty)$. Тогда

$$\lim_{t \rightarrow \infty} \int_0^t Q(t-\tau) dH(\tau) = \frac{1}{\tau} \int_0^\infty Q(t) dt, \quad (13)$$

где $\tau = M\xi_n$.

Доказательство опускаем (см. [36, т. 2], где имеется доказательство более общей теоремы).

С л е д с т в и е. Для любого фиксированного Δ

$$H(t + \Delta) - H(t) \xrightarrow{t \rightarrow \infty} \Delta/\tau. \quad (14)$$

Данное следствие получается из (13): достаточно положить $Q(t)$, равным 1 в интервале $(0; |\Delta|)$ и равным 0 вне этого интервала.

§ 3. Стационарный процесс восстановления

В предыдущем параграфе процесс восстановления был определен как множество моментов времени $s_n, n \geq 1$, для которых $s_n = \xi_1 + \dots + \xi_n$, причем случайные величины ξ_n независимы и

имеют одну и ту же функцию распределения $P(t)$. Общим процессом восстановления назовем множество моментов времени s_n , $n \geq 1$, для которых $s_n = \xi_1 + \dots + \xi_n$, ξ_n — независимы и имеют функции распределения $P_0(t)$ при $n = 1$ и $P(t)$ при $n \geq 2$. Таким образом, общий процесс восстановления превращается в процесс восстановления при $P_0(t) = P(t)$.

Замечание. Элементарная и узловая теоремы восстановления справедливы также для общего процесса восстановления. Доказательство этого основано на том простом факте, что математическое ожидание $\bar{H}(t)$ числа восстановлений в интервале $(0; t)$ для общего процесса восстановления, как легко показать, удовлетворяет уравнению

$$\bar{H}(t) = P_0(t) + \int_0^t H(t-\tau) dP_0(\tau), \quad (1)$$

а следовательно, предельные свойства функций $\bar{H}(t)$ и $H(t)$ при $t \rightarrow \infty$ совпадают.

Возьмем некоторый момент $\theta \geq 0$. Течение процесса после момента θ вполне определяется последовательностью моментов времени $s_1(\theta) < s_2(\theta) < \dots$, где $s_n(\theta)$ — момент n -го восстановления после момента θ . (Ясно, что $s_n(0) = s_n$.) Очевидно, случайные величины $\xi_1(\theta) = s_1(\theta)$, $\xi_2(\theta) = s_2(\theta) - s_1(\theta)$, $\xi_3(\theta) = s_3(\theta) - s_2(\theta)$, ... независимы, причем $\xi_n(\theta)$, $n \geq 2$, имеют функцию распределения $P(t)$, $\xi_1(\theta)$ — некоторую функцию распределения $P_\theta(t)$. Общий процесс восстановления называется *стационарным процессом восстановления*, если

$$P_\theta(t) = P_0(t), \quad \theta > 0. \quad (2)$$

Таким образом, стационарный процесс восстановления характеризуется тем, что его течение после произвольного момента $\theta \geq 0$ описывается одними и теми же вероятностными законами.

Возьмем произвольное $t > 0$ и рассмотрим полуинтервал $[0; Nt) = [0; t) \cup [t; 2t) \cup \dots \cup [(N-1)t; Nt)$. Математическое ожидание числа восстановлений в любом полуинтервале $[kt; (k+1)t)$ в силу стационарности равно $\bar{H}(t)$. Итак,

$$\bar{H}(Nt) = N\bar{H}(t). \quad (3)$$

Отсюда

$$\frac{1}{t} \bar{H}(t) = \frac{1}{Nt} \bar{H}(Nt) \xrightarrow{N \rightarrow \infty} \frac{1}{\tau}. \quad (4)$$

Следовательно, для стационарного процесса восстановления

$$\bar{H}(t) = t/\tau, \quad t \geq 0, \quad (5)$$

т. е.

$$\bar{h}(t) = 1/\tau, \quad t \geq 0, \quad (6)$$

а следовательно,

$$\bar{\varphi}(z) = \frac{1}{\tau z}, \quad \operatorname{Re} z > 0, \quad (7)$$

где $\bar{\varphi}(z)$ — преобразование Лапласа — Стильеса функции $\bar{H}(t)$.

Из уравнения (1)

$$\begin{aligned} \bar{\varphi}(z) &= \psi_0(z) + \psi_0(z) \varphi(z) = \psi_0(z) + \\ &+ \psi_0(z) \psi(z) / [1 - \psi(z)] = \psi_0(z) / (1 - \psi(z)), \end{aligned} \quad (8)$$

где $\psi_0(z)$ — преобразование Лапласа—Стилтьеса распределения $P_0(t)$. Из (7) и (8) следует равенство

$$\psi_0(z) = [1 - \psi(z)] / (\tau z). \quad (9)$$

Обращение преобразования Лапласа — Стилтьеса в виде (9) приводит к формуле

$$P_0(t) = \frac{1}{\tau} \int_0^t (1 - P(\tau)) d\tau. \quad (10)$$

Итак, условием стационарности общего процесса восстановления является равенство (10).

Стационарный процесс восстановления называется также *поток Пальма*; величина s_n называется моментом n -го события потока, $s_n(\theta)$ — моментом n -го события после момента θ . *Параметром потока* называется число

$$\mu = \lim_{\Delta \rightarrow 0} P_0(\Delta) / \Delta. \quad (11)$$

Таким образом, вероятность наличия хотя бы одного события потока в малом интервале Δ эквивалентна $\mu\Delta$. *Интенсивностью потока* называется число

$$\lambda = \bar{H}(\Delta) / \Delta. \quad (12)$$

Поскольку $P(t)$ — функция распределения положительной случайной величины, то $P(t) \rightarrow 0$ при $t \rightarrow 0$. Из (10) находим, что $P_0(\Delta) \sim \Delta/\tau$ при малых Δ . Отсюда $\mu = 1/\tau$. В то же время, как было выяснено, $\bar{H}(t) = t/\tau$, откуда $\lambda = 1/\tau$. Итак,

$$\lambda = \mu, \quad (13)$$

т. е. интенсивность потока Пальма совпадает с его параметром.

Имеем

$$\begin{aligned} \bar{H}(t) &= \mathbf{M}N_t = \sum_{n=1}^{\infty} n \mathbf{P}(N_t = n) = \sum_{n=1}^{\infty} \mathbf{P}(N_t = n) + \\ &+ \sum_{n=2}^{\infty} (n-1) \mathbf{P}(N_t = n) \geq \sum_{n=1}^{\infty} \mathbf{P}(N_t = n) + \sum_{n=2}^{\infty} \mathbf{P}(N_t = n) = \\ &= \mathbf{P}(N_t \geq 1) + \mathbf{P}(N_t \geq 2). \end{aligned} \quad (14)$$

При малых t имеем $\bar{H}(t) \sim \lambda t$ и $\mathbf{P}(N_t \geq 1) = P_0(t) \sim \lambda t$. На основании (14) теперь получаем

$$\mathbf{P}(N_t \geq 2) = o(t), \quad t \rightarrow 0, \quad (15)$$

или, что то же самое,

$$\frac{1}{\Delta} \mathbf{P}(N_{\Delta} \geq 2) \xrightarrow{\Delta \rightarrow 0} 0. \quad (16)$$

Свойство, выражаемое данными соотношениями, называется *ординарностью потока*. Несколько грубо можно сказать так: если поток ординарен, то вероятностью двух или большего числа событий в малом интервале времени можно пренебречь по сравнению с вероятностью единственного события в этом интервале. Потоки Пальма получили широкое применение в теории массового обслуживания. Формула (10) — частный случай *формул Пальма — Хинчина* из теории потоков однородных событий.

Возможен ли случай, когда процесс восстановления стационарен и в то же время $P_0(t) = P(t)$, или, что то же самое, $\psi_0(z) = \psi(z)$? Из формулы (9) находим, что в этом случае

$$\psi(z) = (1 - \psi(z))/(\tau z),$$

т. е.

$$\psi(z) = 1/(1 + \tau z), \quad (17)$$

откуда

$$P(t) = 1 - e^{-t/\tau}, \quad t \geq 0. \quad (18)$$

Согласно § 4.5, в этом случае ($s_n, n \geq 1$) — простейший поток однородных событий с параметром $\lambda = 1/\tau$. Мы обнаруживаем, что в этом случае время от произвольного момента времени до следующего события потока имеет то же распределение, что и интервал между двумя последовательными событиями.

§ 4. Полумарковский процесс

Пусть X — конечное или счетное множество, $(p_i^{(0)}, i \in X)$ — распределение вероятностей, $(P_{ij}(t), j \in X)$ — при каждом $i \in X$ распределение вероятностей случайной величины $(v^{(i)}, \xi^{(i)})$, а именно:

$$P_{ij}(t) = P(v^{(i)} = j, \xi^{(i)} < t).$$

(1)

Определим последовательность

$$v_0, \xi_1, v_1, \xi_2, v_2, \xi_3, \dots \quad (2)$$

по такому вероятностному закону. Величина $v_0 \in X$ имеет распределение $(p_i^{(0)})$. При условии, что $v_n = i$ и величины $\xi_k, k \leq n, v_k, k < n$, приняли произвольные фиксированные значения, величина (v_{n+1}, ξ_{n+1}) имеет распределение $(P_{ij}(t))$. По последовательности (2) однозначно определяется случайная функция времени $v(t)$, которую и назовем

полумарковским процессом, а именно:

$$\begin{aligned} v(t) &= v_0, \quad 0 \leq t < \xi_1; \quad v(t) = v_n, \quad \xi_1 + \dots + \xi_n \leq \\ &\leq t < \xi_1 + \dots + \xi_{n+1}, \quad n \geq 1. \end{aligned} \quad (3)$$

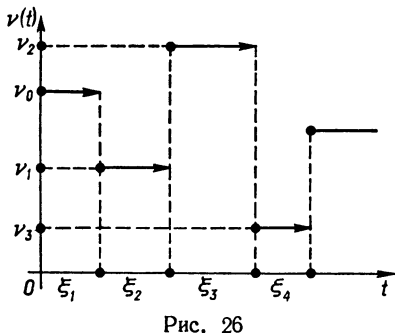


Рис. 26

Таким образом, полумарковский процесс принимает значения из конечного или счетного множества X ; траектории этого процесса — непрерывные справа ступенчатые функции. Очевидно, любую такую траекторию можно задать, указав последовательные состояния процесса v_0, v_1, v_2, \dots и длительности $\xi_1, \xi_2, \xi_3, \dots$ пребывания в этих состояниях (рис. 26). Траектория, согласно описанному выше вероятностному закону, последовательно определяется в интервалах длины ξ_1, ξ_2, \dots , причем в момент t перехода «забывается» все прошлое: дальнейшая часть траектории определяется только состоянием в момент t .

Заметим, что ξ_i и $\xi^{(i)}$ — не одно и то же: ξ_i есть i -й интервал времени, $\xi^{(i)}$ — время пребывания процесса в состоянии i .

Пример 1. Устройство безотказно работает в течение времени $\xi^{(0)}$ с функцией распределения $A(t)$, после чего отказывает. Время восстановления — случайная величина $\xi^{(1)}$ с функцией распределения $B(t)$. Затем все повторяется, но всякий раз значения $\xi^{(0)}, \xi^{(1)}$ независимы от всего предыдущего. Обозначим $v(t) = 0$, если в момент t устройство исправно, $v(t) = 1$ в противном случае. Тогда $v(t)$ — полумарковский процесс.

Имеем $X = \{0, 1\}$, $p_0^{(0)} = 1$, $p_1^{(0)} = 0$ (поскольку при $t = 0$ устройство исправно); $P_{01}(t) = A(t)$, $P_{00}(t) = C$, $P_{10}(t) = B(t)$, $P_{11}(t) = 0$.

Пример 2. Система состоит из m элементов. Вначале все элементы исправны. Длительность безотказной работы k -го элемента — случайная величина $\alpha^{(k)}$ с функцией распределения $A_k(t)$. В момент, когда какой-либо элемент отказал, начинается замена системы. Длительность замены — случайная величина $\beta^{(k)}$ с функцией распределения $B_k(t)$, где k — номер отказавшего элемента. После окончания замены все повторяется с новыми, независимыми от предыдущего значениями $\alpha^{(k)}, \beta^{(k)}$.

В данном случае имеем процесс с $m + 1$ состояниями, а именно: состояние 0 означает исправное состояние всех элементов, состояние k ($1 \leq k \leq m$) — замену системы после отказа k -го элемента. Таким образом, $X = \{0, 1, \dots, m\}$. Далее, $p_0^{(0)} = 1$, $p_k^{(0)} = 0$, $1 \leq k \leq m$. Очевидно, что

$$P_{i0}(t) = B_i(t), \quad 1 \leq i \leq m;$$

$$P_{ij}(t) = 0, \quad 1 < i \leq m, \quad 1 \leq j \leq m;$$

$$P_{00}(t) = 0.$$

Несколько сложнее определяются $P_{0j}(t)$ при $1 \leq j \leq m$. Переход из 0 в j в интервале $(t; t + dt)$, отсчитываемого от момента попадания в состояние 0, происходит в том и только том случае, если $t < \alpha^{(i)} < t + dt$ и $\alpha^{(k)} > \alpha^{(i)}$ при всех $k \neq j$. Итак,

$$P_{0j}(t) = \int_0^t \prod_{k \neq j} (1 - A_k(\tau)) dA_j(\tau), \quad 1 \leq j \leq m. \quad (4)$$

Общий полумарковский процесс $v(t)$ определяется следующим образом. Задано распределение $(P_{ij}^{(0)}(t))$ случайной величины (δ, v_0, ξ_0) , а именно,

$$P_{ij}^{(0)}(t) = P(\delta = i, v_0 = j, \xi_0 < t).$$

При $0 \leq t < \xi_0$ полагаем $v(t) = \delta$; при $t \geq \xi_0$ $v(t)$ равно значе-

нию в момент $t - \xi_0$ полумарковского процесса с начальным состоянием v_0 .

Заметим, что понятие общего полумарковского процесса так же относится к понятию полумарковского процесса, как понятие общего процесса восстановления к понятию процесса восстановления. Именно, отличие состоит в некотором начальном периоде.

Общий полумарковский процесс $v(t)$, как легко видеть, обладает следующим свойством. При любом фиксированном $\theta \geq 0$ случайный процесс $v_\theta(t) = v(t + \theta)$, $t \geq 0$, также является общим полумарковским процессом. Если для этого процесса функции $P_{ij}^{(0)}(t)$, $P_{ij}(t)$ — те же, что и для процесса $v(t)$, случайный процесс $v(t)$ называется *стационарным полумарковским процессом*.

Основной интерес для приложений представляет стационарный полумарковский процесс. Важно знать соотношения между $(P_{ij}^{(0)}(t))$ и $(P_{ij}(t))$ для такого процесса. Для их вывода используем узловую теорему восстановления. Последовательность $(s_n^{(i)})$ моментов попадания процесса $v(t)$ в состояние i (т. е. таких моментов $\xi_0 + \dots + \xi_n$, для которых $v(\xi_0 + \dots + \xi_n) = i$) есть общий процесс восстановления. Легко убедиться в том, что это стационарный процесс восстановления. В таком случае его плотность восстановления h_i есть постоянная.

Возьмем произвольное $T > 0$. В силу стационарности процесса $P_{ij}^{(0)}(t)$ есть вероятность того, что произойдет одно из двух событий: 1) в некоторый момент $\tau < T$ произошел переход $v(t)$ в состояние i ; после этого пребывание в состоянии i продлилось время, большее $T - \tau$, но меньшее $T - \tau + t$, и закончилось переходом в состояние j ; 2) $\delta = i$, $T < \xi_0 < T + t$, $v_0 = j$. Очевидно, вероятность осуществления второго события бесконечно мала при $T \rightarrow \infty$. Вероятность первого события можно записать так:

$$\begin{aligned} & \int_0^T (P_{ij}(T - \tau + t) - P_{ij}(T - \tau)) h_i d\tau = \\ & = h_i \int_0^T (P_{ij}(\infty) - P_{ij}(T - \tau)) d\tau - \\ & - h_i \int_0^T (P_{ij}(\infty) - P_{ij}(T - \tau + t)) d\tau = \\ & = h_i \int_0^t (P_{ij}(\infty) - P_{ij}(x)) dx - h_i \int_{\frac{T}{t}}^{T+t} (P_{ij}(\infty) - P_{ij}(x)) dx. \end{aligned}$$

Вычитаемое, очевидно, стремится к 0 при $T \rightarrow \infty$. Итак,

$$P_{ij}^{(0)}(t) = h_i \int_0^t (P_{ij}(\infty) - P_{ij}(x)) dx. \quad (5)$$

Остается определить постоянные h_i , $i \in X$. Заметим, что при малом Δ выражение $h_i \Delta$ в силу ординарности потока эквивалентно вероятности перехода процесса в состояние i в интервале $(T; T + \Delta)$. Пренебрегая, как и в предыдущем случае, возможностью события $\{\xi_0 > T\}$, получим

$$\begin{aligned} h_i \Delta &\sim \sum_{i \in X} \int_0^T (P_{ij}(T - \tau + \Delta) - P_{ij}(T - \tau)) h_i d\tau = \\ &= \sum_{i \in X} h_i \left(\int_T^{T+\Delta} P_{ij}(x) dx - \int_0^\Delta P_{ij}(x) dx \right) \sim \sum_{i \in X} h_i P_{ij}(\infty). \end{aligned}$$

(В этих выкладках некоторые подробности предельных переходов опущены.)

Итак, для h_i , $i \in X$, имеем систему уравнений

$$h_j = \sum_{i \in X} h_i P_{ij}(\infty), \quad j \in X. \quad (6)$$

Из (5) находим условие нормировки

$$1 = \sum_{i, j \in X} P_{ij}^{(0)}(\infty) = \sum_{i \in X} h_i \sum_{j \in X} \int_0^\infty (P_{ij}(\infty) - P_{ij}(x)) dx.$$

Поскольку

$$\sum_{j \in X} P_{ij}(\infty) = 1, \quad \sum_{j \in X} P_{ij}(x) = \mathbf{P}(\xi^{(i)} < x),$$

то

$$\sum_{i \in X} \int_0^\infty (P_{ij}(\infty) - P_{ij}(x)) dx = \int_0^\infty \mathbf{P}(\xi^{(i)} > x) dx = \tau_i. \quad (7)$$

Отсюда

$$\sum_{i \in X} h_i \tau_i = 1. \quad (8)$$

Во многих интересных для приложений случаях решение системы уравнений (6) с условием (8) единственно, и тогда $P_{ij}^{(0)}(t)$ находятся по формуле (5). В формуле (7) фактически содержится выражение для вероятностей состояний стационарного полумарковского процесса, а именно,

$$\mathbf{P}(v(T) = i) = h_i \tau_i, \quad i \in X. \quad (9)$$

Пример. Пусть в условиях примера 1 случайный процесс $v(t)$ — стационарный полумарковский. Обозначим

$$\tau_0 = \int_0^\infty (1 - A(t)) dt, \quad \tau_1 = \int_0^\infty (1 - B(t)) dt.$$

Тогда, поскольку в данном случае $P_{01}(\infty) = 1$, $P_{00}(\infty) = 0$, $P_{10}(\infty) = 1$, $P_{11}(\infty) = 0$, уравнения (6) и (8) принимают вид $h_0 = h_1$, $h_0 \tau_0 + h_1 \tau_1 = 1$. Отсюда

$$h_0 = h_1 = \frac{1}{\tau_0 + \tau_1}.$$

а следовательно,

$$P(v(T) = 0) = \tau_0 / (\tau_0 + \tau_1) \quad (10)$$

(это выражение называется коэффициентом готовности устройства);

$$P(v(T) = 1) = \tau_1 / (\tau_0 + \tau_1).$$

Коэффициент надежности, равный вероятности исправного состояния устройства на отрезке времени $[T; T + t]$, задается формулой

$$R_T(t) = \frac{1}{\tau_0 + \tau_1} \int_t^{\infty} (1 - A(x)) dx, \quad (11)$$

поскольку

$$R_T(t) = P_{01}^{(0)}(\infty) - P_{01}^{(0)}(t).$$

Задачи

1. Пусть в обозначениях § 1 $\psi(z)$ — дробно-рациональная функция. На основании разложения Хевисайда из операционного исчисления получим явную формулу для $h(t)$. Довести выкладки до конца в случае $\psi(z) = (1+z)/(1+2z)^2$.

2. Время безотказной работы устройства имеет плотность вероятности $a(t)$. Через каждые T единиц времени после включения устройства производится контроль: если устройство отказало, оно заменяется новым. Процесс замены повторяется сколь угодно долго. Пусть $N(t)$ — число отказов устройства за время t . Найти предел по вероятности

$$\lim_{t \rightarrow \infty} \frac{1}{t} N(t).$$

3. Решить предыдущую задачу со следующим изменением в условии. Устройство заменяется новым не только при обнаружении отказа, но и в том случае, если к моменту контроля оно уже проработало rT единиц времени.

4. После включения в работу устройства через время γ с плотностью вероятности $u(x)$ производится его «плановая» замена, занимающая случайное время β_1 с плотностью $v_1(x)$. Если, тем не менее, устройство отказало ранее, производится восстановление продолжительностью β_0 с плотностью $v_0(x)$, после чего включается новое устройство. Время безотказной работы устройства — случайная величина α с плотностью $a(x)$. Предположим, что все указанные случайные величины независимы; при включении каждого нового устройства их значения не зависят от предыдущего. Пусть также математические ожидания величин β_1 , β_2 и хотя бы одной из величин α , γ конечны. Найти предел при $t \rightarrow \infty$ вероятности того, что в момент t : а) устройство исправно; б) устройство исправно и с момента его включения прошло время, не большее Δ .

5. Состояние технического устройства в момент t описывается стационарным полумарковским процессом с состояниями 0, 1. Попад в состояние 0, процесс находится в нем экспоненциально распределенное с параметром λ время, попав в состояние 1 — постоянное время τ . Вероятности перехода: $p_{00} = \alpha$, $p_{01} = 1 - \alpha$, $p_{10} = 1 - \beta$, $p_{11} = \beta$. Найти $P_{ij}(t)$, $P_{ij}^{(0)}(t)$, а также пределы вероятности состояний в момент t при $t \rightarrow \infty$.

6. Найти математическое ожидание и корреляционную функцию случайного процесса, описанного в задаче 5.

7. Пусть $X(t)$ — процесс Пуассона с параметром λ , (s_n) — независимый от него процесс восстановления. Случайные величины $s_n - s_{n-1}$ имеют плотность и конечное математическое ожидание. Определим случайный процесс $Y(t)$ соотношениями

$$Y(t) = X(t) \quad \text{при} \quad 0 \leq t \leq s_1,$$

$$Y(t) = X(t) - X(s_n - 0) \quad \text{при} \quad s_n \leq t < s_{n+1}, \quad n = 1, 2, \dots$$

Найти пределы

$$\lim_{t \rightarrow \infty} P(Y(t) = k), \quad k \geq 0.$$

8. *Обрывающийся процесс восстановления* есть конечная последовательность случайных величин s_1, s_2, \dots, s_v , где $s_1 = \xi_1, s_2 = \xi_1 + \xi_2, \dots; \xi_k$ — независимые положительные случайные величины, v — независимая от (ξ_n) случайная величина с геометрическим распределением

$$P(v = k) = q p^{k-1}, \quad k \geq 1, \quad q = 1 - p > 0.$$

Б. дем. понимать под $H(t), h(t), \Phi(t), \phi(t)$ то же, что и для процесса восстановления (§ 1).

Найти аналоги:

а) формуле (1.9);

б) уравнению (1.14).

Вывести формулу для преобразования Лапласа — Стильтеса случайной величины s_v , а также для $M s_v$ и $D s_v$.

Глава 10

ТЕОРИЯ МАССОВОГО ОБСЛУЖИВАНИЯ

§ 1. Системы обслуживания

В промышленности, транспорте, связи, в бытовом обслуживании и т. п. часто наблюдается явление *скоплений*. Оно состоит в том, что возникает массовый спрос на выполнение какой-либо операции, немедленное удовлетворение которого не всегда возможно из-за случайностей как в процессе образования спроса, так и в процессе выполнения самой операции. Возникают очереди, заторы и тому подобные явления. Изучением скоплений методами теории вероятностей и занимается *теория массового обслуживания* (теория очередей, теория скоплений). Прикладное значение теории — в прогнозировании различного рода потерь из-за скоплений (например, времени простоя автомобилей, станков, задержек в выполнении вычислительного процесса) и оптимизация системы, предназначенной для удовлетворения массового спроса (системы обслуживания) с целью минимизации потерь.

Теория массового обслуживания изучает реальные скопления посредством построения математических моделей. В каждой конкретной постановке задачи в качестве модели, описывающей реально наблюдаемое явление, выбирается некоторая система обслуживания.

В соответствии с потребностями практики исследованы многие сотни типов систем обслуживания. Многие разделы теории массового обслуживания (теория коммутационных сетей, теория запасов, теория транспортных потоков и другие) превратились в самостоятельные развитые теории. В настоящей главе будут рассмотрены лишь наиболее характерные системы обслуживания, ставшие классическими. На них можно продемонстрировать основные идеи и методы теории массового обслуживания.

Система обслуживания включает некоторое число *каналов*, выполняющих обслуживание *требований*. Требование — общее название индивидуального обращения в систему за какой-либо опера-

цией. Например, требованием является телефонный вызов. Для упрощения словесного описания системы часто требование отождествляется с его носителем (например, клиентом предприятия массового обслуживания). Канал — общее название устройства, способного одновременно обслуживать лишь одно требование. В классической модели принимается, что канал обслуживает требование в течение случайного времени η с функцией распределения $B(t)$; длительности η_n обслуживания различных требований независимы. Если в системе имеется несколько каналов, то длительности обслуживания на них также независимы. Система называется *одноканальной*, *двухканальной* и т. п., судя по числу m имеющихся каналов. Не исключается и случай $m = \infty$. Наиболее изучены одноканальные системы обслуживания. Все каналы идентичны в том смысле, что распределение времени обслуживания для них одно и то же. В классической системе обслуживания поступающее требование может немедленно занять любой свободный канал; это свойство называется *полностью доступностью*. Кроме каналов, в системе имеется некоторое число r *мест для ожидания*. Если в момент поступления требования все каналы заняты, требование занимает одно из свободных мест для ожидания. Требования, занимающие места для ожидания, образуют *очередь*. Когда канал освобождается, его занимает требование из очереди, поступившее ранее других. Если в момент поступления требования заняты как все каналы, так и все места для ожидания, требование теряется (получает *отказ*). При $r = \infty$ имеем *систему с ожиданием*, при $r = 0$ — *систему с потерями*, при $1 \leq r < \infty$ — *систему с ограниченной очередью*. Под *величиной очереди* по традиции понимают общее число требований, находящихся в системе, т. е. обслуживаемых и ожидающих начала обслуживания.

Для кодирования систем обслуживания принята *классификация Д. Кендалла*. Согласно этой классификации, система обслуживания кодируется словом

$$\bigcirc \mid \bigcirc \mid m \mid r,$$

где на месте первого кружочка ставится символ входящего потока, на месте второго — символ распределения времени обслуживания, m и r определены выше. При $r = \infty$ систему кодируют проще:

$$\bigcirc \mid \bigcirc \mid m.$$

Употребительные символы входящего потока:

G — общий стационарный поток (неубывающая последовательность (t_n) моментов времени, обладающая тем свойством, что если $t_{n(\theta)}$ — первый член этой последовательности, больший θ , то при любом $k \geq 0$ распределение $(k+1)$ -мерной величины

$$(t_{n(\theta)} - \theta, t_{n(\theta)+1} - \theta, \dots, t_{n(\theta)+k} - \theta)$$

не зависит от θ),

GI («general independent») — рекуррентный поток, M («Markovian») — простейший поток, E_k — эрланговский поток порядка k

(множество моментов вида t_{nk} , $n \geq 1$, где t_1, t_2, \dots — моменты событий простейшего потока), D — детерминированный поток (требования поступают в моменты nc , $n \geq 1$).

Для распределения времени обслуживания употребляются следующие символы; G — распределение общего вида, E_k — распределение Эрланга порядка k , M — экспоненциальное распределение, D — распределение постоянной величины, называемое еще *вырожденным распределением*.

§ 2. Стационарный поток однородных событий

Определение стационарного потока было дано в § 1 при объяснении символа G . Этому определению эквивалентно следующее. Пусть v_1, \dots, v_n — числа событий потока в полуинтервалах времени $[a_1 + \theta; b_1 + \theta), \dots, [a_n + \theta; b_n + \theta)$. Тогда при любом n и любых a_i, b_i распределение n -мерной случайной величины (v_1, \dots, v_n) не зависит от θ . Эквивалентность обоих определений устанавливается тем, что между $(t_{n(\theta)+k})$ и (v_i) существует взаимно однозначное соответствие. Так, $v_k = j$ в том и только том случае, если

$$t_{n(\theta)+m} - \theta < a_k \leq t_{n(\theta)+m+1} - \theta, \quad t_{n(\theta)+m+j} - \theta < b_k \leq t_{n(\theta)+m+j+1} - \theta \quad \text{для некоторого } m \geq 0.$$

Обратно, $t_{n(\theta)+m} - \theta < x$ в том и только том случае, если $v_1 \geq m$ при $a_1 = 0, b_1 = x$. Следовательно, можно пользоваться обоими определениями в зависимости от того, какое окажется удобнее.

Итак, пусть имеется стационарный поток однородных событий. Обозначим $v(a, b)$ число событий этого потока в полуинтервале $[a; b)$.

Интенсивность стационарного потока однородных событий называется числом

$$\lambda = Mv(0, 1). \quad (1)$$

(Не исключается и случай $\lambda = \infty$.)

Параметром стационарного потока однородных событий называется число

$$\mu = \lim_{t \rightarrow 0} \frac{1}{t} P(v(0, t) \geq 1). \quad (2)$$

Стационарный поток однородных событий называется *ординарным*, если

$$\lim_{t \rightarrow 0} \frac{1}{t} P(v(0, t) \geq 2) = 0, \quad (3)$$

или, что то же самое,

$$P(v(0, t) \geq 2) = o(t), \quad t \rightarrow 0. \quad (4)$$

Вначале установим равенство

$$Mv(a, b) = \lambda(b - a), \quad a \leq b. \quad (5)$$

Поскольку в силу стационарности потока, очевидно, $Mv(a, b) = Mv(0, b - a)$, то достаточно установить равенство $Mv(0, t) = \lambda t$, $t \geq 0$. Имеем

$$v(0, t) = v\left(0, \frac{1}{n}\right) + v\left(\frac{1}{n}, \frac{2}{n}\right) + \dots + v\left(\frac{n-1}{n}, 1\right),$$

откуда $Mv(0, 1) = nMv\left(0, \frac{1}{n}\right)$. Итак, $Mv\left(0, \frac{1}{n}\right) = \frac{\lambda}{n}$. Взяв

теперь k/n вместо 1, получаем $Mv(0, k/n) = \lambda k/n$. Поскольку, очевидно, $Mv(0, t)$ — неубывающая функция t , то равенство $Mv(0, t) = \lambda t$ справедливо для любого $t \geq 0$, что и требовалось доказать. Данное рассуждение справедливо и при $\lambda = \infty$.

Установим важное свойство ординарного потока. Если стационарный поток однородных событий ординарен, то с вероятностью 1 $t_n \neq t_m$ для любых $n \neq m$, где t_n — моменты событий потока. Предположим, что $P\left(\bigcup_{n \neq m} \{t_n = t_m\}\right) > 0$. Тогда найдутся такие $n \neq m$, что $P(t_n = t_m) > 0$. По аксиоме непрерывности найдется такое T , что

$$P(t_n = t_m < T) = \delta > 0.$$

Имеем

$$\{t_n = t_m < T\} \subset \bigcup_{k=1}^N \left\{v\left(\frac{k-1}{N}T, \frac{k}{N}T\right) \geq 2\right\},$$

откуда

$$\delta \leq NP\left(v\left(0, \frac{T}{N}\right) \geq 2\right).$$

Это неравенство можно записать так:

$$P\left(v\left(0, \frac{T}{N}\right) \geq 2\right)/(T/N) \geq \delta/T.$$

Устремив N к бесконечности, получаем противоречие с (3), т. е. поток не ординарен.

Справедлива следующая теорема Хинчина.

Теорема. Любой стационарный поток однородных событий обладает параметром μ (не исключается и случай $\mu = \infty$).

Полное доказательство теоремы можно найти в книге [11].

Приведем доказательство теоремы Хинчина при условии, что поток ординарен. Обозначим $v_{nk} = 1$, если в полуинтервале $\left[\frac{k-1}{2^n}; \frac{k}{2^n}\right)$ произошло хотя бы одно событие потока; $s_n = v_{n1} + v_{n2} + \dots + v_{n,2^n}$. Очевидно, $(s_n, n \geq 1)$ — неубывающая последовательность. Поэтому (см. «Дополнение») по теореме Лебега существует предел

$$\mu = \lim_{n \rightarrow \infty} Ms_n = \lim_{n \rightarrow \infty} 2^n Mv_{n1} = \lim_{n \rightarrow \infty} P(v(0, 2^{-n}) \geq 1)/2^{-n}. \quad (6)$$

Пусть $t > 0$, $m \geq 0$. Возьмем r из условия $2^{-r} \leq t < 2^{-r+1}$ и затем j из условия $2^{-r-m}j \leq t < 2^{-r-m}(j+1)$. Очевидно, $j \geq 2^m$. Далее, поскольку

$$[0; t) \subset [0; 2^{-r-m}) \cup [2^{-r-m}; 2 \cdot 2^{-r-m}) \cup \dots \\ \dots \cup [j \cdot 2^{-r-m}; (j+1) \cdot 2^{-r-m}),$$

то

$$\mathbf{P}(v(0, t) \geq 1) \leq (j+1) \mathbf{P}(v_{r+m,1} = 1) = \\ = 2^{-r-m}(j+1) \mathbf{P}(v_{r+m,1} = 1)/2^{-r-m}. \quad (7)$$

При достаточно большом m выражение $2^{r+m}(j+1)$ приблизительно равно t ; при достаточно малом t величина r достаточно велика, и тогда $\mathbf{P}(v_{r+m,1} = 1)/2^{-r-m}$ приблизительно равно μ . Итак,

$$\lim_{t \rightarrow 0} \frac{1}{t} \mathbf{P}(v(0, t) \geq 1) \leq \mu. \quad (8)$$

Заметим теперь, что

$$\mathbf{P}(v(0, t) \geq 1) \geq \sum_{k=1}^j \mathbf{P}(v_{m+r,k} = 1) - \sum_{1 \leq k < l \leq j} \mathbf{P}(v_{m+r,k} = v_{m+r,l} = 1) \geq \\ \geq j \mathbf{P}(v_{m+r,1} = 1) - 2^{2m+1} \mathbf{P}(v(0, t) \geq 2).$$

Первый член этой суммы оценивается, как и в (7), последний представляет собой $o(t)$ в силу ординарности потока. Это доказывает, что вместо верхнего предела в (8) можно взять нижний. Итак, существует параметр потока.

Справедлива также теорема Королюка.

Теорема. Для стационарного ординарного потока однородных событий интенсивность совпадает с параметром:

$$\lambda = \mu. \quad (9)$$

Докажем эту теорему, полагая, что λ конечно.

Поскольку $t_n \neq t_m$ для всех $n \neq m$, то $s_n = v(0, 1)$ для достаточно большого n . При этом, как уже было отмечено, (s_n) — неубывающая последовательность неотрицательных случайных величин. Отсюда по теореме Лебега

$$\lambda = \mathbf{M}v(0, 1) = \lim_{n \rightarrow \infty} \mathbf{M}s_n = \mu,$$

что и требовалось доказать.

Резюмируем установленные результаты.

Для стационарного ординарного потока с конечной интенсивностью λ :

а) вероятность появления события потока в бесконечно малом интервале $(t; t + \Delta)$ равна $\lambda \Delta + o(\Delta)$;

б) вероятность появления двух или большего числа событий в данном интервале есть величина порядка $o(\Delta)$;

в) математическое ожидание числа событий в данном интервале равно $\lambda \Delta$.

При выводе дифференциальных уравнений и других соотношений свойства а) и б) удобно формулировать проще:

а) дифференциал вероятности появления события потока в интервале $(t; t + dt)$ равен λdt ;

б) дифференциал вероятности появления не менее двух событий в интервале $(t; t + dt)$ равен 0.

§ 3. Системы обслуживания $M|M|m|r$

Пусть имеется система $M|M|m|r$. Обозначим $v(t)$ величину очереди в момент t . Тогда возможными значениями $v(t)$ являются числа $0, 1, \dots, m+r$ при $m+r < \infty$ и все целые неотрицательные числа при $m+r = \infty$. Пусть $v(t) = k$. Тогда, как легко сообразить, траектория процесса $v(t)$, начиная с момента t , полностью определяется тремя наборами величин:

1) остаточными длительностями обслуживания тех $\min\{k, m\}$ требований, которые в момент t находятся на обслуживании;

2) длительностями обслуживания требований, которые поступят после момента t ;

3) моментами поступления этих требований.

Все эти величины в совокупности не зависят от течения процесса до момента t вследствие свойства экспоненциального распределения, независимости длительностей обслуживания и свойства простейшего потока. Итак, $v(t)$ — марковский процесс. Рассмотрим лишь *стационарный режим* системы обслуживания, иными словами, положим, что $v(t)$ — стационарный марковский процесс (цепь Маркова с непрерывным временем).

Найдем интенсивности перехода данного процесса. Очевидно, функция $v(t)$ изменяется лишь скачками величины ± 1 (в противном случае было бы возможно одновременное поступление или окончание обслуживания нескольких требований). Следовательно, можно ограничиться лишь переходами $i \rightarrow i+1$ и $i \rightarrow i-1$, считая, что $v(t) = i$. За время dt первый происходит с вероятностью λdt при $i < m+r$ и с вероятностью 0 при $i = m+r$, второй — с вероятностью $\bar{i}\mu dt$, где $\bar{i} = \min\{i, m\}$. В самом деле, при $v(t) = i$ в момент t происходит \bar{i} операций обслуживания. Каждая из них за время dt может закончиться с вероятностью μdt , а так как две или большее число операций могут закончиться с вероятностью, дифференциал которой равен 0, то вероятность окончания одной операции составляет $\bar{i}\mu dt$.

Итак,

$$\lambda_{i,i+1} = \lambda, \quad i < m+r; \quad \lambda_{i,i-1} = \bar{i}\mu. \quad (1)$$

Отсюда

$$\lambda_{ii} = -(\lambda + \bar{i}\mu), \quad i < m+r; \quad \lambda_{m+r,m+r} = -m\mu. \quad (2)$$

Следовательно, $v(t)$ — процесс размножения и гибели (§ 4.7).

Согласно формулам § 4.7,

$$\pi_k = \pi_0 \frac{\lambda_{0,1} \lambda_{1,2} \dots \lambda_{k-1,k}}{\lambda_{1,0} \lambda_{2,1} \dots \lambda_{k,k-1}}, \quad k \geq 0. \quad (3)$$

Обозначив $\rho = \lambda/\mu$, из (1) найдем

$$\pi_k = \pi_0 \rho^k / k!, \quad 0 \leq k \leq m; \quad (4)$$

$$\pi_k = \pi_0 \rho^k / (m! m^{k-m}), \quad m < k \leq m+r; \quad (5)$$

$$\pi_k = 0, \quad k > m+r. \quad (6)$$

Постоянная π_0 определяется условием нормировки $\sum_k \pi_k = 1$. При $m+r < \infty$

$$\pi_0^{-1} = \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{k=m+1}^{m+r} \left(\frac{\rho}{m} \right)^{k-m} = \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^m (1 - (\rho/m)^{r+1})}{m! (1 - \rho/m)}. \quad (7)$$

При $m = \infty$

$$\pi_0^{-1} = \sum_{k=0}^{\infty} \frac{\rho^k}{k!} = e^\rho. \quad (8)$$

При $m < \infty$, $r = \infty$

$$\pi_0^{-1} = \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^m}{m! (1 - \rho/m)}, \quad (9)$$

если только выполняется дополнительное условие

$$\rho < m. \quad (10)$$

§ 4. Основные стационарные характеристики

Обозначим \bar{v} среднюю величину очереди в системе $M|M|_m$:

$$\bar{v} = \sum_{k=0}^{\infty} k \pi_k.$$

На основании формул (3.4) и (3.5) находим

$$\pi_0^{-1} \bar{v} = \sum_{k=0}^{m-1} k \rho^k / k! + \sum_{k=m}^{\infty} k \rho^k / (m! m^{k-m}).$$

Во второй сумме, подставив $k = m + (k - m)$, получим

$$\begin{aligned} \pi_0^{-1} \bar{v} &= \rho \sum_{k=0}^{m-1} \frac{\rho^{k-1}}{(k-1)!} \frac{\rho^m}{(m-1)!} \sum_{k=m}^{\infty} \left(\frac{\rho}{m} \right)^{k-m} + \\ &+ \frac{\rho^{m+1}}{m!} \sum_{k=m}^{\infty} (k-m) \left(\frac{\rho}{m} \right)^{k-m-1}. \end{aligned} \quad (1)$$

Вторая сумма есть сумма сходящейся геометрической прогрессии; к третьей сумме применим тождество

$$\sum_{i=0}^{\infty} i z^{i-1} = \sum_{i=0}^{\infty} (z^i)' = \left(\sum_{i=0}^{\infty} z^i \right)' = \left(\frac{1}{1-z} \right)' = \frac{1}{(1-z)^2}$$

при $z = \rho/m$. Итак, при $\rho < m$ имеем

$$\pi_0^{-1} \bar{v} = \rho \sum_{k=0}^{m-2} \frac{\rho^k}{k!} + \frac{\rho^m}{(m-1)! (1-\rho/m)} + \frac{\rho^{m+1}}{m! (1-\rho/m)^2},$$

или после очевидных преобразований

$$\pi_0^{-1} \bar{v} = \rho \sum_{k=0}^{m-2} \frac{\rho^k}{k!} + \frac{m^2 \rho^m}{(m-1)! (m-\rho)^2}. \quad (2)$$

Как ведет себя \bar{v} при $\rho \uparrow m$? Из формулы (3.9) находим

$$\pi_0^{-1} \sim \frac{\rho^m}{(m-1)! (m-\rho)};$$

из формулы (2) находим

$$\pi_0^{-1} \bar{v} \sim \frac{m^2 \rho^m}{(m-1)! (m-\rho)^2}.$$

Отсюда

$$\bar{v} \sim \frac{m^2}{m-\rho}, \quad \rho \uparrow m. \quad (3)$$

Мы подходим к одному из важнейших в практическом отношении выводов теории массового обслуживания: при $\rho \uparrow m$ средняя величина очереди неограниченно возрастает.

За время T в систему поступает в среднем λT требований. Если бы система обслуживала требования без перерывов, то множество моментов времени, когда на любом из m каналов заканчивается какое-либо обслуживание, образовывало бы простейший поток с параметром μ , т. е. за время T обслуживалось бы в среднем $m\mu T$ требований. Чтобы очередь не накапливалась, очевидно, нужно, чтобы поступало в среднем не больше требований, чем система может обслужить: $\rho \leq m$. Однако, как оказывается, и при $\rho = m$ очередь с течением времени растет до бесконечности по вероятности. Интуитивно это ясно: при плотном графике любая задержка (например, в разгрузке судна в порту) сказывается на всей очереди. Соотношение $\rho \uparrow m$ называется *условием большой загрузки* системы обслуживания.

Зададимся вопросом: каково суммарное время, теряемое требованиями в системе обслуживания в интервале длины T ? (Пусть, например, в парикмахерской работают два мастера и в течение 6 ч постоянно было трое ожидающих в очереди. Тогда за это время клиентами было потеряно 30 ч времени). Если $v(t) = k$ и момент t — точка непрерывности $v(t)$, то в интервале $(t; t + dt)$ требованиями теряется $k dt$ единиц времени. Усреднив по всем возможным

k , находим, что средние потери составляют $\bar{v}dt$ в интервале dt , а стало быть, $\bar{v}T$ в интервале длины T . Поскольку за это время в систему поступает в среднем λT требований, то средние потери времени в расчете на одно требование составляют

$$\bar{v} = \bar{v}/\lambda. \quad (4)$$

Величина \bar{v} называется *средним временем пребывания* требования в системе.

Если в интервале dt значение $v(t)$ равно k , то на ожидание в очереди в этом интервале тратится $(k - m)dt$ единиц времени при $k > m$; при $k \leq m$ потери отсутствуют. Следовательно, средние потери на ожидании в очереди в единицу времени составляют

$$\sum_{k=m}^{\infty} (k - m) \pi_k = \sum_{k=m}^{\infty} k \pi_k - m \sum_{k=m}^{\infty} \pi_k = \bar{v} - \left(\sum_{k=0}^{m-1} k \pi_k + m \sum_{k=m}^{\infty} \pi_k \right). \quad (5)$$

Из формул (3.4), (3.5) находим

$$\begin{aligned} \sum_{k=0}^{m-1} k \pi_k / \pi_0 + m \sum_{k=m}^{\infty} \pi_k / \pi_0 &= \sum_{k=0}^{m-1} k \rho^k / k! + \frac{m}{m!} \sum_{k=m}^{\infty} (\rho/m)^k = \\ &= \rho \sum_{k=0}^{m-2} \rho^k / k! + \frac{\rho^m m}{(m-1)! (m-\rho)} = \rho \sum_{k=0}^{m-1} \rho^k / k! + \rho^{m+1} / (m-1)(m-\rho), \end{aligned}$$

что на основании формулы (3.9) совпадает с π_0^{-1} . Итак, вычитаемое в (5) равно ρ . Отсюда находим, что средние потери на ожидание в очереди в единицу времени составляют $\bar{v} - \lambda$. В расчете на одно требование имеем

$$\bar{w} = \frac{1}{\lambda} (\bar{v} - \rho). \quad (6)$$

Величина w называется *средним временем ожидания* требования.

Пусть теперь $m + r < \infty$. В интервале $(t; t + dt)$ происходит потеря некоторого требования, если $v(t) = m + r$ и в данном интервале в систему поступит требование. Следовательно, параметр потока потерянных требований равен $\lambda \pi_{m+r}$. По теореме Королука находим, что это же число есть интенсивность потока потерянных требований. Таким образом, за время T теряется в среднем $\lambda \pi_{m+r} T$ требований. Поскольку за это же время поступает в среднем λT требований, то на одно требование приходится в среднем π_{m+r} потерь. Эта величина называется *вероятностью потери* требования. В частности, при $r = 0$ получаем знаменитую *формулу Эрланга* вероятности потери в системе $M | M | m | 0$:

$$P_{\text{пот}} = \frac{\rho^m}{m!} / \sum_{k=0}^m \frac{\rho^k}{k!}. \quad (7)$$

§ 5. Метод вложенных цепей Маркова

Пусть $v(t)$ — величина очереди в системе обслуживания. Если входящий поток не простейший либо распределение времени обслуживания не экспоненциальное, то $v(t)$ не является марковским процессом. Тем не менее, очень часто можно выбрать такую возрастающую последовательность моментов времени (t_n) , что последовательность $v_n = v(t_n)$ является цепью Маркова. При этом моменты t_n — случайные величины, определяемые траекторией случайного процесса $v(t)$. Оказывается, что, изучая данную цепь Маркова, можно получить нужную информацию о характеристиках системы обслуживания. Последовательность (v_n) называется *вложенной цепью Маркова* процесса $v(t)$. Метод исследования систем обслуживания с помощью вложенных цепей Маркова разработан Д. Кендаллом. На примере важных систем обслуживания этот метод фактически применял еще ранее А. Я. Хинчин. Продемонстрируем метод вложенных цепей Маркова на важном примере — исследовании величин очереди в системе $M|G|1$.

Пусть $v(t)$ — величина очереди в указанной системе в момент t , причем в моменты скачков $v(t)$ определяется по непрерывности справа: $v(t) = v(t+0)$. Обозначим t_n момент окончания обслуживания n -го требования и положим $v_n = v(t_n)$. Таким образом, v_n — число требований, оставшихся в системе после окончания обслуживания n -го требования. Легко видеть, что (v_n) — цепь Маркова с состояниями $0, 1, 2, \dots$. Найдем ее вероятности перехода, обозначив через γ_n число требований, поступающих в систему в интервале $(t_{n-1}; t_n)$. При $v_{n-1} = i$ имеем $v_n = i + \gamma_n - 1$. При $i \geq 1$ в момент t_{n-1} начинается обслуживание требования, так что γ_n есть число требований, поступивших в систему в течение времени η обслуживания упомянутого требования. Следовательно,

$$P(\gamma_n = k) = \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) = f_k, \quad (1)$$

где $B(t)$ — функция распределения времени обслуживания требования, λ — интенсивность входящего потока. При $i = 0$ величина γ_n есть единица плюс число требований, поступивших в интервале длительности η , т. е.

$$P(\gamma_n = k) = \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!} dB(t) = f_{k-1}, \quad k \geq 1. \quad (2)$$

Обозначим p_{ij} вероятность перехода (v_n) из состояния i в состояние j . На основании сказанного

$$p_{ij} = \begin{cases} f_{j-i+1} & \text{при } i \geq 1, \\ f_j & \text{при } i = 0. \end{cases} \quad (3)$$

Обозначим (π_j^*) стационарное распределение цепи Маркова (v_n) .

Тогда на основании формулы (3) находим

$$\pi_j^* = \pi_0^* p_{0j} + \sum_{i=1}^{j+1} \pi_i^* p_{ij} = \pi_0^* f_j + \sum_{i=1}^{j+1} \pi_i^* f_{j-i+1}. \quad (4)$$

Обозначим через $\varphi(z)$ производящую функцию последовательности (f_j) , через $\pi^*(z)$ — производящую функцию последовательности (π_j^*) . Из (4) находим

$$\pi^*(z) = \pi_0^* \varphi(z) + (\pi^*(z) - \pi_0^*) \varphi(z)/z,$$

или

$$\pi^*(z) = \pi_0^* (1 - z) \varphi(z) / (\varphi(z) - z). \quad (5)$$

Найдем выражение для $\varphi(z)$. Имеем

$$\varphi(z) = \sum_{j=0}^{\infty} z^j \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} dB(t) = \int_0^{\infty} e^{-\lambda(1-z)t} dB(t) = \psi(\lambda(1-z)), \quad (6)$$

где $\psi(s)$ — преобразование Лапласа — Стильеса времени обслуживания:

$$\psi(s) = \int_0^{\infty} e^{-st} dB(t).$$

Пусть $\tau = M\eta < \infty$. Тогда, поскольку

$$\psi'(s) = - \int_0^{\infty} te^{-st} dB(t),$$

то $\psi'(0) = -\tau$ (имеется в виду правосторонняя производная). Отсюда

$$\begin{aligned} \psi(\lambda(1-z)) - z &= (1-z) - (1 - \psi(\lambda(1-z))) \sim \\ &\sim (1-z)(1 - \lambda\tau), \quad z \uparrow 1. \end{aligned}$$

На основании этого соотношения из (5) находим

$$1 = \pi^*(1) = \pi_0^* / (1 - \lambda\tau)$$

или

$$\pi_0^* = 1 - \lambda\tau.$$

Теперь (5) можно записать в окончательном виде

$$\pi^*(z) = \frac{(1 - \lambda\tau)(1 - z)\psi(\lambda(1 - z))}{\psi(\lambda(1 - z)) - z}. \quad (7)$$

Формула (7) называется *формулой Поллачека — Хинчина*. Методами теории цепей Маркова доказывается, что при $\lambda\tau < 1$ цепь Маркова (v_n) имеет эргодическое распределение (π_j^*) с производящей функцией вида (7). Более того,

$$\pi_j = \lim_{t \rightarrow \infty} P(v(t) = j) = \pi_j^*.$$

Таким образом, производящая функция (π_j) также имеет вид (7).

В качестве следствия из приведенных утверждений найдем среднее время \bar{w} ожидания требования. За время dt требования, находящиеся в системе обслуживания, теряют в среднем $Mv(t) dt$ единиц времени. Из этого количества $(1 - \pi_0^*) dt$ теряется на обслуживание некоторого требования, т. е. на ожидание в очереди в единицу времени остается величина $Mv(t) - 1 + \pi_0^* = Mv(t) - \lambda\tau$. При стационарном распределении $v(t)$

$$Mv(t) = \frac{d}{dz} \pi^*(z) |_{z=1}.$$

Из (7), разложив числитель и знаменатель правой части по формуле Тейлора, при $z \rightarrow 1$ получаем

$$Mv(t) = \lambda\tau + \lambda^2(\tau^2 + \sigma^2)/(2(1 - \lambda\tau)),$$

где σ^2 — дисперсия времени обслуживания требования. В силу предыдущего второй член правой части этого равенства есть величина потерь на ожидание в единицу времени. Отнеся это число к среднему числу требований, поступающих в единицу времени, находим

$$\bar{w} = \lambda(\tau^2 + \sigma^2)/(2(1 - \lambda\tau)). \quad (8)$$

Пример. К билетной кассе подходит в среднем один клиент в минуту, длительность обслуживания постоянна и равна 0,5 мин. Имеем $\lambda = 1$, $\tau = 0,5$, $\sigma = 0$, отсюда

$$\bar{w} = \frac{1 \cdot 0,5^2}{2(1 - 0,5)} = 0,25 \text{ (мин)}.$$

Допустим теперь, что у 1 % требований время обслуживания увеличилось до 10 мин. Каким будет новое значение \bar{w} ?

В данном случае

$$\begin{aligned} \lambda &= 1, \quad \tau = 0,99 \cdot 0,5 + 0,01 \cdot 10 = 0,595, \\ \tau^2 + \sigma^2 &= M\eta^2 = 0,99 \cdot 0,5^2 + 0,01 \cdot 10^2 = 1,2475. \end{aligned}$$

Отсюда

$$\bar{w} = \frac{1 \cdot 1,2475}{2(1 - 0,595)} \approx 1,54 \text{ (мин)}.$$

Видим, что \bar{w} увеличилось более, чем в 6 раз; тогда как $\lambda\tau$ возросло лишь на 19 %.

Задачи

1. В одноканальную систему обслуживания с ожиданием поступило 10 требований в моменты 7, 21, 22, 28, 39, 40, 50, 68, 73. Длительности их обслуживания равны соответственно 12, 25, 7, 1, 40, 20, 4, 4, 1, 1. Составить график функции $v(t)$ — числа требований в системе в момент t до того момента (чему он равен?), когда все требования будут обслужены. Каково суммарное время ожидания указанных десяти требований? Как определить его по графику функции $v(t)$?

2. Решить предыдущую задачу при условии, что система двухканальная. Сравнить результаты.

3. Пусть $X(t)$ — обобщенный процесс Пуассона с параметром λ ; (f_k) — распределение величины скачка процесса. Данный процесс порождает поток однородных событий, если понимать под $X(t)$ число событий в интервале $(0; t)$. Найти выражения интенсивности и параметра данного потока. Какое условие необходимо и достаточно для ординарности потока?

4. Найти значение π_0 для системы $M | M | m | r$ при $\bar{\rho} = m$, перейдя к пределу в правой части формулы (3.7).

5. Обозначим $\xi_p = (m - \rho) v$, где v — число требований в системе $M | M | m$ в стационарном режиме при $\rho < m$. Найти предельное распределение случайной величины ξ_0 при $\rho \rightarrow 1$.

6. В системе обслуживания $M | M | m | r$ эксплуатационные расходы в единицу времени составляют $c_0 m + c_1 r$ (руб.). За единицу времени ожидания требованием начала обслуживания система платит c_2 руб., за потерю требования c_3 руб. Указать способ нахождения значений m, r , при которых убыток в единицу времени минимален. Значения λ, μ считаются заданными.

7. Сравнить по вероятностям потери системы $M | M | m | 0$, $1 \leq m \leq 4$, с одним и тем же λ и значениями μ , обратно пропорциональными числу каналов.

8. При проектировании системы обслуживания конкурируют два проекта: $M | M | 4 | 0$ и $M | M | 4 | 1$. Место для ожидания предусматривать целесообразно лишь в том случае, если вследствие этого интенсивность потерь снижается хотя бы на λ_0 . В каких областях значений параметров (λ, μ) предпочтителен первый и соответственно второй проект?

9. Найти \bar{w} по формуле Хинчина в случае, если время обслуживания распределено по закону Эрланга порядка k .

10. В системе $M | M | 2 | 0$ требование вначале поступает на первый канал и только в случае его занятости — на второй. Обозначим: λ_0 — интенсивность входящего потока системы, $\lambda_1 (\lambda_0)$ — интенсивность потерь первого канала, $\lambda_2 (\lambda_1)$ — интенсивность потерь второго канала. Доказать неравенство

$$\lambda_1 (\lambda_0) < \lambda_2 (\lambda_0).$$

11. Пусть $\varphi_m(z)$ — преобразование Лапласа — Стильтеса интервала γ между последовательными моментами потери требований в системе $M | M | m | 0$. Вывести рекуррентное соотношение для $\varphi_m(z)$, основываясь на рассмотрении упорядоченной системы обслуживания, описанной в случае $m = 2$ в задаче 10. При $m = 1, 2$ проверить соотношение

$$M\gamma = 1/(\lambda\pi_m),$$

взяв выражение π_m из формулы Эрланга.

12. Каковы минимальное и максимальное значения \bar{w} при заданных λ, τ , где $\lambda\tau < 1$, если известно, что время обслуживания η с вероятностью 1 принадлежит отрезку $[0; 4\tau]$?

Глава 11

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

§ 1. Задача статистического решения

Теория вероятностей представляет собой мощный аппарат исследования явлений реального мира. Однако в рамках самой теории вероятностей решить какую-либо задачу с реальным смыслом было бы невозможно, не умей мы определить некоторые «изначальные» вероятности из эксперимента. Например, теория вероятностей указывает, как найти вероятность поражения цели при данных среднеквадратических ошибках σ_i величин, от которых зависит точность наведения ракеты. Но откуда получить значения σ_i ?

Это делается на основании эксперимента.

Математическая статистика — наука о способах получения выводов о распределениях вероятностей на основании наблюдений случайных величин.

Статистической моделью называется множество $\mathcal{F} = \{F\}$ распределений одномерной или многомерной случайной величины. *Задача статистического решения* обычно формулируется так. Известно, что случайная величина ξ имеет одно из распределений множества \mathcal{F} . На основании выборки объема n , т. е. совокупности $X = (x_1, \dots, x_n)$ независимых наблюдений ξ , требуется принять решение d из множества решений D . Независимость наблюдений x_1, \dots, x_n означает, что x_k независимые случайные величины. Распределение каждой из этих величин такое же, как и распределение величины ξ .

Пример 1. Автомат считывает цифры почтового индекса. Наблюдаемая случайная величина $\xi = (\xi_1, \dots, \xi_9)$, где ξ_i — «степень затемненности» i -го прямолинейного элемента изображения. При истинном значении цифр 0, 1, ..., 9 распределение ξ будет соответственно F_0, F_1, \dots, F_9 . В результате наблюдения x величины ξ нужно принять одно из решений: 0, 1, ..., 9, т. е. прочитать цифру как 0, 1, ..., 9.

Пример 2. Производятся испытания элементов на долговечность. Известно, что время безотказной работы ξ имеет экспоненциальное распределение: для положительных значений аргумента $F_\xi(t) = 1 - e^{-\lambda t}$, однако параметр λ не известен. В результате испытаний наблюдаются x_1, \dots, x_n — длительности безотказной работы n элементов, подвергшихся испытанию. Множество \mathcal{F} в данном случае есть семейство распределений $\{1 - e^{-\lambda t}, \lambda > 0\}$. Решение d — любое неотрицательное число; таким образом, $D = \mathbb{R}^+$. Если принято решение d , то это означает, что истинное значение параметра λ оценено числом d .

Пример 3. Производится пристрелка наземной цели по дальности. После каждого выстрела получается наблюдение «+» (перелет) или «—» (недолет). Обозначим θ разность между средней дальностью падения снаряда на данном делении прицела и расстоянием до цели. Распределение результата выстрела имеет следующий вид:

$$P(+)=P_{\theta}(+)=\frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\theta/\sigma}^{\infty}e^{-x^2/2}dx,$$

$$P(-)=P_{\theta}(-)=1-P(+).$$

Таким образом, \mathcal{F} есть совокупность распределений случайной величины со значениями «+» и «—»; каждое из этих распределений характеризуется определенным значением θ . В результате пристрелки¹ получаем некоторое число знаков + и —, на основании которых указывается так называемый район возможного положения цели ($R_1; R_2$). Смысл решения $d = (R_1; R_2)$ в следующем: полагают, что цель находится в интервале ($R_1; R_2$) по дальности.

Статистическая модель называется *параметрической*, если каждое входящее в нее распределение однозначно определяется значением одномерного или многомерного параметра θ , т. е.

$$\mathcal{F} = \{F(x, \theta), \theta \in \Theta\},$$

где Θ — подмножество \mathbb{R}^m . В примерах 2 и 3 статистическая модель — параметрическая.

¹ Для упрощения формулы нами принято, что в процессе пристрелки деление прицела не меняется.

На примерах 1—3 показаны три основные задачи математической статистики. В примере 1 имеем *задачу статистического выбора*, когда имеется некоторое множество гипотез (в данном случае 0, 1, ..., 9) о распределении случайной величины и по выборке надлежит выбрать (принять) одну из них. В примере 2 имеем *задачу оценки параметра* (более подробно: *точечной* оценки параметра), когда по выборке $X = (x_1, \dots, x_n)$ требуется построить оценку параметра λ , т. е. такую функцию выборки $\hat{\theta}(X)$, которая может служить приближением этого параметра. Наконец, пример 3 иллюстрирует *задачу интервальной оценки* — такого интервала $(R_1(X); R_2(X))$, который с некоторой вероятностью содержит значение неизвестного параметра.

Любую числовую или векторную функцию выборки $\eta = f(X)$ назовем *статистикой*, если при любом $F \in \mathcal{F}$ она является случайной величиной. Функцию распределения, плотность вероятности, математическое ожидание данной статистики при данном распределении $F \in \mathcal{F}$ наблюдаемой случайной величины ξ будем обозначать $F_{\eta, F}(x)$, $p_{\eta, F}(x)$, $M_F \eta$. В случае параметрической статистической модели соответственно записываем $F_{\eta, \theta}(x)$, $p_{\eta, \theta}(x)$, $M_{\theta} \eta$. (Если оговорено, о какой величине η идет речь, символ η в нижнем индексе не указывается.)

Математическая статистика дает рекомендации о выборе статистических решений, оптимальных в том или ином смысле. Чтобы уточнить смысл оптимальности, нужно, наряду со статистической моделью, задать некоторую количественную меру предпочтения того или иного решения. В последующих параграфах такие меры будут конкретизированы.

Сделаем замечание об отношении выборки (x_k) к реально наблюдаемым явлениям. Обычно выводы математической статистики применяются к выборочному исследованию больших популяций, называемых генеральными совокупностями. *Генеральная совокупность* — это множество N объектов, каждый из которых обладает некоторым признаком x . Чтобы сделать вывод о некоторой характеристике генеральной совокупности (например, о среднем значении переменной x по всем N объектам), производят выборку объема n и измеряют значения x_1, \dots, x_n признака объектов, попавших в выборку.

Выборка образуется в соответствии с некоторым механизмом. Интуитивно ясно, что если попадание объекта в выборку зависит от значения признака x , то статистика будет непредставительной, т. е. выборка не будет представлять среднюю характеристику генеральной совокупности. Такой механизм характерен для всякого рода фальсификаций, «жонглирования цифрами». В противоположность этому при честном подходе к решению прикладной задачи стремятся как можно более уравнивать вероятности попадания в выборку любых объектов генеральной совокупности. В качестве идеализированных математических моделей беспристрастного образования выборки приняты две схемы выбора: 1) выбор без возвращения; 2) выбор с возвращением. В первом случае вероятность попада-

ния в выборку объема n объектов i_1, i_2, \dots, i_n генеральной совокупности, где $1 \leq i_1 < i_2 < \dots < i_n \leq N$, равна $1/C_N^n$. Во втором случае первый объект выбирается из генеральной совокупности по равновероятному закону и после измерения значения параметра вновь возвращается в генеральную совокупность. Второй объект снова выбирается равновероятно из всей совокупности и т. д. Таким образом, в выборку могут по несколько раз попадать одни и те же объекты.

Пусть ω — случайная величина, принимающая значения $1, 2, \dots, N$ с вероятностью $1/N$, $\xi = \xi(\omega)$ — значение признака x для объекта с номером ω . Тогда, как легко видеть, при выборе с возвращением x_1, \dots, x_n — независимые случайные величины, распределенные так же, как и величина ξ . При выборе с возвращением вероятность попадания в выборку хотя бы двух одинаковых объектов

$$q \leq C_n^2 \frac{1}{N} < \frac{n^2}{2N}$$

(обосновать эту оценку предлагается читателю), а стало быть, при достаточно большом объеме генеральной совокупности выбор с возвращением и выбор без возвращения — практически одно и то же. В данной главе будет рассматриваться лишь выбор с возвращением, что согласуется с определенным в начале этого параграфа понятием статистической модели.

§ 2. Выбор между двумя гипотезами

Пусть $\mathcal{F} = \{F_0, F_1\}$. Это означает, что наблюдаемая случайная величина ξ имеет либо распределение $F_0(x)$ (гипотеза H_0), либо распределение $F_1(x)$ (гипотеза H_1). Статистическое решение о выборе между гипотезами H_0 и H_1 может быть нерандомизированным или рандомизированным. *Нерандомизированное решение* определяется разбиением множества значений X на два подмножества: A_0 и A_1 . Если X попадает в множество A_i , принимается гипотеза H_i . Множество A_1 называется *критической областью* гипотезы H_0 (соответственно A_0 — критическая область H_1). Можно сказать так: гипотеза отклоняется, если выборка попала в критическую область этой гипотезы.

Рандомизированное решение определяется статистикой $\varphi(X)$, значения которой принадлежат отрезку $[0; 1]$ при любом значении X , и формулируется так: при данном значении X гипотеза H_1 принимается с вероятностью $\varphi(X)$, гипотеза H_0 — с вероятностью $1 - \varphi(X)$. (Таким образом, для принятия решения нужно «бросить жребий».) В частности, положив $\varphi(X) = 0$ при $X \in A_0$, $\varphi(X) = 1$ при $X \in A_1$, получим нерандомизированное решение, соответствующее данным A_0 и A_1 . Таким образом, класс нерандомизированных решений входит в класс рандомизированных решений. Решение принять гипотезу H_i отождествим со значением $d = i$ ($i = 0, 1$). *Ошибкой первого рода* назовем принятие решения $d = 1$ при

условии, что справедлива гипотеза H_0 . Ошибкой второго рода назовем принятие решения $d = 0$ при справедливости H_1 . Вероятности ошибок первого и второго рода имеют вид

$$\alpha = M_0 \varphi(X), \quad \beta = M_1(1 - \varphi(X)) = 1 - M_1 \varphi(X). \quad (1)$$

Естественно, желательно сделать и α , и β как можно меньше. Рассмотрим следующие задачи.

Задача I. Заданы положительные числа c_0, c_1 . Требуется выбрать рандомизированное решение, минимизирующее взвешенную сумму вероятностей ошибок первого и второго рода:

$$c_0 \alpha + c_1 \beta \Rightarrow \min. \quad (2)$$

При экономическом подходе к выбору между гипотезами c_0 и c_1 интерпретируются как убытки от ошибок первого и второго рода. В практических задачах эти убытки часто сильно различаются. Например, при специальном исследовании деталей машин наличие скрытых дефектов несоизмеримо опаснее не заметить имеющуюся трещину, чем ошибочно заключить о наличии таковой.

Задача II. Задано число α_0 , $0 < \alpha_0 < 1$. Требуется выбрать рандомизированное решение из условия

$$\beta \Rightarrow \min, \quad \alpha \leq \alpha_0. \quad (3)$$

По формуле (1) находим

$$c_0 \alpha + c_1 \beta = c_1 + c_0 M_0 \varphi(X) - c_1 M_1 \varphi(X). \quad (4)$$

Допустим, что распределения F_0 и F_1 имеют плотности $p_0(x)$ и $p_1(x)$ относительно некоторой меры, т. е.

$$\int_C dF_i(x) = \int_C p_i(x) d\Phi(x), \quad i = 0, 1. \quad (5)$$

В частности, если F_i — непрерывные распределения, то в качестве $p_i(x)$ можно взять соответствующие им плотности вероятностей; в этом случае $d\Phi(x) = dx$. Если F_i — дискретные распределения, сосредоточенные на одних и тех же последовательностях (x_k) , можно положить $\Phi(A) = 1$ для одноточечных множеств $A = \{x_k\}$, $\Phi(A) = 0$ для A , не содержащих ни одной точки x_k . Тогда если $p_i(x) = P(\xi = x)$ при распределении F_i , то выполняется равенство (5).

Принятая символика как бы уравнивает непрерывное и дискретное распределения, чем избавляет от повторения выкладок отдельно для одного и другого случаев.

Из (4) находим

$$c_0 \alpha + c_1 \beta = c_1 + \int [c_0 p_0(X) - c_1 p_1(X)] \varphi(X) d\Phi(X), \quad (6)$$

где для сокращения обозначено $\int = \int \dots \int$, $p_i(X) = p_i(x_1) \dots$

¹ Эти названия привязаны к определенной нумерации гипотез H_0, H_1 : гипотезу H_0 называют «основной», H_1 — «альтернативной». В математическом отношении картина вполне симметрична.

... $p_i(x_n)$, $d\Phi(X) = d\Phi(x_1) \dots d\Phi(x_n)$. Очевидно, минимум выражения (6) получим, положив $\varphi(X) = 0$ там, где выражение в квадратных скобках положительно, и $\varphi(X) = 1$ там, где оно отрицательно. В точках, где [...] = 0, $\varphi(X)$ можно задать произвольным образом. Отношение

$$L(X) = p_1(X)/p_0(X) = \prod_{k=1}^n [p_1(x_k)/p_0(x_k)] \quad (7)$$

назовем *отношением правдоподобия* (гипотезы H_1 против гипотезы H_0). Знак выражения в квадратных скобках равенства (6) противоположен знаку выражения $L(X) - c_0/c_1$. Отсюда находим

$$\varphi(X) = \begin{cases} 1 & \text{при } L(X) > c_0/c_1, \\ 0 & \text{при } L(X) < c_0/c_1. \end{cases} \quad (8)$$

В частности, если $P(c_0 p_0(X) - c_1 p_1(X) = 0) = 0$, получаем следствие, что решением задачи оптимизации I является нерандомизированное решение, для которого

$$A_1 = \{X : L(X) \geq c_0/c_1\}. \quad (9)$$

Полезно дать вероятностную интерпретацию полученного решения. Допустим, что статистическое решение принимается многократно, всякий раз по отношению к новой выборке X . Скажем, производится N реализацией статистического решения. Предположим также, что известны вероятности u_0, u_1 того, что в данной

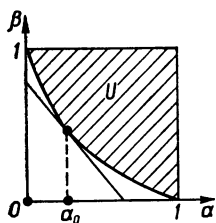


Рис. 27

реализации справедлива соответственно гипотеза H_0 или H_1 . Эти вероятности называются *априорными вероятностями* гипотез. Пусть, далее, ошибка первого рода приносит убыток w_0 , ошибка второго рода — убыток w_1 . Тогда в расчете на одну реализацию с данным статистическим решением связан средний убыток

$$\bar{w} = u_0 w_0 \alpha + u_1 w_1 \beta. \quad (10)$$

Легко видеть, что решение, минимизирующее \bar{w} , есть решение задачи I при $c_0 = u_0 w_0$, $c_1 = u_1 w_1$. Это решение называется *байесовским статистическим решением*. Применяя такое решение в неограниченно возрастающем числе реализаций, получим минимальное значение среднего убытка в расчете на одну реализацию.

Перейдем к решению задачи II. Любому решению $\varphi(X)$ соответствуют некоторые α, β . Обозначим через U множество точек (α, β) единичного квадрата, каждая из которых соответствует некоторому $\varphi(X)$. Множество U выпукло. В самом деле, как легко видеть, если φ_1, φ_2 соответствуют точки $(\alpha_1, \beta_1), (\alpha_2, \beta_2)$, то при любом x , $0 < x < 1$, $x\varphi_1 + (1-x)\varphi_2$ также будет решением и ему будет соответствовать точка $(x\alpha_1 + (1-x)\alpha_2, x\beta_1 + (1-x)\beta_2)$; это и есть свойство выпуклости множества. Теперь остается использовать свойство выпуклого множества: решение задачи II сводится к решению задачи I при надлежащим образом выбранных c_0, c_1 (рис. 27).

Итак, доказано следующее утверждение, называемое *леммой Неймана — Пирсона*.

Лемма. Критерий выбора между двумя гипотезами (задача II) характеризуется функцией

$$\varphi(X) = \begin{cases} 1, & \text{если } L(X) > z, \\ 0, & \text{если } L(X) < z \end{cases} \quad (11)$$

при некотором z .

Указанный критерий называется *критерием Неймана — Пирсона (критерием отношения правдоподобия)*. В частности, если функция распределения статистики $L(X)$ непрерывна при обеих гипотезах H_0, H_1 , то задача II имеет нерандомизированное решение

$$\begin{aligned} A_0 &= \{x : L(x) < z\}, \\ A_1 &= \{x : L(x) \geq z\}, \end{aligned} \quad (12)$$

где постоянная z определяется (любым) решением уравнения

$$P_0(L(X) > z) = \alpha, \quad (13)$$

или, что то же самое,

$$\int_{L(X) > z} dF_0(X) = \alpha. \quad (14)$$

Поскольку логарифм — монотонная функция, то критерий Неймана — Пирсона может быть также выражен формулой

$$\varphi(X) = \begin{cases} 1 & \text{при } l(X) > \ln z, \\ 0 & \text{при } l(X) < \ln z, \end{cases} \quad (15)$$

где $l(X)$ — логарифм отношения правдоподобия:

$$l(X) = \sum_{k=1}^n \ln(p_1(x_k)/p_0(x_k)). \quad (16)$$

Пример 1. Пусть F_l — нормальное распределение с математическим ожиданием a_l и дисперсией σ^2 . Для определенности положим $a_0 < a_1$.
Имеем

$$\ln(p_1(x)/p_0(x)) = \frac{1}{2\sigma^2} ((x - a_0)^2 - (x - a_1)^2) = \frac{a_1 - a_0}{\sigma^2} x + \frac{a_0^2 - a_1^2}{2\sigma^2}.$$

Обозначив $\bar{x} = (x_1 + \dots + x_n)/n$, получим

$$\ln(X) = n \left(\frac{a_1 - a_0}{\sigma^2} \bar{x} + \frac{a_0^2 - a_1^2}{2\sigma^2} \right).$$

Таким образом, $l(X)$ есть линейная функция \bar{x} . Легко понять, что решение задачи II определяется условием

$$A_1 = \{X : \bar{x} \geq \gamma\}.$$

Постоянная γ , вследствие (13), определяется условием

$$P_0(\bar{x} > \gamma) = \alpha_0. \quad (17)$$

Поскольку $\sqrt{n}(\bar{x} - a_0)/\sigma$ имеет стандартное нормальное распределение, то вместо (17) имеем уравнение

$$\frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}(\bar{y}-a_0)/\sigma}^{\infty} e^{-x^2/2} dx = \alpha.$$

При данном решении $\alpha = \alpha_0$. Для β имеем выражение

$$\beta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{n}(\bar{y}-a_1)/\sigma} e^{-x^2/2} dx.$$

Отметим, что при $c_0 = c_1$ решение задачи I выглядит особенно просто:

$$A_1 = \left\{ X : \bar{x} \geq \frac{a_0 + a_1}{2} \right\}.$$

Пример 2. При гипотезах H_0 , H_1 распределение наблюдаемой случайной величины ξ экспоненциально со значениями параметров λ_0 и λ_1 . Положим для определенности $\lambda_0 > \lambda_1$. Имеем

$$l(X) = (\lambda_0 - \lambda_1)(x_1 + \dots + x_n) + n \ln(\lambda_1/\lambda_0).$$

Отсюда очевидно, что решение задачи II имеет вид $A_1 = \{x_1 + \dots + x_n > \gamma\}$. Поскольку при гипотезе H_0 сумма n независимых экспоненциальных случайных величин с параметром λ имеет распределение Эрланга, то постоянную γ находим из уравнения

$$e^{-\lambda_0 \gamma} \sum_{k=0}^{n-1} (\lambda_0 \gamma)^k / k! = \alpha_0.$$

Как и в примере 1, $\alpha = \alpha_0$. Для β имеем выражение

$$\beta = 1 - e^{-\lambda_1 \gamma} \sum_{k=0}^{n-1} (\lambda_1 \gamma)^k / k!.$$

Пример 3. Наблюдается пуассоновская случайная величина ξ . Относительно значения параметра закона Пуассона имеются две гипотезы: $H_0 = \{a = a_0\}$ и $H_1 = \{a = a_1\}$. Для определенности полагаем $a_0 < a_1$.

Определим меру Φ следующим образом. Мера любого одноточечного множества $\{n\}$ при целых неотрицательных n равна 1, мера любого множества, не содержащего целых неотрицательных точек, равна 0. Тогда F_i имеют плотности $p_i(x)$ относительно меры Φ :

$$p_i(x) = e^{-a_i} a_i^x / x!, \quad i = 0, 1.$$

Отсюда

$$\ln(p_1(x)/p_0(x)) = x \ln(a_1/a_0) + a_0 - a_1,$$

откуда

$$l(X) = (x_1 + \dots + x_n) \ln(a_1/a_0) + n(a_1 - a_0).$$

В отличие от предыдущих двух примеров распределение $l(X)$ дискретно. Вначале предположим, что при некотором целом m

$$q_0(m) + q_0(m+1) + \dots = \alpha_0,$$

где

$$q_i(k) = e^{-na_i} (na_i)^k / k!.$$

(Заметим, что сумма $x_1 + \dots + x_n$ имеет распределение $(q_i(k))$ при гипотезе H_i .) Тогда решение задачи II имеет вид

$$A_1 = \{X : x_1 + \dots + x_n \geq m\}.$$

Если при некотором целом m

$$q_0(m) + q_0(m+1) + \dots < \alpha_0 < q_0(m-1) + q_0(m) + \dots,$$

то решение задачи II имеет следующий вид. При $x_1 + \dots + x_n \geq m$ принимается гипотеза H_1 , при $x_1 + \dots + x_n \leq m-2$ — гипотеза H_0 . Если же $x_1 + \dots + x_n = m-1$, «бросаем жребий»: гипотеза H_1 принимается с вероятностью t , определяемой из уравнения

$$q_0(m-1)t + q_0(m) + q_0(m+1) + \dots = \alpha_0.$$

(Впрочем, в некоторых случаях можно избежать рандомизации, привлекая значения x_1, \dots, x_n в отдельности, а не только их сумму.)

§ 3. Проверка статистической гипотезы

Пусть имеется статистическая модель \mathcal{F} . *Статистической гипотезой* называется любое непротиворечивое утверждение относительно $F \in \mathcal{F}$. С гипотезой H можно отождествить подмножество \mathcal{F}_H множества \mathcal{F} : это множество распределений, совместных с гипотезой H . Гипотеза H называется *простой*, если \mathcal{F}_H состоит из единственного элемента, в противном случае — *сложной*. В § 2 рассматривались только простые гипотезы H_0, H_1 .

Итак, пусть H — гипотеза, которую мы называем *основной*. Гипотезу K , отождествленную со всеми остальными F , назовем *альтернативной*. Таким образом, $\mathcal{F}_K = \mathcal{F} \setminus \mathcal{F}_H$. В случае принятия гипотезы H или K будем писать соответственно $d = 0$ и $d = 1$.

Критерием (тестом) гипотезы H называется статистическое решение вида

$$d = d(X) = \begin{cases} 0 & \text{при } X \in A, \\ 1 & \text{при } X \in B, \end{cases}$$

где A и B — множества, составляющие разбиение множества возможных значений X , либо вида $\varphi(X)$. Смысл $\varphi(X)$ — тот же, что и в § 2. В первом случае критерий гипотезы назовем *нерандомизированным*, во втором — *рандомизированным*. Пусть

$$\mathbf{P}_F(d = 1) \leq \alpha, \quad \alpha \in \mathcal{F}_H. \quad (1)$$

Тогда α называется *уровнем значимости* данного критерия. Функция $\mathbf{P}_F(d = 1)$, определенная при $F \in \mathcal{F}_K$, называется *функцией мощности* данного критерия. При выборе критерия гипотезы H обычно отбрасывают любой критерий уровня значимости α , для которого существует более мощный критерий, т. е. такой критерий уровня значимости α , для которого $\mathbf{P}_F(d = 1)$ на множестве \mathcal{F}_K всюду не меньше и хотя бы в одной точке больше, чем для первого критерия. Критерий, для которого не существует более мощного, называется *допустимым*. Если данный критерий уровня значимости α обладает тем свойством, что для любого другого критерия с этим уровнем значимости значение $\mathbf{P}_F(d = 1)$ не больше, чем для данного критерия, при любых $F \in \mathcal{F}_K$, то данный критерий называется *равномерно наиболее мощным*. Для такого критерия функция мощности представляет собой верхнюю огибающую семейства функций мощности всех критериев данного уровня значимости.

Укажем простой случай, когда равномерно наиболее мощный критерий существует. Предположим, что $\mathcal{F} = \{F_\theta\}$, где F_θ — распределения одномерных случайных величин, зависящие от одномерного параметра θ , причем выполняется соотношение

$$p_\theta(x) = \exp\{a(x) + b(x)u(\theta) + c(\theta)\}, \quad (2)$$

где $p_\theta(x)$ — плотность вероятности данного распределения¹. Допустим также, что в формуле (2) функции $b(x)$ и $u(\theta)$ — строго монотонные, $H = \{F_\theta, \theta \leq \theta_0\}$, $K = \{F_\theta, \theta > \theta_0\}$. Докажем, что это действительно так.

Пусть для определенности $b(x)$ и $u(\theta)$ — возрастающие функции. Построим нерандомизированный критерий $d_t(X)$ как решение задачи II § 2 при простых гипотезах F_θ и F_t ($t > \theta_0$). Имеем

$$\ln(p_1(x)/p_0(x)) = b(x)(u(t) - u(\theta_0)) + c(t) - c(\theta_0).$$

Отсюда

$$l(X) = (u(t) - u(\theta_0)) \sum_{k=1}^n b(x_k) + n(c(t) - c(\theta_0)). \quad (3)$$

Легко видеть, что критерий $d_t(X)$ определяется условием

$$d_t(X) = \begin{cases} 1 & \text{при } \sum_{k=1}^n b(x_k) \geq z, \\ 0 & \text{в противном случае.} \end{cases}$$

Поскольку z определяется из условия

$$P_{\theta_0}\left(\sum_{k=1}^n b(x_k) \geq z\right) = \alpha_0,$$

то z не зависит от t ; следовательно, все d_t совпадают. Поскольку $d_t(X)$ есть решение задачи II § 2, то

$$P_t(d(X) = 1) \leq P_t(d_t(X) = 1)$$

для любого критерия $d(X)$ уровня значимости α . Это и означает, что $d_t(X)$ — равномерно наиболее мощный критерий.

В качестве следствия получаем, что критерий отношения правдоподобия, полученный в примере 1 § 2, является равномерно наиболее мощным критерием уровня α_0 проверки гипотезы $\{a \leq a_0\}$ при альтернативе $\{a > a_0\}$.

§ 4. Непараметрические критерии проверки гипотез

Часто возникает задача проверки простой гипотезы H , состоящей в том, что наблюдаемая случайная величина ξ имеет функцию распределения $F(x)$. Альтернативные гипотезы при этом явно не указываются. Обычно данная задача решается следующим образом.

¹ Множество распределений с данным свойством называется *экспоненциальным семейством распределений*.

Выбирают некоторую одномерную статистику $\varphi(X)$, распределение которой при гипотезе H известно:

$$P_F(\varphi(X) < x) = G(x).$$

Назначают уровень значимости α и по нему определяют квантиль z этого распределения:

$$G(z) = 1 - \alpha.$$

Это означает, что

$$P_F(\varphi(X) \geq z) = \alpha.$$

Критерий проверки гипотезы H , называемый также *критерием согласия*, формулируется следующим образом: *если $\varphi(X) < z$, гипотеза принимается; если $\varphi(X) \geq z$, гипотеза отклоняется*. Данный критерий обладает уровнем значимости α . В случае, если гипотеза принята, это не означает, что она в самом деле верна, а означает лишь то, что она не находится в грубом противоречии с результатами эксперимента. Распределение F , связанное с гипотезой H , называется *теоретическим распределением* случайной величины ξ . Статистика $\varphi(X)$ называется *статистикой критерия*. В математической статистике большую роль играют *непараметрические критерии*, характеризующиеся тем, что статистика критерия при гипотезе H имеет распределение, не зависящее от F . (Имеется в виду, что вид самой функции $\varphi(X)$ зависит от F .) Обычно непараметрический критерий применяется в следующем приближенном варианте.

Статистика $\varphi(X)$ зависит от объема выборки n : $\varphi(X) = \varphi_n(X)$. Соответственно $G(x) = G_n(x)$. Допустим, что $\varphi_n(X)$ при гипотезе H имеет непрерывное предельное распределение, т. е.

$$G_n(x) \xrightarrow{n \rightarrow \infty} G_0(x). \quad (1)$$

Тогда критерий $\varphi_n(X) < z$ принятия гипотезы H при объеме выборки n , где $G_0(z) = 1 - \alpha$, будет иметь уровень значимости, приблизительно равный α при большом n . В практических случаях ошибкой в значении уровня значимости пренебрегают, т. е. допредельное распределение $\varphi(X)$ заменяют предельным. Так, описанный ниже непараметрический критерий Колмогорова рекомендуется применять при объеме выборки $n \geq 30$.

Пусть ξ — одномерная случайная величина. По выборке X объема n построим *эмпирическую функцию распределения* $F_n(x)$, определенную при любом $x \in \mathbb{R}$ как частота наблюдений со значениями, меньшими x . Формулой можно записать так:

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n E(x - x_k),$$

где $E(x) = 0$ при $x \leq 0$, $E(x) = 1$ при $x > 0$.

Эмпирическая функция распределения — неубывающая функция, непрерывная слева, возрастающая лишь скачками, кратными $1/n$, равная нулю при x , меньших минимального наблюдения, и равная единице при x , больших максимального наблюдения. Если

x совпадает с k наблюдениями из числа x_1, \dots, x_n , то $F_n(x+0) - F_n(x) = k/n$. Типичный график эмпирической функции распределения показан на рис. 28.

Пусть $F(x) = P_H(\xi < x)$ — непрерывная функция. Обозначим через Δ_n максимальное абсолютное отклонение эмпирической функции распределения от теоретической:

$$\Delta_n = \sup_x |F_n(x) - F(x)|. \quad (2)$$

Чтобы вычислить эту величину, не нужно перебирать все значения. Действительно, как легко видеть,

$$\begin{aligned} \Delta_n = \max \{ & |F_n(x_1) - F(x_1)|, |F_n(x_1+0) - F(x_1+0)|, \\ & |F_n(x_2) - F(x_2)|, |F_n(x_2+0) - F(x_2+0)|, \dots \\ & \dots, |F_n(x_n) - F(x_n)|, |F_n(x_n+0) - F(x_n+0)| \}, \end{aligned} \quad (3)$$

т. е. нужно перебрать лишь $2n$ чисел. Справедлива следующая предельная теорема Колмогорова.

Теорема. При любом $z \in \mathbb{R}$

$$P(\sqrt{n} \Delta_n < z) \xrightarrow{n \rightarrow \infty} K(z), \quad (4)$$

где $K(z)$ — функция распределения, называемая функцией Колмогорова и задаваемая формулой

$$K(z) = \begin{cases} \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 z^2} & \text{при } z > 0, \\ 0 & \text{при } z \leq 0. \end{cases}$$

Критерий Колмогорова формулируется так: по заданному α , $0 < \alpha < 1$ выбирают z из условия $K(z) = 1 - \alpha$.

Гипотезу H принимают при $\Delta_n < z/\sqrt{n}$ и отклоняют при $\Delta_n \geq z/\sqrt{n}$. В силу (4) критерий Колмогорова, примененный при большом объеме выборки, имеет уровень значимости, близкий к α .

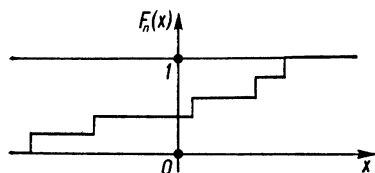


Рис. 28

Значения $K(z)$ табулированы. Функция $K(z)$, как и обратная к ней функция, не выражается в замкнутом виде через элементарные функции. Впрочем, при малых α можно вывести простую прибли-

женную формулу. Именно, при $\alpha \rightarrow 0$, очевидно, $z \rightarrow \infty$. Из (5) находим, что при $z \rightarrow \infty$

$$1 - K(z) \sim 2e^{-2z^2}$$

(слагаемые с $k = \pm 1$ «подавляют» все слагаемые с $|k| \geq 2$). Отсюда получаем, что при малых α $e^{-2z^2} \approx \alpha/2$, или

$$z \approx \left(\frac{1}{2} \ln \frac{2}{\alpha} \right)^{1/2}. \quad (6)$$

Доказательство теоремы Колмогорова основано на теории диффузионных процессов (см. [8]). Существует также комбинаторное доказательство теоремы Колмогорова (см. Д. Дюге. Теоретическая и прикладная статистика.— М., 1972).

Рассматривая эмпирическую функцию распределения $F_n(x)$ одномерной случайной величины ξ , необходимо отметить ее важнейшее свойство, имеющее принципиальное значение и выражаемое теоремой Гливленко — Кантелли.

Теорема. Пусть x_1, x_2, \dots — независимые наблюдения одномерной случайной величины ξ , $F_n(x)$ — эмпирическая функция распределения, составленная по наблюдениям x_1, \dots, x_n . Тогда с вероятностью 1

$$\Delta_n = \sup_x |F_n(x) - F(x)| \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство теоремы Гливленко — Кантелли приведем лишь для случая, когда теоретическая функция распределения $F(x)$ непрерывна. В этом случае по данному $\varepsilon > 0$ найдем такие точки $-\infty = t_0 < t_1 < \dots < t_m < \infty = t_{m+1}$, где m зависит от ε , что $F(t_{k+1}) - F(t_k) < \varepsilon/2$ для всех $k = 1, 2, \dots, m+1$. Вследствие усиленного закона больших чисел $F_n(t_k) \xrightarrow{n \rightarrow \infty} F(t_k)$. Следовательно, с вероятностью 1 найдется такое случайное n_ε , что $|F_n(t_k) - F(t_k)| < \varepsilon/2$ при всех $k = 1, 2, \dots, m$ и всех $n \geq n_\varepsilon$. Поскольку $F_n(x)$ и $F(x)$ монотонно не убывают, то для всех x из отрезка $[t_k; t_{k+1}]$ при $n \geq n_\varepsilon$

$$\begin{aligned} F_n(x) - F(x) &\leq F_n(t_{k+1}) - F(t_k) = (F_n(t_{k+1}) - F(t_{k+1})) + \\ &+ (F(t_{k+1}) - F(t_k)) < \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Аналогично при $n \geq n_\varepsilon$

$$\begin{aligned} F(x) - F_n(x) &\leq F(t_{k+1}) - F_n(t_k) = (F(t_{k+1}) - F(t_k)) + \\ &+ (F(t_k) - F_n(t_k)) < \varepsilon. \end{aligned}$$

Итак с вероятностью 1 при $n \geq n_\varepsilon$ выполняется неравенство

$$|F_n(x) - F(x)| < \varepsilon$$

для всех x одновременно, т. е. $\Delta_n < \varepsilon$. Поскольку $\varepsilon > 0$ произвольно, то это означает, что $\Delta_n \xrightarrow{n \rightarrow \infty} 0$ с вероятностью 1.

Приведем еще непараметрический критерий согласия *хи-квадрат* (χ^2) Пирсона. Пусть $F(x)$ — теоретическое распределение. Разобьем пространство значений случайной величины на r подмножеств A_1, \dots, A_r (например, если наблюдаемая случайная величина — одномерная, то обычно $A_1 = (-\infty; a_1)$, $A_2 = [a_1; a_2)$, ..., $A_{r-1} = [a_{r-2}; a_{r-1})$, $A_r = [a_{r-1}; \infty)$). Пусть $p_k = \int A_k dF(x)$ — теоретическая вероятность попадания наблюдения в множество A_k при основной гипотезе H . Статистикой *хи-квадрат* назовем функцию выборки X вида

$$\chi^2 = \sum_{k=1}^r (s_{nk} - np_k)^2 / (np_k), \quad (7)$$

где s_{nk} — число наблюдений выборки объема n со значениями из множества A_k .

В примере 3 § 6.1 показано, что $(r - 1)$ -мерная случайная величина

$$\eta_n = ((s_{n1} - np_1)/\sqrt{np_1}, \dots, (s_{nr-1} - np_{r-1})/\sqrt{np_{r-1}})$$

асимптотически нормальна с нулевым математическим ожиданием и корреляционной матрицей $K = (k_{ij})$, где $k_{ii} = 1 - p_i$, $k_{ij} = -\sqrt{p_i p_j}$ при $i \neq j$. Отсюда следует, что

$$\begin{aligned} P(\chi^2 < z) &= \int \dots \int_{x_1^2 + \dots + x_r^2 < z} d\Phi_n(x_1, \dots, x_{r-1}) \xrightarrow{n \rightarrow \infty} \\ &\xrightarrow{n \rightarrow \infty} \int \dots \int_{x_1^2 + \dots + x_r^2 < z} d\Phi(x_1, \dots, x_{r-1}), \end{aligned} \quad (8)$$

где Φ_n — функция распределения случайной величины η_n , Φ — функция распределения величины $N(0, K)$, которую мы обозначим $\zeta = (\zeta_1, \dots, \zeta_{r-1})$, переменная x_r в обоих интегралах формулы (8)

определяется из уравнения $\sum_{k=1}^r x_k \sqrt{p_k} = 0$. Соотношение (8) можно переформулировать так:

$$P(\chi^2 < z) \xrightarrow{n \rightarrow \infty} P\left(\sum_{k=1}^r \zeta_k^2 < z\right).$$

В дальнейшем будем понимать под ζ не $(r - 1)$ -мерную, а r -мерную нормальную величину с нулевым математическим ожиданием и корреляционной матрицей $(\delta_{ij} - \sqrt{p_i p_j})$, где $\delta_{ij} = 1$ при $i = j$, $\delta_{ij} = 0$ в противном случае.

При любом ортогональном преобразовании $\zeta = C\gamma$ имеем $\sum_{k=1}^r \zeta_k^2 = \sum_{k=1}^r \gamma_k^2$. Подберем C так, чтобы γ_k были независимы. Имеем

$$M \zeta_k^2 = k_{ii} = 1 - p_i.$$

Отсюда, просуммировав по k , получим

$$M \sum_{k=1}^r \gamma_k^2 = r - 1. \quad (9)$$

Далее,

$$M \left(\sum_{k=1}^r \gamma_k^2 \right)^2 = M \left(\sum_{k=1}^r \zeta_k^2 \right)^2 = \sum_{k=1}^r M \zeta_k^4 + \sum_{i \neq j} M \zeta_i^2 \zeta_j^2. \quad (10)$$

Поскольку ζ_k — нормальная величина с параметрами $0, 1 - p_k$, то $M \zeta_k^4 = 3(1 - p_k)^2$. Характеристическая функция случайной ве-

личины (ξ_i, ξ_j) имеет вид

$$\varphi(t_1, t_2) = \exp \left\{ -\frac{1}{2} ((1-p_i)t_1^2 - 2\sqrt{p_i p_j} t_1 t_2 + (1-p_j)t_2^2) \right\}.$$

Отсюда

$$M_{\xi_1 \xi_2}^2 = \frac{\partial^4 \varphi(t_1, t_2)}{\partial t_1^2 \partial t_2^2} \Big|_{t_1=t_2=0} = 1 - p_i - p_j + 3p_i p_j.$$

Теперь из (10) имеем

$$M \left(\sum_{k=1}^r \gamma_k^2 \right)^2 = 3 \sum_{k=1}^r (1-p_k)^2 + \sum_{i \neq j} (1-p_i - p_j + 3p_i p_j).$$

Это равенство легко сводится к виду

$$M \left(\sum_{k=1}^r \gamma_k^2 \right)^2 = r^2 - 1. \quad (11)$$

Обозначим $x_k = D\alpha_k$. Тогда из (9) имеем

$$\sum_{k=1}^r x_k = r - 1. \quad (12)$$

Из (11), учитывая независимость γ_k , получаем

$$3 \sum_{k=1}^r x_k^2 + \sum_{i \neq k} x_i x_k = r^2 - 1,$$

или после упрощения и подстановки (12)

$$\sum_{k=1}^r x_k^2 = r - 1. \quad (13)$$

Поскольку

$$\begin{aligned} D \left(\sum_{k=1}^r \xi_k \sqrt{p_k} \right) &= \sum_{k=1}^r p_k D\xi_k + \sum_{i \neq j} \text{cov}(\xi_i, \xi_j) \sqrt{p_i p_j} = \\ &= \sum_{k=1}^r p_k (1-p_k) - \sum_{i \neq j} p_i p_j = 0, \end{aligned}$$

то ξ имеет вырожденное нормальное распределение. Тогда одна из γ_k равна нулю с вероятностью 1. Положим для определенности $\gamma_r = 0$. Тогда из (9) и (13) имеем

$$\sum_{k=1}^{r-1} (x_k - 1)^2 = 0,$$

а следовательно, $x_1 = \dots = x_{r-1} = 1$. Итак, χ^2 распадается в сумму квадратов $r-1$ независимых стандартных нормальных величин, т. е. имеет распределение χ_{r-1}^2 . Окончательно получаем, что статистика χ^2 при $n \rightarrow \infty$ имеет предельное распределение

$$P(\chi^2 < z) \xrightarrow{n \rightarrow \infty} P(\chi_{r-1}^2 < z). \quad (14)$$

Точку z , определяющую критерий хи-квадрат, при большом n следует выбирать из условия

$$P(\chi^2_{r-1} \geq z) = \alpha. \quad (15)$$

Распределение χ^2 табулировано, так что уравнение (15) можно решать табличным способом.

§ 5. Статистическая оценка параметров

Рассмотрим статистическую модель $\{p(x, \theta), \theta \in \Theta\}$, где Θ — область r -мерного пространства, $p(x, \theta)$ — плотность вероятности m -мерной случайной величины, зависящая от параметра θ . Задача статистической оценки параметра формулируется так. Считается известным, что наблюдаемая величина имеет плотность $p(x, \theta)$ при некотором «истинном» значении параметра θ , однако это значение не известно. По выборке X объема n , т. е. по независимым наблюдениям x_1, \dots, x_n случайной величины ξ , требуется построить оценку $\tilde{\theta}$ параметра θ — такую статистику, которую в определенных условиях можно было бы использовать вместо неизвестного θ .

В математической статистике формулируют некоторые желательные свойства статистических оценок, среди которых на первом месте — состоятельность. Обозначим $\tilde{\theta} = \tilde{\theta}_n$, указывая этим зависимость оценки от объема выборки. Оценка $\tilde{\theta}_n$ называется *состоятельной*, если

$$\tilde{\theta}_n \xrightarrow[n \rightarrow \infty]{p} \theta. \quad (1)$$

Таким образом, состоятельность — свойство не одной оценки $\tilde{\theta}_n$, а целой последовательности оценок $(\tilde{\theta}_n)$. Если оценка состоятельна, то это еще ничего не говорит о ее качестве при данном фиксированном объеме выборки n , но указывает на принципиальную возможность оценки параметра со сколь угодно высокой точностью при достаточно большом n .

Пусть $\tau(\theta)$ — числовая или векторная функция параметра θ . Статистика $\tilde{\tau} = \tilde{\tau}_n$ называется *состоятельной оценкой* $\tau(\theta)$, если

$$\tilde{\tau}_n \xrightarrow[n \rightarrow \infty]{p} \tau(\theta). \quad (2)$$

Это определение — обобщение (1).

Так, пусть $\tau = Mf(\xi)$, где $f(x)$ — числовая функция, причем $M_{\theta}f(\xi)$ конечно. Рассмотрим оценку

$$\tilde{\tau}_n = \frac{1}{n} \sum_{k=1}^n f(x_k),$$

где x_k — независимые наблюдения случайной величины ξ . Состоятельность этой оценки следует из закона больших чисел в формуле Хинчина.

Следующее желательное свойство статистической оценки — *несмещенность* — относится к фиксированному объему выборки n . Оценка $\tilde{\theta}$ ($\tilde{\tau}$) называется *несмещенной оценкой* параметра θ (его числовой (векторной) функции $\tau(\theta)$), если

$$M_{\theta}\tilde{\theta} = \theta \quad (3)$$

и соответственно

$$M_{\theta}\tilde{\tau} = \tau(\theta). \quad (4)$$

Так, рассмотренная в предыдущем примере оценка $\bar{\tau}_n$ является несмещенной.

Оценка $\tilde{\theta}$ одномерного параметра θ (оценка $\tilde{\tau}$ числовой функции $\tau(\theta)$) называется *эффективной*, если она несмещенная и обладает наименьшей дисперсией среди всех несмещенных оценок.

Таким образом, условие эффективности оценки записывается так:

$$D\tilde{\theta} \Rightarrow \min, M_{\theta}\tilde{\theta} = \theta \quad (5)$$

и соответственно

$$D\tilde{\tau} \Rightarrow \min, M_{\theta}\tilde{\tau} = \tau(\theta). \quad (6)$$

В статистической модели $\{p(x, \theta), \theta \in \Theta\}$ будем считать, что Θ — конечный или бесконечный числовой интервал, $p(x, \theta)$ — плотность в \mathbb{R}^m относительно некоторой меры. (Если ξ — непрерывная случайная величина, то $p(x, \theta)$ — обычная плотность; если ξ — дискретная, то $p(x, \theta) = P_{\theta}(\xi = x)$.) Предположим, что при любом $\theta \in \Theta$ плотность положительна в одной и той же области из \mathbb{R}^m .

Считая для определенности, что $p(x, \theta)$ — обычная плотность, запишем условие несмещенности оценки $\tilde{\tau} = \tilde{\tau}(X)$ параметра $\tau(\theta)$ в виде

$$\int \tilde{\tau}(X) p(X, \theta) dX = \tau(\theta), \quad (7)$$

где для краткости обозначено $p(X, \theta) dX = p(x_1, \theta) \dots p(x_n, \theta) \times \times dx_1 \dots dx_n$. Далее, очевидно,

$$\int p(X, \theta) dX = 1. \quad (8)$$

Предположим, что левые части равенств (7) и (8) можно продифференцировать по θ под знаком интеграла. Производя замену

$$\frac{\partial p(X, \theta)}{\partial \theta} = \frac{\partial \ln p(X, \theta)}{\partial \theta} p(X, \theta),$$

получим

$$\int \tilde{\tau}(X) \frac{\partial \ln p(X, \theta)}{\partial \theta} p(X, \theta) dX = \tau'(\theta), \quad (9)$$

$$\int \frac{\partial \ln p(X, \theta)}{\partial \theta} p(X, \theta) dX = 0. \quad (10)$$

Из равенств (9) и (10) находим

$$\int (\tilde{\tau}(X) - \tau(\theta)) \frac{\partial \ln p(X, \theta)}{\partial \theta} p(X, \theta) dX = \tau'(\theta), \quad (11)$$

или, что то же самое,

$$M_{\theta}(\tilde{\tau}(X) - \tau(\theta)) \frac{\partial \ln p(X, \theta)}{\partial \theta} = \tau'(\theta). \quad (12)$$

Неравенство Коши—Буняковского

$$|M\alpha\beta|^2 \leq M\alpha^2 \cdot M\beta^2$$

позволяет вывести из (12) неравенство

$$(\tau'(\theta))^2 \leq M(\tilde{\tau}(X) - \tau(\theta))^2 \cdot M\left(\frac{\partial \ln p(X, \theta)}{\partial \theta}\right)^2,$$

или, что то же самое,

$$D\tilde{\tau} \geq (\tau'(\theta))^2 / M\left(\frac{\partial \ln p(X, \theta)}{\partial \theta}\right)^2. \quad (13)$$

Согласно предыдущему,

$$\int \frac{\partial \ln p(X, \theta)}{\partial \theta} p(X, \theta) dX = 0.$$

Предположим, что возможно дифференцирование под знаком интеграла. Получим

$$\int \frac{\partial^2 \ln p(X, \theta)}{\partial \theta^2} p(X, \theta) dX + \int \left(\frac{\partial \ln p(X, \theta)}{\partial \theta}\right)^2 p(X, \theta) dX = 0.$$

Отсюда знаменатель правой части (13) преобразуется к виду

$$I = - \int \frac{\partial^2 \ln p(X, \theta)}{\partial \theta^2} p(X, \theta) dX.$$

Поскольку $\ln p(X, \theta) = \ln p(x_1, \theta) + \dots + \ln p(x_n, \theta)$, то

$$I = -n \frac{\partial^2 \ln p(x, \theta)}{\partial \theta^2} p(x, \theta) dx = -nM_{\theta} \frac{\partial^2 \ln p(\xi, \theta)}{\partial \theta^2}.$$

Подстановка в (13) приводит к неравенству

$$D_{\theta}\tilde{\tau} \geq (\tau'(\theta))^2 / \left(-nM_{\theta} \frac{\partial^2 \ln p(\xi, \theta)}{\partial \theta^2}\right), \quad (14)$$

называемому *неравенством Рао — Крамера*. Подставив $\tau(\theta) = \theta$, для дисперсии любой несмещенной оценки $\hat{\theta}$ параметра θ получаем неравенство

$$D_{\theta}\hat{\theta} \geq 1 / \left(-nM_{\theta} \frac{\partial^2 \ln p(\xi, \theta)}{\partial \theta^2}\right), \quad (15)$$

которое также называется *неравенством Рао — Крамера*. Обозначив

$$-M_{\theta} \frac{\partial^2 \ln p(\xi, \theta)}{\partial \theta^2} = i(\theta),$$

можно записать (14), (15) более кратко:

$$D_6 \tilde{\tau} \geq (\tau'(\theta))^2 / (n i(\theta)), \quad (16)$$

$$D_6 \tilde{\theta} \geq 1 / (n i(\theta)). \quad (17)$$

При выводе неравенства Рао — Крамера мы предполагали, что некоторые выражения допускают дифференцирование под знаком интеграла. Можно доказать, что неравенство (13) выполняется в любом случае, если справедливы формулы (9) и (10); что же касается неравенства Рао — Крамера в форме (15), то для его справедливости достаточно дополнительное условие: существование $\frac{\partial^2 p(x, \theta)}{\partial \theta^2}$.

Можно сделать практический вывод из неравенства Рао — Крамера. Если $\tilde{\theta}$ или $\tilde{\tau}$ — несмещенная оценка θ (соответственно $\tau(\theta)$), причем дисперсия этой оценки равна правой части (14) (соответственно (15)), то $\tilde{\theta}$ ($\tilde{\tau}$) есть эффективная оценка.

Пример 1. Рассматривается статистическая модель $\{p(x; a, \sigma^2), a \in \mathbb{R}\}$, где $p(x; a, \sigma^2)$ — плотность нормального распределения с параметрами a, σ^2 . Предположим, что σ^2 фиксировано. Статистика $\bar{x} = \frac{1}{n} (x_1 + \dots + x_n)$, очевидно, является несмещенной оценкой параметра a . Докажем, что она эффективна.

Поскольку

$$\ln p(x; a, \sigma^2) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x-a)^2}{2\sigma^2},$$

то

$$\frac{\partial \ln p(x; a, \sigma^2)}{\partial a^2} = \frac{1}{\sigma^2}; \quad I = n/\sigma^2.$$

Из (15) находим, что для любой несмещенной оценки \tilde{a} параметра a $D\tilde{a} \geq \sigma^2/n$. В то же время, очевидно, $D\bar{x} = \sigma^2/n$. Следовательно, \bar{x} — эффективная оценка параметра a .

Пример 2. Пусть математическое ожидание a нормальной величины известно, σ^2 не известно. Имеем

$$\frac{\partial^2 \ln p(x; a, \sigma^2)}{\partial (\sigma^2)^2} = \frac{1}{2\sigma^4} - \frac{(x-a)^2}{\sigma^6}.$$

Взяв математическое ожидание, получим

$$M \frac{\partial^2 \ln p(x; a, \sigma^2)}{\partial (\sigma^2)^2} = -\frac{1}{2\sigma^4}.$$

Итак, для несмещенной оценки $\tilde{\sigma}^2$ параметра σ^2 должно быть $D\tilde{\sigma}^2 \geq 2\sigma^4/n$. Взяв несмещенную оценку

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2,$$

получим

$$D\tilde{\sigma}^2 = \frac{1}{n} D(\xi - a)^2 = \frac{1}{n} (M(\xi - a)^4 - (M(\xi - a)^2)^2) = 2\sigma^4/n.$$

Снова получили эффективную оценку.

Неравенство Рао — Крамера позволяет доказать, что эффективными являются:

оценка вероятности успеха частотой в схеме независимых испытаний;

оценка параметра a распределения Пуассона средним арифметическим результатов независимых наблюдений пуассоновской случайной величины;

оценка параметра T экспоненциального закона

$$p(x; T) = \frac{1}{T} e^{-x/T}, \quad x \geq 0,$$

средним арифметическим результатов независимых наблюдений.

Пример 3. Пусть $p(x, \theta)$ — плотность равномерного распределения в интервале $(0; \theta)$, где $\theta > 0$ — неизвестный параметр. Как несложно подсчитать, формальная нижняя граница для дисперсии несмещенной оценки в неравенстве Рао — Крамера равна θ^2/n . В то же время, для несмещенной оценки $T_n^* = \frac{n+1}{n} \max\{x_1, \dots, x_n\}$ имеем

$$DT_n^* = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n}.$$

Кажущееся противоречие объясняется невозможностью дифференцировать левые части (9), (10) под знаком интеграла.

§ 6. Асимптотические свойства статистических оценок

Пусть $\tilde{\theta} = \tilde{\theta}_n$ — оценка числового параметра θ по выборке X объема n . Эффективная оценка $\tilde{\theta}_n^*$ существует не всегда, но если она существует, то выполняется равенство

$$M_{\theta}(\tilde{\theta}_n^* - \theta)^2 = D_{\theta}\tilde{\theta}_n^* = 1 / \left(-n M_{\theta} \frac{\partial^2 \ln p(x, \theta)}{\partial \theta^2} \right). \quad (1)$$

Отношение правой части равенства (1) к $M_{\theta}(\tilde{\theta} - \theta)^2$ называется *эффективностью* оценки $\tilde{\theta}_n$ и обозначается $e_{\theta}(\tilde{\theta})$.

Пусть требуется, чтобы математическое ожидание квадрата отклонения оценки от истинного значения параметра было не больше ε . Сколько для этого требуется наблюдений? Для эффективной оценки достаточно, чтобы выполнялось условие $n \geq 1/(\varepsilon i(\theta))$, для

оценки $\tilde{\theta}$ — условие $n \geq 1/(\varepsilon i(\theta) e_{\theta}(\tilde{\theta}))$. Пренебрегая эффектом округления, скажем так: эффективность — обратная величина к увеличению числа наблюдений по сравнению с эффективной оценкой при той же точности оценок. Скажем, при $e(\tilde{\theta}) = 0,2$ требуется в 5 раз больше наблюдений, чем в случае эффективной оценки. Если существует $\bar{e}_{\theta}(\tilde{\theta}) = \lim_{n \rightarrow \infty} e_{\theta}(\tilde{\theta}_n)$, то это число называется *асимптотической эффективностью* оценки $\tilde{\theta} = (\tilde{\theta}_n)$.

Оценка $\tilde{\theta} = (\tilde{\theta}_n)$ называется *асимптотически эффективной*, если $\bar{e}_{\theta}(\tilde{\theta}) = 1$ при всех

$\theta \in \Theta$. Понятия эффективности и асимптотической эффективности для оценок $\tau(\theta)$ вводятся аналогично. При достаточно большом объеме выборки целесообразно использование асимптотически эффективных оценок.

Оценка $\tilde{\theta} = (\tilde{\theta}_n)$ называется *асимптотически нормальной*, *асимптотически эффективной*, если при $n \rightarrow \infty$ случайная величина

$$\sqrt{ni}(\theta)(\tilde{\theta}_n - \theta)$$

асимптотически нормальна с параметрами 0, 1, т. е.

$$P_{\theta}(\tilde{\theta}_n - \theta < x/\sqrt{ni}(\theta)) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt. \quad (2)$$

Пусть, например, $\tilde{\theta} = \tilde{\theta}_n$ — оценка максимального правдоподобия параметра θ , т. е. решение задачи на экстремум:

$$p(X, \tilde{\theta}) = \max_{\theta} p(X, \theta). \quad (3)$$

Тогда при соответствующих аналитических условиях оценка $\tilde{\theta}$ удовлетворяет условию (2). Отметим основные моменты доказательства.

Для отличия от остальных значений параметра θ его истинное значение обозначим θ_0 . Обозначим также

$$f(\theta) = \frac{1}{n} \frac{\partial \ln p(X, \theta)}{\partial \theta}.$$

Тогда $f(\tilde{\theta}) = 0$. Имеем

$$f(\theta_0) = \frac{1}{n} \sum_{k=1}^n \left. \frac{\partial \ln p(x_k, \theta)}{\partial \theta} \right|_{\theta=\theta_0}.$$

Слагаемые этой суммы независимы, одинаково распределены. Каждое из них имеет математическое ожидание

$$M_{\theta_0} \left. \frac{\partial \ln p(x_k, \theta)}{\partial \theta} \right|_{\theta=\theta_0} = 0$$

и дисперсию

$$M_{\theta_0} \left(\left. \frac{\partial \ln p(x_k, \theta)}{\partial \theta} \right|_{\theta=\theta_0} \right)^2 = i(\theta_0).$$

Итак, по центральной предельной теореме для одинаково распределенных слагаемых

$$f(\theta_0) = \sqrt{\frac{i(\theta_0)}{n}} \zeta, \quad (4)$$

где ζ — асимптотически нормальная с параметрами 0, 1 случайная величина.

Далее,

$$f'(\theta_0) = \frac{1}{n} \sum_{k=1}^n \left. \frac{\partial^2 \ln p(x_k, \theta)}{\partial \theta^2} \right|_{\theta=\theta_0}.$$

Математическое ожидание каждого слагаемого этой суммы равно $-i(\theta_0)$. По закону больших чисел $f'(\theta_0) \xrightarrow[n \rightarrow \infty]{p} -i(\theta_0)$. Если третья производная функции $\ln p(x, \theta)$ по θ ограничена, то в окрестности точки $\theta = \theta_0$ имеем

$$f(\theta) = f(\theta_0) + f'(\theta_0)(\theta - \theta_0) + O((\theta - \theta_0)^2).$$

Учитывая сказанное ранее о $f(\theta_0)$ и $f'(\theta_0)$, приходим к выводу, что с вероятностью, стремящейся к 1 при $n \rightarrow \infty$, существует решение уравнения $f(\tilde{\theta}) = 0$, имеющее вид

$$\tilde{\theta} = \theta_0 + \zeta' / \sqrt{ni(\theta_0)},$$

где ζ' — «слегка» искаженное ζ , т. е. $\zeta'/\zeta \xrightarrow[n \rightarrow \infty]{p} 1$. Для доказательства сформулированного утверждения (2) достаточно ввести условие: решение уравнения $f(\tilde{\theta}) = 0$ с вероятностью, стремящейся к единице при $n \rightarrow \infty$, единственно.

§ 7. Метод наименьших квадратов

Если имеется несколько неизвестных параметров ($p(x; \theta_1, \dots, \theta_r)$ вместо $p(x, \theta)$), то оценка максимального правдоподобия определяется решением системы уравнений

$$\frac{\partial \ln p(X; \theta_1, \dots, \theta_r)}{\partial \theta_j} = 0, \quad 1 \leq j \leq r. \quad (1)$$

Важнейший пример — обоснование известного метода наименьших квадратов, используемого для сглаживания экспериментальных зависимостей. Пусть имеется две числовые переменные x и y . Предполагается, что

$$y = a_0 \varphi_0(x) + \dots + a_r \varphi_r(x), \quad (2)$$

где $\varphi_i(x)$ — известные функции, a_i — неизвестные параметры, которые и нужно оценить.

Пусть в точках x_1, \dots, x_n (не обязательно различных) произведены наблюдения y_1, \dots, y_n переменной y ; отклонения наблюдений от истинных значений этой переменной — независимые $N(0, \sigma^2)$ случайные величины. Логарифм совместной плотности y_1, \dots, y_n с точностью до постоянной равен выражению $-z^2/\sigma^2$, где

$$z^2 = \frac{1}{2} \sum_{k=1}^n \left(y_k - \sum_{j=0}^r a_j \varphi_j(x_k) \right)^2. \quad (3)$$

Дифференцирование z^2 по a_0, \dots, a_r приводит к уравнениям

$$\sum_{k=1}^n \left(y_k - \sum_{j=0}^r a_j \varphi_j(x_k) \right) \varphi_i(x_k) = 0, \quad 0 \leq i \leq r,$$

или после группировки слагаемых

$$\sum_{j=0}^r \left(\sum_{k=1}^n \varphi_j(x_k) \varphi_i(x_k) \right) a_j = \sum_{k=1}^n y_k \varphi_i(x_k), \quad 0 \leq i \leq r. \quad (4)$$

Пример. Зависимость между y и x — параболическая: $y = a_0 + a_1x + a_2x^2$. Получены наблюдения

x	— 4	— 3	— 2	— 1	0	1	2	3	4	5
y	10,14	8,28	6,85	5,79	5,02	4,60	4,77	5,32	6,21	7,54

Требуется оценить параметры a_0, a_1, a_2 по методу наименьших квадратов. Система уравнений (4) в данном случае принимает вид

$$\begin{aligned} na_0 + (\Sigma x_k) a_1 + (\Sigma x_k^2) a_2 &= \Sigma y_k, \\ (\Sigma x_k) a_0 + (\Sigma x_k^2) a_1 + (\Sigma x_k^3) a_2 &= \Sigma y_k x_k, \\ (\Sigma x_k^2) a_0 + (\Sigma x_k^3) a_1 + (\Sigma x_k^4) a_2 &= \Sigma y_k x_k^2. \end{aligned}$$

В данном случае имеем $n = 10$, $\Sigma x_k = 5$, $\Sigma x_k^2 = 85$, $\Sigma x_k^3 = 125$, $\Sigma x_k^4 = 1333$, $\Sigma y_k = 64,52$, $\Sigma y_k x_k = 7,75$, $\Sigma y_k x_k^2 = 629,37$. Таким образом, получена система уравнений

$$\begin{aligned} 10a_0 + 5a_1 + 85a_2 &= 64,52, \\ 5a_0 + 85a_1 + 125a_2 &= 7,75, \\ 85a_0 + 125a_1 + 1333a_2 &= 629,37. \end{aligned}$$

Решение этой системы имеет вид $a_0 \approx 5,0027$, $a_1 \approx -0,4968$, $a_2 \approx 0,1997$. В результате получаем приближенную зависимость переменных:

$$y \approx 5,0027 - 0,4968x + 0,1997x^2.$$

§ 8. Метод моментов

Пусть $X = (x_1, \dots, x_n)$ — выборка независимых наблюдений случайной величины ξ . Допустим, что конечен момент порядка r :

$$\alpha_r = \alpha_r(\xi) = M\xi^r.$$

Легко видеть, что несмещенной оценкой α_r служит статистика

$$\alpha_r^* = \frac{1}{n} \sum_{k=1}^n x_k^r, \quad (1)$$

называемая *эмпирическим моментом* порядка r . Дисперсия этой оценки

$$D\alpha_r^* = \frac{1}{n} (\alpha_{2r} - \alpha_r^2). \quad (2)$$

При $n \rightarrow \infty$ оценка α_r^* асимптотически нормальна; таким образом, отклонение

$$\alpha_r^* - \alpha_r = \zeta \sqrt{(\alpha_{2r} - \alpha_r^2)/n},$$

где ζ имеет стандартное нормальное предельное распределение при $n \rightarrow \infty$.

Пусть теперь f — некоторая числовая характеристика распределения, определяемая его моментами:

$$f = f(\alpha_1, \dots, \alpha_m).$$

Подставив вместо α_r эмпирические моменты α_r^* , получим оценку $f^* = f(\alpha_1^*, \dots, \alpha_m^*)$ характеристики f , называемую *моментной оцен-*

кой. Если функция $f(\alpha_1, \dots, \alpha_m)$ дифференцируема, то моментная оценка состоятельна и асимптотически нормальна. Она, тем не менее, не всегда несмещенная; это связано с нелинейностью функции $f(\alpha_1, \dots, \alpha_m)$. Так, рассмотрим моментную оценку $D^*\xi$ дисперсии σ^2 случайной величины ξ . Имеем

$$\sigma^2 = \alpha_2 - \alpha_1^2.$$

Таким образом, моментная оценка имеет вид

$$D^*\xi = \alpha_2^* - (\alpha_1^*)^2, \quad (3)$$

или более подробно

$$D^*\xi = \frac{1}{n} \sum_{k=1}^n x_k^2 - \bar{x}^2, \quad (4)$$

где $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$. Обозначим $a = M\xi$. Тогда $Mx_k = a$, $Mx_k^2 = a^2 + \sigma^2$, $Mx_i x_k = a^2$ при $i \neq k$. Используя равенство (4), находим

$$\begin{aligned} MD^*\xi &= \frac{1}{n} \sum_{k=1}^n Mx_k^2 - \frac{1}{n^2} \sum_{i,k=1}^n Mx_i x_k = \\ &= a^2 + \sigma^2 - \frac{1}{n^2} (n(a^2 + \sigma^2) + n(n-1)a^2) = \frac{n-1}{n} \sigma^2. \end{aligned}$$

Итак, оценка $D^*\xi$ — смещенная. Легко видеть, что

$$s^2 = \frac{n}{n-1} D^*\xi = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$$

является несмещенной оценкой σ^2 .

§ 9. Доверительные интервалы

Пусть имеется статистическая модель $\{F(x, \theta), \theta \in \Theta\}$, $\tau(\theta)$ — числовая функция числового параметра θ ; X — выборка объема n . Если $a_1(X)$, $a_2(X)$ — статистики, удовлетворяющие условию

$$P_\theta(a_1(X) \leq \tau(\theta) \leq a_2(X)) \geq \alpha, \quad \theta \in \Theta, \quad (1)$$

то случайный интервал $[a_1(X); a_2(X)]$ называется *доверительным интервалом* (для $\tau(\theta)$) с доверительным уровнем α ¹.

Стандартный прием построения доверительного интервала для параметра θ состоит в следующем. Пусть $\varphi(X)$ — числовая статистика, функция распределения которой непрерывна и монотонна по параметру θ (например, $F(x, \theta)$ убывает по аргументу θ). Возьмем два числа $\varepsilon_1, \varepsilon_2$, в сумме равные $\varepsilon = 1 - \alpha$, и при любом θ

¹ Если функции распределения статистик $a_i(X)$ непрерывны, то замкнутый и открытый интервалы в вероятностном смысле эквивалентны.

определим $b_1(\theta)$, $b_2(\theta)$ из условий $F_\theta(b_1(\theta)) = \varepsilon_1$, $1 - F_\theta(b_2(\theta)) = \varepsilon_2$, так что

$$P_\theta(b_1(\theta) \leq \varphi(X) \leq b_2(\theta)) = \alpha. \quad (2)$$

В плоскости $(\varphi(X), \theta)$ образовалась «полоса», определенная условиями $b_1(\theta) \leq \varphi(X) \leq b_2(\theta)$. Определим теперь случайный отрезок $[a_1(X); a_2(X)]$ как пересечение «полосы» с прямой $\varphi(X) = \text{const}$. Он и будет искомым доверительным интервалом. В самом деле, при любом истинном значении θ параметра это значение попадает на отрезок $[a_1(X); a_2(X)]$ в том и только в том случае, если $b_1(\theta) \leq \varphi(X) \leq b_2(\theta)$ (рис. 29), вероятность же такого события, согласно (2), равна α . Подобным же образом строится доверительный интервал и для монотонной функции $\tau(\theta)$ неизвестного параметра.

Пример. Пусть θ — неизвестное математическое ожидание нормальной случайной величины с известной дисперсией σ^2 . Определим z из условия

$$\frac{1}{\sqrt{2\pi}} \int_{-z}^z e^{-x^2/2} dx = \alpha.$$

Возьмем, далее, $\varphi(X) = \bar{x}$. Положим $b_1(\theta) = \theta - \sigma z / \sqrt{n}$, $b_2(\theta) = \theta + \sigma z / \sqrt{n}$. Тогда $P_\theta(b_1(\theta) \leq \bar{x} \leq b_2(\theta)) = \alpha$. Пересечение прямой $\bar{x} = \text{const}$ с «полосой» происходит по отрезку $[\bar{x} - \sigma z / \sqrt{n}; \bar{x} + \sigma z / \sqrt{n}]$; это и есть искомым доверительный интервал.

В выборе значений ε_1 , ε_2 имеется определенный произвол. Например, можно положить $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$ (такой доверительный интервал называется *симметричным*); можно положить $\varepsilon_1 = \varepsilon$, $\varepsilon_2 = 0$ или, наоборот, $\varepsilon_1 = 0$, $\varepsilon_2 = \varepsilon$ (*односторонний* доверительный интервал). Конкретные значения ε_1 , ε_2 определяются соображениями опасности выхода θ за пределы доверительного интервала в ту или иную сторону. Так, если строится интервал для вероятности безотказности изделия, выпускаемого заводом, то переоценить эту вероятность значительно опаснее, чем недооценить; поэтому в подобных случаях предпочитают односторонние доверительные интервалы.

Часто встречается статистическая модель $\{F(x; \theta_1, \theta_2)\}$, где θ_1 — числовой параметр, для которого требуется построить доверительный интервал, θ_2 — числовой или векторный параметр, создающий неоднозначность при рассмотрении тех или иных вероятностей и поэтому называемый *мешающим параметром*. В такой ситуации пытаются найти статистику $\varphi(X)$ с распределением, зависящим только от θ_1 , но не от θ_2 . Коль скоро такая статистика найдена, пытаются применить описанную выше процедуру.

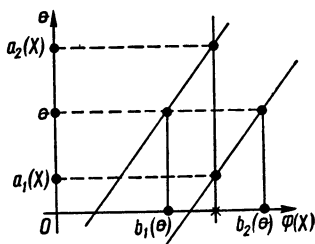


Рис. 29

Пример 1. Оба параметра a , σ^2 нормального распределения не известны; требуется построить доверительный интервал уровня α для параметра σ^2 .

Как известно, статистика $\sum_{k=1}^n (x_k - \bar{x})^2 / \sigma^2$ имеет распределение χ^2 с $n - 1$ степенями свободы. Поэтому обозначим эту статистику χ_{n-1}^2 . Найдем такие z_1, z_2 , что

$$P(z_1 \leq \chi_{n-1}^2 \leq z_2) = \alpha.$$

Это то же самое, что

$$P\left(\frac{1}{z_2} \leq \sigma^2 / \sum_{k=1}^n (x_k - \bar{x})^2 \leq \frac{1}{z_1}\right) = \alpha.$$

Видим, что доверительный интервал уровня α для параметра σ^2 имеет вид

$$\left[\frac{1}{z_2} \sum_{k=1}^n (x_k - \bar{x})^2; \frac{1}{z_1} \sum_{k=1}^n (x_k - \bar{x})^2 \right] = \left[\frac{(n-1)s^2}{z_2}; \frac{(n-1)s^2}{z_1} \right].$$

Пример 2. В той же модели требуется построить доверительный интервал уровня α для параметра a . На этот раз решающим параметром является σ^2 .

Обозначим $s = \left(\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2 \right)^{1/2}$. В § 3.10 было доказано, что n -мерная случайная величина $(x_1 - \bar{x}, \dots, x_n - \bar{x})$ и случайная величина \bar{x} независимы. Далее, можно записать

$$\begin{aligned} \sum_{k=1}^n (x_k - a)^2 &= \sum_{k=1}^n (x_k - \bar{x} + \bar{x} - a)^2 = \\ &= \sum_{k=1}^n (x_k - \bar{x})^2 + 2(\bar{x} - a) \sum_{k=1}^n (x_k - \bar{x}) + n(\bar{x} - a)^2. \end{aligned}$$

Поскольку средняя сумма, очевидно, равна нулю, то

$$\sum_{k=1}^n (x_k - a)^2 = \sum_{k=1}^n (x_k - \bar{x})^2 + n(\bar{x} - a)^2.$$

В левой части имеем $\sigma^2 \chi_n^2$, где χ_n^2 — сумма n независимых стандартных нормальных величин. Таким образом,

$$\chi_n^2 = \frac{n-1}{\sigma^2} s^2 + \frac{n}{\sigma^2} (\bar{x} - a)^2.$$

Второе слагаемое правой части этого равенства есть квадрат стандартной нормальной величины, т. е. χ_1^2 . Поскольку слагаемые независимы, то

$$\varphi_{\chi_1^2}^n(t) = \varphi_{\chi_n^2}(t) = \varphi_{ns^2/\sigma^2}(t) \varphi_{\chi_1^2}(t),$$

откуда

$$\varphi_{ns^2/\sigma^2}(t) = \varphi_{\chi_1^2}^{n-1}(t) = \varphi_{\chi_{n-1}^2}(t).$$

Итак, случайная величина ns^2/σ^2 имеет распределение χ_{n-1}^2 . Теперь очевидно, что случайная величина

$$t_{n-1} = \sqrt{n} (\bar{x} - a)/s,$$

где $s = \left(\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2 \right)^{1/2}$, имеет распределение Стьюдента с $n-1$ степенями свободы. Найдем такое z , при котором

$$P(|t_{n-1}| < z) = \alpha.$$

Тогда

$$P(|\bar{x} - a| < zs/\sqrt{n}) = \alpha.$$

Следовательно, случайный отрезок

$$\left[\bar{x} - \frac{zs}{\sqrt{n}}; \bar{x} + \frac{zs}{\sqrt{n}} \right]$$

является симметричным доверительным интервалом уровня α параметра a .

Приближенные доверительные интервалы при большом n строятся из более простых соображений, и притом при произвольном (но с конечной дисперсией) распределении ξ . Имеем $D\xi \approx s$, откуда

$$P\left(\bar{x} - \frac{z_0 s}{\sqrt{n}} \leq a \leq \bar{x} + \frac{z_0 s}{\sqrt{n}}\right) \approx \alpha,$$

где z_0 определяется решением уравнения

$$\frac{1}{\sqrt{2\pi}} \int_{-z_0}^{z_0} e^{-x^2/2} dx = \alpha.$$

Задачи

1. Имеется n независимых наблюдений над двумерной нормальной случайной величиной (ξ, η) с нулевым математическим ожиданием. При гипотезах H_0 и H_1 корреляционная матрица наблюдаемой величины равна соответственно

$$H_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 1 & 0,4 \\ 0,4 & 1 \end{pmatrix}.$$

Выписать выражение для логарифма отношения правдоподобия. Найти вид порогового уровня z в критерии Неймана — Пирсона.

2. Распознается изображение по проступающим пятнышкам. При гипотезе H_k , $k = 0, 1$, изображение есть треугольник с вершинами в точках $e^{ik\pi/3}$, $e^{i(k+2)\pi/3}$, $e^{i(k+4)\pi/3}$ комплексной плоскости. Проступило n пятнышек пренебрежимо малого размера, распределенных по площади треугольника (истинного изображения) по равномерному закону.

а) Решить байесовскую задачу различения гипотез при одинаковом убытке от ошибок первого и второго рода.

б) Найти вид критерия Неймана — Пирсона.

3. Построить критерий Неймана — Пирсона для различения гипотез $H_0 = \{\xi \text{ распределена равномерно в интервале } (-\Delta; \Delta)\}$ и $H_1 = \{\xi = N(0, \sigma^2)\}$ по n независимым наблюдениям величины ξ . Построить также байесовский критерий. При каком значении отношения σ/Δ минимально достижимое значение суммы вероятностей ошибок первого и второго рода достигает максимума?

4. Наблюдается пуассоновская случайная величина с параметром λ . Как по единственному наблюдению различать гипотезы $H_0 = \lambda_0$ и $H_1 = \lambda_1$, где $\lambda_1 > \lambda_0$, если ошибка второго рода в k раз опаснее, чем ошибка первого рода?

5. Получить в качестве следствия теоремы § 10.3 равномерно наиболее мощные критерии в следующих постановках задачи:

а) $p_\lambda(x) = \lambda e^{-\lambda x}$, $x > 0$, $H = \{\lambda \leq \lambda_0\}$, $K = \{\lambda > \lambda_0\}$;

б) $p_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\theta)^2/2\sigma^2}$, $H = \{\theta \leq \theta_0\}$, $K = \{\theta > \theta_0\}$;

$$в) p_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}, H = \{\sigma \leq \sigma_0\}, K = \{\sigma > \sigma_0\};$$

$$г) p_a(x) = e^{-a|x|}, x = 0, 1, \dots, H = \{a \leq a_0\}, K = \{a > a_0\}.$$

6. При гипотезе H случайная величина ξ имеет функцию распределения $F(x) = x, 0 \leq x \leq 1$. Найти распределение статистики $\Delta_n = \sup_x |F_n(x) - F(x)|$ при гипотезе H в частных случаях $n = 1$ и $n = 2$.

7. Как ведет себя статистика $(1/r^2)\chi^2$, где χ^2 задана формулой (4.7), когда $p_k = 1/r, n$ фиксировано и $r \rightarrow \infty$, если теоретическое распределение непрерывно?

8. Случайная величина ξ имеет в точке x непрерывную плотность, значение которой в данной точке равно θ . Для оценки θ используем статистику

$$\hat{\theta}_n = \frac{1}{2\Delta_n} (F_n(x + \Delta_n) - F_n(x - \Delta_n)),$$

где $F_n(x)$ — эмпирическая функция распределения. Чтобы $\hat{\theta}_n$ была состоятельной оценкой θ , в общем случае необходимо потребовать, чтобы Δ_n стремилось к нулю при $n \rightarrow \infty$, иначе получим оценку не точечного значения плотности, а усреднения по некоторому интервалу. Однако слишком быстрое убывание Δ_n к нулю также нарушает состоятельность: скажем, будут преобладать случаи, когда $\hat{\theta}_n = 0$. Указать допустимый порядок убывания Δ_n .

9. Пусть θ — неизвестный параметр, $\tau(\theta)$ — функция, непрерывная в точке $\theta = \theta_0$, где θ_0 — истинное значение θ . Доказать, что из состоятельности оценки $\hat{\theta}_n$ параметра θ следует состоятельность оценки $\tau(\hat{\theta}_n)$ значения его функции $\tau(\theta)$.

10. Выборочной медианой μ_n называется среднее по величине наблюдений при нечетном объеме выборки $n = 2m + 1$. Доказать, что если наблюдаемая величина имеет плотность $p(x)$ и функцию распределения $F(x)$, то плотность медианы имеет вид

$$p_m(x) = n C_{2m}^m F^m(x) (1 - F(x))^m p(x).$$

11. На основании результата задачи 10 вывести предельное распределение при $n \rightarrow \infty$ случайной величины $\zeta_n = \sqrt{n}(\mu_n - \theta)$, где θ — математическое ожидание наблюдаемой нормальной случайной величины с дисперсией σ^2 .

12. На основании результата задачи 11 найти асимптотическую эффективность μ_n как оценки параметра θ .

13. Наблюдения в точках $x = 0, 1, 2, \dots, 9$, имеют соответственно значения $y = 0,3; 2,0; 3,4; 5,0; 5,9; 6,8; 7,3; 6,6; 5,5; 4,5$. Оценить параметры a_0, a_1, a_2 приближения $y = a_0 + a_1x + a_2x^2$ по методу наименьших квадратов. Вычислить соответствующее минимальное значение z^2 по формуле (7.3).

14. Случайная величина ξ имеет функцию распределения Вейбулла:

$$F(x) = 1 - \exp\{-(\lambda x)^\alpha\}, x \geq 0.$$

Построить оценки неизвестных параметров λ и α по методу моментов.

15. Вывести распределение Стьюдента для величины t_{n-1} из примера § 9.

16. Построить доверительный интервал вида $\{\lambda \leq a(x)\}$, где x — наблюдение:

а) пуассоновской случайной величины с параметром λT ;

б) n независимых величин с плотностью $\lambda e^{-\lambda x}, x > 0$.

§ 1. М

В механике изучаются различные распределения масс. Так, в известной задаче N тел, решаемой в небесной механике, имеется N масс m_1, \dots, m_N , находящихся в данный момент в точках P_1, \dots, P_N трехмерного пространства; в теории колебаний рассматривается колебание стержня с линейной плотностью $p(x)$: масса элементарного участка между точками x и $x + dx$ равна $p(x) dx$; в теории тонких оболочек масса распределена на поверхности с некоторой поверхностной плотностью; при расчете движения жидкотопливной ракеты имеем объемную плотность массы $p(x, y, z)$ в точке (x, y, z) топливного бака.

В электростатике изучаются распределения зарядов: могут быть точечные заряды, заряды, распределенные на поверхностях диэлектриков, и т. п. Математическим понятием, объединяющим в себе свойства всевозможных распределений материи и энергии, а также таких математических понятий, как длина, площадь, является понятие *меры*.

Пусть Ω — пространство произвольной природы, т. е. любое непустое множество (в частности, числовая прямая \mathbb{R} , m -мерное пространство \mathbb{R}^m). Функция множества $\mu(A)$, где A — различные подмножества Ω , называется *мерой*, если $\mu(A) \geq 0$ и

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n) \quad (1)$$

как для конечных, так и для бесконечных последовательностей непересекающихся множеств A_n .

Условие (1) называется *аксиомой счетной аддитивности*. В примере с N точечными массами масса $\mu(A)$, попадающая в пространственное множество A , определяется формулой

$$\mu(A) = \sum_{i: M_i \in A} m_i.$$

Если A_n — непересекающиеся множества, то, очевидно, масса, попавшая в их объединение, равна сумме масс, попавших в A_n , откуда следует свойство (1). При линейном распределении массы с плотностью $p(x)$, взяв в качестве A отрезок или объединение конечного числа отрезков прямой, для массы $\mu(A)$ в этом множестве получаем

$$\mu(A) = \int_A p(x) dx;$$

свойство (1) для случая конечного числа A_n следует из того, что интеграл по сумме отрезков равен сумме интегралов по этим отрезкам.

Мера μ *сосредоточена на подмножестве* $\Omega^* \subset \Omega$, если $\mu(\Omega \setminus \Omega^*) = 0$. Мера μ называется *дискретной*, если она сосредоточена на конечном или счетном множестве (последовательности) точек $\omega_n \in \Omega$. Обозначив через p_n меру множества, состоящего из единственной точки ω_n , будем иметь

$$\mu(A) = \sum_{n: \omega_n \in A} p_n \quad (2)$$

для любого подмножества A множества Ω .

Как видим, дискретная мера определена для всех возможных подмножеств A множества Ω . В случае не дискретных мер определить $\mu(A)$ для любого A не удается. Так, если $\mu(A)$ — площадь фигуры A на плоскости, то $\mu(A)$ определена только для так называемых *квадрируемых* (т. е. имеющих определенную площадь) фигур. В случае любой меры μ требуется, чтобы $\mu(A)$ была определена для всех A из некоторой σ -алгебры (читается: сигма-алгебра) подмножеств Ω — такого класса \mathcal{F} подмножеств, что

$$\Omega \in \mathcal{F}, \quad (3)$$

$$A \in \mathcal{F} \Rightarrow \bar{A} = \Omega \setminus A \in \mathcal{F}, \quad (4)$$

$$A_n \in \mathcal{F}, n = 1, 2, \dots \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}. \quad (5)$$

Из аксиом (3) — (5), определяющих σ -алгебру, следует, что

$$A_n \in \mathcal{F}, n = 1, 2, \dots \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}. \quad (6)$$

Таким образом, с множествами, входящими в \mathcal{F} , можно производить какие угодно операции: объединения, пересечения, взятия дополнения, всякий раз получая множества из того же \mathcal{F} . Для наглядности полезно сопоставить это свойство со свойством действительных чисел: производя с ними действия сложения, умножения и т. п., всегда получаем в результате действительные числа.

Если Ω — непустое множество, \mathcal{F} — некоторая σ -алгебра его подмножеств, то пара $\langle \Omega, \mathcal{F} \rangle$ называется *измеримым пространством*, а множества из \mathcal{F} — *измеримыми множествами*. Таким образом, мера $\mu(A)$ определена лишь для измеримых множеств A . Тройка $\langle \Omega, \mathcal{F}, \mu \rangle$ называется *пространством с мерой*. Мера μ называется *конечной*, если $\mu(\Omega) < \infty$ (тогда, очевидно, $\mu(A)$ конечно и не превосходит $\mu(\Omega)$ для любого $A \in \mathcal{F}$).

Мера μ называется *нормированной*, если $\mu(\Omega) = 1$. Пример такой меры — дискретная мера, сосредоточенная на множестве точек \mathbb{Z}^+ и задаваемая формулой $p_n = 2^{-n-1}$, где p_n — мера множества, состоящего из единственной точки n . Действительно, по формуле (2) $\mu(\Omega) = \sum_{n=0}^{\infty} p_n = 1$.

Другой пример меры $\mu(A) = V(A \cap S)$, где A — фигуры в пространстве, S — фиксированная фигура единичного объема, V — символ объема. Поскольку, очевидно, $\Omega \cap S = S$, где Ω — все пространство, то $\mu(\Omega) = V(S) = 1$.

§ 2. Интеграл Лебега

В математике существует много понятий интеграла, каждое служит своим целям. Все понятия основаны на:

определении интеграла от некоторых «простейших» функций;

распространении определения на функции более сложной природы с помощью предельного перехода.

Пусть $\langle \Omega, \mathcal{F}, \mu \rangle$ — пространство с мерой. Числовая функция $f(\omega)$, $\omega \in \Omega$, называется *измеримой*, если при любом x множество тех ω , при котором $f(\omega) < x$, измеримо. (Тогда и множество тех ω , при которых $a \leq f(\omega) < b$, измеримо при любых a, b .)

Интеграл измеримой функции $f(\omega)$ по мере μ , обозначаемый $\int f(\omega) \mu(d\omega)$ ¹, определяется следующими свойствами.

1. Назовем функцию

$$I_A(\omega) = \begin{cases} 1 & \text{при } \omega \in A, \\ 0 & \text{при } \omega \notin A \end{cases}$$

¹ Вместо ω можно использовать любой другой символ, например x, y .

индикатором множества A ($A \in \mathcal{F}$). Тогда

$$\int I_A(\omega) \mu(d\omega) = \mu(A). \quad (7)$$

2. Линейность относительно функции f : если $\int f_i(\omega) \mu(d\omega)$, $1 \leq i \leq n$, определены, то определен и интеграл от линейной комбинации этих функций, причем

$$\begin{aligned} & \int (c_1 f_1(\omega) + \dots + c_n f_n(\omega)) \mu(d\omega) = \\ & = c_1 \int f_1(\omega) \mu(d\omega) + \dots + c_n \int f_n(\omega) \mu(d\omega). \end{aligned} \quad (8)$$

3. Пусть $(f_n(\omega))$ — последовательность функций, для которых интеграл определен, монотонно сходящаяся к измеримой функции $f(\omega)$. Тогда интеграл от этой последней определяется равенством

$$\int f(\omega) \mu(d\omega) = \lim_{n \rightarrow \infty} \int f_n(\omega) \mu(d\omega) \quad (9)$$

в случае, если этот предел конечен.

Интеграл, определяемый свойствами 1—3, называется *абстрактным интегралом Лебега*.

Свойство 3, выраженное также равенством

$$\int \sum_{n=1}^{\infty} f_n(\omega) \mu(d\omega) = \sum_{n=1}^{\infty} \int f_n(\omega) \mu(d\omega) \quad (10)$$

для неотрицательных $f_n(\omega)$, называется *теоремой Лебега*.

По определению

$$\int f(\omega) I_A(\omega) \mu(d\omega) = \int_A f(\omega) \mu(d\omega). \quad (11)$$

Таким образом, символ \int обозначает то же, что и \int_{Ω} .

§ 3. Основные свойства интеграла

1. *Монотонность*: при $f(\omega) \leq g(\omega)$

$$\int f(\omega) \mu(d\omega) \leq \int g(\omega) \mu(d\omega). \quad (12)$$

2. *Теорема о среднем*:

при $a \leq f(\omega)$, $\omega \in \Omega$,

$$a\mu(\Omega) \leq \int f(\omega) \mu(d\omega); \quad (13)$$

при $f(\omega) \leq b$, $\omega \in \Omega$,

$$\int f(\omega) \mu(d\omega) \leq b\mu(\Omega). \quad (14)$$

3. Интеграл от ступенчатой функции. Функция $f(\omega)$ называется *ступенчатой*, если Ω разбивается на множества A_n , на которых она постоянна: $f(\omega) = x_n$, $\omega \in A_n$. Тогда если

$$\sum_{n=1}^{\infty} |x_n| \mu(A_n) < \infty, \quad (15)$$

то интеграл от функции f существует и равен

$$\int f(\omega) \mu(d\omega) = \sum_{n=1}^{\infty} x_n \mu(A_n). \quad (16)$$

4. Для данной функции $f(\omega)$ интегралы

$$\int f(\omega) \mu(d\omega), \quad \int |f(\omega)| \mu(d\omega)$$

существуют либо не существуют одновременно. В этом свойстве имеется отличие от интеграла Римана, для которого допускается условная сходимость.

В предположении, что интеграл $\int f(\omega) \mu(d\omega)$ существует, справедливы следующие четыре утверждения.

5. Пусть A_n — измеримые множества, $A_n \subset A_{n+1}$, $\bigcup_{n=1}^{\infty} A_n = \Omega$. Тогда

$$\lim_{n \rightarrow \infty} \int_{A_n} f(\omega) \mu(d\omega) = \int f(\omega) \mu(d\omega). \quad (17)$$

6. Пусть $\mu(A_n) \xrightarrow{n \rightarrow \infty} \mu(\Omega) < \infty$. Тогда выполняется (17).

7. Пусть A_n измеримы, $A_{n+1} \subset A_n$, $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Тогда

$$\lim_{n \rightarrow \infty} \int_{A_n} f(\omega) \mu(d\omega) = 0. \quad (18)$$

8. Пусть $\mu(A_n) \xrightarrow{n \rightarrow \infty} 0$. Тогда выполняется (18).

Утверждения 5—8 называются *свойствами непрерывности интеграла*.

Пусть $f(\omega) \geq 0$, $\omega \in \Omega$, и для любого n найдется такая ступенчатая функция $f_n(\omega)$, что $f_n(\omega) < f(\omega)$, $\omega \in \Omega$, и $\int f_n(\omega) \mu(d\omega) > n$. Тогда по определению $\int f(\omega) \mu(d\omega) = \infty$.

Аналогично, если $f(\omega) \leq 0$, $f_n(\omega) \geq f(\omega)$ и $\int f_n(\omega) \mu(d\omega) < -n$ для любого n , полагаем $\int f(\omega) \mu(d\omega) = -\infty$.

Если $f(\omega)$ — функция, меняющая знак, то положим $f_+(\omega) = f(\omega)$ при $f(\omega) > 0$, $f_+(\omega) = 0$ в противном случае; $f_-(\omega) = f(\omega) - f_+(\omega)$. Интегралу функции f приписывается значение

$$\int f(\omega) \mu(d\omega) = \int f_+(\omega) \mu(d\omega) + \int f_-(\omega) \mu(d\omega), \quad (19)$$

в том числе $+\infty$ или $-\infty$, за исключением случая, когда в правой части (19) — неопределенность вида « $\infty - \infty$ ». Тем не менее, всюду в дальнейшем выражение «интеграл существует» означает, что значение интеграла конечно.

9. Замена переменной в интеграле. Пусть $\langle \Omega, \mathcal{F}, \mu \rangle$ — пространство с мерой, $\langle X, \mathcal{X} \rangle$ — измеримое пространство, $g(\omega)$ — отображение Ω в X . Это отображение называется *измеримым*, если прообраз любого множества из \mathcal{X} принадлежит \mathcal{F} . На $\langle X, \mathcal{X} \rangle$ можно определить меру $\nu(A) = \mu(A^{-1})$, $A \in \mathcal{X}$, где A^{-1} — прообраз A . Пусть далее $f(x)$ — измеримая числовая функция, заданная на X .

Справедливо равенство

$$\int_{\Omega} f(g(\omega)) \mu(d\omega) = \int_X f(x) \nu(dx) \quad (20)$$

(если один из этих интегралов существует, то существует и другой и оба совпадают).

10. Приведение кратного интеграла к повторному. Пусть $\langle \Omega, \mathcal{F}, \mu \rangle$ — пространство с мерой, причем $\Omega = X \times Y$, т. е. любое $\omega \in \Omega$ имеет вид (x, y) , где $x \in X$, $y \in Y$. Предположим, что $\langle X, \mathcal{F}_X \rangle, \langle Y, \mathcal{F}_Y \rangle$ — измеримые пространства, а \mathcal{F} — наименьшая σ -алгебра, содержащая все множества вида $\{(x, y) : x \in A, y \in B\}$, где $A \in \mathcal{F}_X, B \in \mathcal{F}_Y$. Наконец, предположим, что заданы мера μ_0 на $\langle X, \mathcal{F}_X \rangle$ и семейство мер $\{\nu_x\}$ на $\langle Y, \mathcal{F}_Y \rangle$, обладающее следующим свойством. Если $C = A \times B$, т. е. $C = \{(x, y) : x \in A, y \in B\}$, где $A \in \mathcal{F}_X, B \in \mathcal{F}_Y$, то

$$\mu(C) = \int_A \nu_x(B) \mu_0(dx). \quad (21)$$

Тогда

$$\int_{\Omega} f(\omega) \mu(d\omega) = \int_X \left\{ \int_Y f(x, y) \nu_x(dy) \right\} d\mu_0(x) \quad (22)$$

(если существует одно из этих выражений, то существует и другое и совпадает с первым).

В частности, мера μ_0 может быть взята в виде $\mu_0(A) = \mu(A \times Y)$, тогда $\nu_x(B)$ при фиксированном $B \in \mathcal{F}_Y$ определяется равенством (21) однозначно (с точностью до изменения на каком-либо множестве Λ , $\mu_0(\Lambda) = 0$) и является так называемой *мерой Радона*.

§ 4. Интеграл и мера в \mathbb{R}^m

Пусть $\Omega = \mathbb{R}^m$, \mathcal{F} — σ -алгебра борелевских множеств в \mathbb{R}^m , т. е. наименьшая σ -алгебра, содержащая все множества вида $\{(t_1, \dots, t_m) : t_1 < x_1, \dots, t_m < x_m\}$. Тогда конечная мера μ на $\langle \mathbb{R}^m, \mathcal{F} \rangle$ полностью определяется, или, как говорят, *порождается* функцией распределения меры $F(x) = F(x_1, \dots, x_m)$, которая при данных x_1, \dots, x_n равна мере множества $\{(t_1, \dots, t_n) : t_1 < x_1, \dots, t_n < x_n\}$. (Пусть, например, $m = 2$, $A = \{(t_1, t_2) : a \leq t_1 < a + h, b \leq t_2 < b + h\}$. Тогда, как легко видеть, $\mu(A) = F(a + h, b + h) - F(a + h, b) - F(a, b + h) + F(a, b)$.)

Интеграл $\int f(x) \mu(dx)$ в рассматриваемом случае обозначают

$$\int f(x) dF(x) = \int f(x_1, \dots, x_m) dF(x_1, \dots, x_m)$$

и называют *интегралом Лебега — Стильеса функции f по функции F* .

К определению интеграла f по F можно подойти и несколько иначе — аналогично тому, как определяется интеграл Римана — предел интегральных сумм. Это определение было дано Стильесом в связи с задачами механики.

Пусть задан гиперпараллелепипед $A = \{a_i \leq x_i < b_i, 1 \leq i \leq m\}$ в пространстве \mathbb{R}^m . Разобьем его на n^m равных гиперпараллелепипедов Δ_j , подобных A . Возьмем произвольное $t_j \in \Delta_j$ и составим интегральную сумму $\sum_j f(t_j) \mu(\Delta_j)$. Предел таких сумм при $n \rightarrow \infty$, если он существует и не зависит от $\{t_n\}$, называется *интегралом Римана — Стильеса* и обозначается

$$(R) \int_a^b f(x) dF(x) = (R) \int_{a_1}^{b_1} \dots \int_{a_m}^{b_m} f(x_1, \dots, x_m) dF(x_1, \dots, x_m).$$

Интеграл по всему пространству

$$(R) \int f(x) dF(x) = (R) \int \dots \int f(x_1, \dots, x_m) dF(x_1, \dots, x_m)$$

определяется как предел предыдущего при $a_i \rightarrow -\infty, b_i \rightarrow \infty$ в предположении, что существует определенный таким же образом предел для $|f(x)|$ вместо $f(x)$.

Если $f(x)$ непрерывна и ограничена, то интеграл Римана — Стильеса существует. В любом случае, когда интеграл Римана — Стильеса существует, существует также интеграл Лебега — Стильеса и значения их обоих совпадают:

$$(R) \int f(x) dF(x) = \int f(x) dF(x).$$

Рассмотрим случай $m = 1$. Пусть выполнены следующие условия:

- 1) $f(x)$ — кусочно-непрерывна; допускаются лишь разрывы первого рода;
- 2) точки разрыва $f(x)$ не совпадают с точками разрыва $F(x)$;
- 3) $f(x)$ ограничена: $|f(x)| \leq M, x \in \mathbb{R}$.

Тогда $(R) \int f(x) dF(x)$ существует.

Условие 3 можно заменить следующим:

$$f(x) \leq c_0 |x|^\alpha, \quad |x| > x_0;$$

$$F(x) \leq c_1 |x|^\beta, \quad x < -x_0;$$

$$F(\infty) - F(x) \leq c_2/x^\beta, \quad x > x_0,$$

где $\beta > \alpha$.

Мера μ называется *абсолютно непрерывной*, если существует такая функция $p(x) = p(x_1, \dots, x_m)$, что для любого множества $A = \{(x_1, \dots, x_m) : a_i \leq x_i \leq b_i, 1 \leq i \leq m\}$

$$\mu(A) = \int_A p(x) dx = \int_A \dots \int p(x_1, \dots, x_m) dx_1 \dots dx_m. \quad (23)$$

(В таком случае (23) выполняется для любого борелевского множества $A \subset \mathbb{R}^m$.) Функция $p(x) \geq 0$ называется *плотностью меры* μ .

Если μ абсолютно непрерывна, то

$$\begin{aligned} \int f(x) dF(x) &= \int_{-\infty}^{\infty} f(x) p(x) dx = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_m) p(x_1, \dots, x_m) dx_1 \dots dx_m. \end{aligned} \quad (24)$$

В общем случае этот интеграл понимается как интеграл Лебега $\int f(x) p(x) v(dx)$, где $v(A)$ — m -мерный объем множества A (лебегова мера в \mathbb{R}^m). В приложениях, как правило, (24) можно понимать как интеграл Римана,

§ 1. Первоначальные данные

Сейчас уже трудно установить, кто впервые поставил вопрос, пусть и в несовершенной форме, о возможности количественного измерения уверенности в появлении случайного события. Ясно одно, что мало-мальски удовлетворительный ответ на этот вопрос потребовал длительного времени и значительных усилий ряда поколений выдающихся исследователей. В течение долгого периода исследователи ограничивались рассмотрением разного рода игр, особенно игр в кости, поскольку их изучение позволяет ограничиваться простыми и прозрачными математическими моделями. Однако следует заметить, что многие отлично понимали то, что позднее было прекрасно сформулировано Христианом Гюйгенсом: «...я полагаю, что при внимательном изучении предмета читатель заметит, что имеет дело не только с игрой, но что здесь закладываются основы очень интересной и глубокой теории».

Мы увидим, что при дальнейшем прогрессе теории вероятностей глубокие соображения как естественнонаучного, так и общеполитического характера играли большую роль. Эта тенденция продолжается и в наши дни: мы постоянно наблюдаем, как вопросы практики — научной, производственной, оборонной — выдвигают перед теорией вероятностей новые проблемы и приводят к необходимости расширения арсенала ее идей, понятий и методов исследования.

На первом этапе изучения случайных явлений внимание ученых было сосредоточено на трех задачах: 1) подсчет числа различных возможных исходов при бросании нескольких костей; 2) раздел ставки между игроками, когда игра прервана где-то посередине; 3) определение числа бросаний двух или нескольких костей, при котором число случаев, благоприятствующих выпадению на всех костях одинаковых граней (например, «шестерок»), хотя бы при одном бросании было бы больше, чем число случаев, когда это событие не появится ни разу.

Число различных исходов при бросании трех игральных костей было определено в 960 г. епископом Виболдом из города Камбрэ. Он считал, что таких исходов 56 (он не принимал во внимание то обстоятельство, что данное число очков может появиться на любой из трех костей).

Попытка подсчитать число исходов при бросании трех игральных костей, включая и перестановки, имеется в поэме Ричарда де Форниваля (1200—1250) «De Vetula». Хотя в тексте явно указано лишь число случаев по Виболду (56), но фактически Ричард де Форниваль полностью подготовил подсчет общего числа равновероятных случаев при бросании трех костей, а именно: $6 \cdot 1 + 30 \cdot 3 + 20 \cdot 6 = 216$. Описание правильных подсчетов было дано в XI в. летописцем Балдериком, а появилось оно в печати лишь в 1615 г.

Заслуживает специального упоминания одна из первых математических книг начала эпохи итальянского Возрождения, написанная Лукой Пачоли (ок. 1445—ок. 1514) и называемая «Сумма знаний по арифметике, геометрии, отношениям и пропорциональности». Написана эта книга была в 1487 г., но издана лишь через семь лет в Венеции. Поскольку задачи Луки Пачоли сыграли определенную роль в формировании интереса к теории вероятностей, мы приведем их формулировку. В разделе «Необычные задачи» в упомянутой книге были помещены такие две задачи:

1. Компания играет в мяч до 60 очков и делает ставку в 22 дуката. В связи с некоторыми обстоятельствами игра не может быть окончена, причем одна сторона в этот момент имеет 50, другая — 30 очков. Спрашивается, какую часть общей ставки должна получить каждая сторона?

2. Трое соревнуются в стрельбе из арбалета. Кто первым достигнет 6 лучших попаданий, тот выигрывает. Ставка 10 дукатов. Когда первый получил 4 лучших

попадания, второй 3, а третий 2, они не хотят продолжать и решают разделить приз справедливо. Спрашивается, какой должна быть доля каждого?

Пачоли предложил решение, которое позднее многократно оспаривалось, поскольку оно было признано ошибочным.

§ 2. Исследования Дж. Кардано и Н. Тарталья

Несомненно, что существенное продвижение в решении первичных задач теории вероятностей связано с именами итальянских ученых Дж. Кардано (1501—1576) и Н. Тарталья (ок. 1499—1557). В рукописи «Книга об игре в кости», датированной самим Кардано 1526 г., но изданной лишь в 1563 г., были решены многие задачи, связанные с бросанием игральных костей и выпадением на них того или иного числа очков. Он правильно подсчитал числа различных случаев, которые могут произойти при бросании двух и трех костей. Словесные формулировки при этом достаточно сложны. Вот, например, что он писал в главе XI «О бросании двух костей»: «При бросании двух костей возможны 6 случаев по два одинаковых числа и 15 случаев выпадения разного числа очков, т. е., считая и двойные, 30. Следовательно, всего возможно 36 случаев». Под двойными выпадениями он понимает выпадение на двух костях очков, получаемых перестановкой. Например, двойным к случаю выпадения на первой кости двух очков, а на второй пяти будет выпадение пяти очков на первой кости и двух на второй.

Кардано указал далее число возможных случаев появления хотя бы на одной кости определенного числа очков. Таких случаев оказалось 11. Заслуживают упоминания слова Кардано: «Число это меньше, чем число случаев отсутствия данного числа очков. По отношению к общему числу случаев при бросании двух костей оно составляет больше одной шестой и меньше одной четверти». Здесь у Кардано ошибка: нужно было сказать меньше одной трети, поскольку $\frac{11}{36}$ не меньше, а больше $\frac{1}{4}$.

Это место заслуживает пристального внимания, поскольку Кардано дважды предложил рассматривать отношение, которое теперь мы называем классическим определением вероятности. А именно: $\frac{1}{6}$ это вероятность появления заданного числа очков при бросании одной кости, $\frac{11}{36}$ — вероятность получить хотя бы на одной кости грань с заданным числом очков. Означает ли это, что Кардано решил рассматривать вместо чисел благоприятствующих шансов вероятности случайных событий, т. е. ввел в рассмотрение классическое определение вероятности? Судя по всему, это было озарение только для данного примера. По-видимому, Кардано хотел выяснить вопрос: что чаще происходит — при бросании одной кости выпадение заданного числа очков или же при бросании двух костей выпадение этого числа очков хотя бы на одной кости? Ответ был найден и на этом Кардано успокоился. Единичное наблюдение он не сделал основой для общего заключения. В результате он не заметил, что стоял на пороге введения важного понятия для всего дальнейшего развития большой главы математики, да и всего количественного естествознания.

В XI главе имеется одно предложение, которое рядом авторов трактуется весьма широко, хотя, как мы сейчас увидим, его формулировка достаточно неопределенная. Вот эти слова Кардано: «Целая серия игр не дает отклонения, хотя в одной игре это может случиться... при большом числе игр оказывается, что действительность весьма приближается к этому предположению». Позднее О. Оре в своей книге¹ писал, что Кардано формулировал и использовал закон больших чисел в рудиментарной форме. Вполне возможно, что мнение Оре имеет некоторые основания, но следует заметить, что формулировка Кардано весьма неопределенная. Заметим также, что Кардано приблизился к определению безобидной игры.

Задача Пачоли о разделе ставки до окончания игры интересовала также и Кардано. В книге «Практика общей арифметики», изданной в 1539 г., Кардано привел ряд критических замечаний по вопросу решения Пачоли. Он указал на то, что Пачоли, предлагая делить ставку пропорционально числу уже выигранных партий, никак не учитывает, как много партий еще нужно выиграть каждому из игроков. Однако решение, предложенное Кардано, в общем случае ошибочно

¹ О. Ore. Cardano. The gambling scholar.— Princeton, 1953.

и лишь в некоторых весьма частных случаях оно приводит к правильному результату.

К задаче о разделе ставки вновь вернулся Н. Тарталья в книге «Общий трактат о мере и числе», которая была опубликована в 1556 г. Критические замечания Тарталья верны и имеют серьезный здравый смысл, но решение, предложенное им, также ошибочно. Следует согласиться с тем, что трудно было бы требовать от него самого и его предшественников правильного решения, поскольку в науке для этого еще не было выработано необходимых понятий — понятия вероятности и математического ожидания.

§ 3. Исследования Галилео Галилея

Мы видим, что уже в XVI в. возникли задачи чисто вероятностного характера и упорно разыскивались подходы к их решению. Это неизбежно приводило к необходимости развития, с одной стороны, комбинаторных методов, а с другой — к поиску тех понятий, терминами которых можно было бы описывать возникающие ситуации. Ошибки, допущенные одними исследователями, подмечались другими. Эти другие предлагали свои способы решения, которые, в свою очередь, подвергались критическому анализу. Постепенно вырабатывались подходы, которые позднее становились основой новой теории и, во всяком случае, позволяли решать отдельные задачи.

Заслуживает внимания вклад в этот процесс известного естествоиспытателя, ученого широких интересов и взглядов — Галилео Галилея (1564—1642). Его работа «О выходе очков при игре в кости», увидевшая свет только в 1718 г., была посвящена подсчету числа возможных случаев при бросании трех костей. Число всех возможных случаев Галилей подсчитал самым простым и естественным путем — оно возвел 6 (число различных возможностей при бросании одной кости) в третью степень и получил $6^3 = 216$, что неоднократно непосредственным подсчетом получалось и ранее.

Далее Галилей подсчитал число различных способов, которыми может быть получено то или иное значение суммы выпавших на костях очков. При подсчете Галилей пользовался полезной идеей — кости нумеровались (первая, вторая, третья) и возможные исходы записывались в виде троек чисел. Эта простая мысль для своего времени была весьма полезной.

Галилей нашел объяснение факту, обнаруженному ранее в результате наблюдения игр: значения 10 и 11 более выигрышны, чем 9 и 12, хотя числа комбинаций во всех случаях совпадают. Объяснение состоит в необходимости подсчета числа возможных перестановок той или иной комбинации очков.

Заметим, что Галилей, в сущности, повторил результаты, полученные значительно раньше рядом предшественников — епископом Виболдом, Ричардом де Форнивалем и рядом других. Однако эта, теперь такая простая для студента второго курса университета задача в ту пору была серьезным испытанием и для мыслителя столь высокого ранга как Галилей.

Заметим, что и Галилей, как и его предшественники, рассуждал не над вероятностями случайных событий, а над числами шансов, которые им благоприятствуют.

Для теории вероятностей и математической статистики большее значение, чем только что рассмотренная работа, имеют его соображения по поводу теории ошибок наблюдений. До него никто этим не занимался. Таким образом, все, что он написал на эту тему, ново для его времени и важно даже в наши дни. Свои мысли и выводы Галилей достаточно подробно изложил в одном из основных своих произведений «Диалог о двух главнейших системах мира — птоломеевой и коперниковой» (М.; Л., 1948 г.).

Согласно Галилею, ошибки наблюдений являются неизбежными спутниками каждого измерения, каждого экспериментального исследования. Он писал: «В каждой комбинации наблюдений будет какая-нибудь ошибка; я думаю, что это неизбежно...» (см. «Диалог...», с. 214). При этом ошибки могут быть двух типов: систематические, связанные прочно со способом измерений и с используемыми инструментами, и случайные, которые меняются непредсказуемым образом от одного измерения к другому. Эта классификация сохранилась до нашего времени.

Случайные ошибки измерений обладают некоторыми характерными особенностями. Их Галилей старательно выделил и проанализировал. Так, малые ошибки встречаются чаще, чем большие, поэтому, как правило, в результаты измерений следует вносить лишь небольшие поправки. Положительные ошибки встречаются так же часто, как и отрицательные. Далее Галилей отметил, что около истинного результата должно группироваться наибольшее число измерений.

Эти исследования Галилея имеют принципиальное значение, поскольку они положили начало новой научной дисциплине — теории ошибок наблюдений. Эта теория, несомненно, сыграла важную роль в формировании теории вероятностей, но еще большее значение она имела для развития математической статистики. Это тем более так, что теория «случайных ошибок» наблюдений в настоящее время рассматривается в качестве естественной задачи математической статистики.

§ 4. Вклад Б. Паскаля и П. Ферма в развитие теории вероятностей

Обычно считают, что теория вероятностей зародилась в переписке двух великих ученых — Б. Паскаля (1623—1662) и П. Ферма (1601—1665). От этой переписки сохранилось лишь три письма Паскаля (от 29 июля, 24 августа и 27 октября 1654 г.) и четыре письма Ферма (одно письмо без даты и письма от 9 августа, 29 августа, 25 сентября 1654 г.). Самое первое письмо Паскаля утеряно и о его содержании можно судить лишь по ответу Ферма.

Приведем в этой связи фрагмент одной из рукописей М. В. Остроградского (1801—1862).

«Теорию вероятностей должно отнести к наукам нового времени, ибо настоящее ее начало восходит дальше половины XVII столетия. Правда, некоторые предметы, относящиеся к этой науке, были известны во времена весьма отдаленные. Так, постоянно делали расчеты, основанные на продолжительности средней жизни, известны были морские страхования, знали число случайностей в азартных играх, но только в самых простых, найдены были величины ставок или закладов, безобидных для игроков, но подобные выводы не были подчинены никаким правилам. Однако же теорию вероятностей считают наукой нового времени и ее начало относится к первой половине XVII столетия, ибо прежде этой эпохи вопросы о вероятностях не были подчинены математическому анализу и не имелось никаких точных общих правил для решения их.

Паскаль и Ферма, геометры XVII столетия, по справедливости, считаются основателями науки о вероятностях. Первый вопрос, относящийся к этой науке, и довольно сложный, решен Паскалем. Вопрос, о котором говорим, был предложен Паскалю кавалером де Мере и состоял в следующем условии. Два игрока начали игру, состоящую из данного числа партий, положим 30-ти, розыгрыш каждой партии непременно выигрывается одним из игроков, и тот из них, кто выиграл бы прежде другого 30 партий, считался окончательно выигравшим и взял бы обе ставки, внесенные в начале игры. Но игроки согласились прекратить игру, не окончив ее, т. е. одному не хватало до выигрыша тридцати партий некоторого числа, например, трех партий, а другому, положим, пятнадцати партий. Внесенные ставки для безобидности, конечно, должны быть разделены между игроками так, чтобы тот, кому недостает до выигрыша большего числа партий, получил бы меньшую сумму, а противник его большую, именно безобидный раздел требует, чтобы каждый игрок получил часть внесенной суммы, пропорциональную вероятности своего выигрыша. Итак, нужно найти эту вероятность. Паскаль нашел ее, а потом вопрос де Мере предложил Ферма. Последний немедленно нашел решение и даже для случая более сложного, когда игра происходит не между двумя только, а между произвольным числом игроков.

Замечательно, что имя кавалера де Мере, человека светского и не имевшего никакого преуспеяния на поприще математических наук, остается навсегда в истории этих наук».

Мы увидим теперь, что оценка, данная роли Паскаля и Ферма Остроградским, несколько завышена. Впрочем, такой же точки зрения придерживаются многочисленные историки науки. Однако в переписке Паскаля с Ферма еще отсутствует понятие вероятности и оба они ограничиваются рассмотрением числа благоприятствующих событию шансов. Конечно, у этих авторов впервые в ис-

тории имеется правильное решение задачи о разделе ставки, которой, как мы знаем, прилагали много усилий исследователи в течение длительного времени. Оба они исходили из одной и той же идеи: раздела ставки в отношении пропорциональном, как мы теперь сказали бы, вероятностям окончательного выигрыша каждого игрока. В предложенных ими решениях можно увидеть начатки использования математического ожидания и в весьма несовершенной форме теорем о сложении и умножении вероятностей. Это был серьезный шаг в создании предпосылок и интересов к задачам теоретико-вероятностного характера.

Второй шаг был сделан также Паскалем, когда он существенно продвинул развитие комбинаторики и указал на ее значение для зарождающейся теории вероятностей. Результатом этого явился «Трактат об арифметическом треугольнике», опубликованный в 1665 г. и внесший серьезный вклад в развитие комбинаторики. В этом трактате имеется параграф, в котором изложены правила использования комбинаторных результатов в задаче о разделе ставки. Правило, предложенное Паскалем, состоит в следующем: пусть игроку A до выигрыша всей игры не хватает m партий, а игроку B — n партий, тогда ставка должна делиться между игроками в таком отношении:

$$\sum_{i=1}^{n-1} C_{m+n-1}^i : \sum_{i=1}^{m-1} C_{m+n-1}^i.$$

§ 5. Работа Х. Гюйгенса

Несомненно, что на развитие теории вероятностей значительное влияние оказала работа Х. Гюйгенса (1629—1695). Интерес Гюйгенса к этим вопросам был вызван его поездкой в Париж в 1655 г., где он познакомился с рядом видных ученых и услышал от них сведения относительно задач о разделе ставки в азартных играх, которые разрабатывались Паскалем и Ферма. По-видимому, ему стали известны и идеи, которыми они руководствовались при решении. Задачи заинтересовали Гюйгенса и он самостоятельно занялся размышлениями над подобными же вопросами. Поскольку ни Паскаль, ни Ферма не опубликовали разработанных ими методов, ему пришлось самому искать пути решения. Результатом явилась работа Гюйгенса, опубликованная в 1656 г. в виде дополнения к книге его учителя Ф. ван Схоутена «Математические этюды». Схоутен настолько высоко оценил эту работу Гюйгенса, что сам перевел ее на латинский язык.

Работа Гюйгенса состоит из небольшого введения и 14 предложений. Приведем первые три предложения, составляющие идейную основу всего сочинения Гюйгенса.

Предложение 1. Если я имею равные шансы получить a или b , то это мне стоит $(a + b)/2$.

Предложение 2. Если я имею равные шансы на получение a , b или c , то это мне стоит столько же, как если бы я имел $(a + b + c)/3$.

Предложение 3. Если число случаев, в которых получается сумма a , равно p , а число случаев, в которых получается сумма b , равно q , то стоимость моего ожидания равна $(ap + bq)/(p + q)$.

Для нас ясно, что этими предложениями Гюйгенс ввел понятие математического ожидания для случайной величины, принимающей два или три значения. У Гюйгенса понятие вероятности еще не выделено, и он все время оперирует с числами шансов, благоприятствующих тому или иному событию. Гюйгенс предпочел, как бы, коммерческую терминологию и говорил о стоимости, за которую он готов уступить свое право на получение выигрыша. Термин «ожидание» был введен в употребление учителем Гюйгенса Схоутеном при переводе. В той же работе Гюйгенс сформулировал пять задач, предложенных читателям для самостоятельного решения. Их решение он опубликовал лишь в 1665 г. Приведем одну из них.

A и B , каждый имеющий по 12 монет, играют тремя костями на условиях: если A выбросит 11 очков, то он должен дать B одну монету, но если он выбросит 14, тогда B должен дать одну монету A . Тот игрок выигрывает, который первым получит все монеты. Здесь шансы A относятся к шансам B как 244 140 625 к 282 429 536 581.

Эта задача является разновидностью широко исследованной в теории вероятностей задачи о разорении игрока.

Спустя десять лет после кончины известного философа Б. Спинозы (1632—1677) в Гааге была опубликована анонимная работа, состоящая из двух частей, различных по содержанию: «Исследование о радуге» и «Заметки о математической вероятности». Проведенные исследования подтверждают предположение о том, что эти сочинения были написаны Спинозой. Во второй части работы содержалось решение одной из задач Гюйгенса и были приведены формулировки остальных четырех. В названии работы уже говорится о математической вероятности, но в самой работе вероятность не определяется и рассуждения ведутся над числом благоприятствующих событию случаев.

В 1692 г. Дж. Арбутнот (1667—1735) осуществил издание английского перевода книги Гюйгенса и к этому переводу он добавил ряд новых задач, в том числе задачу иной природы. Формулировка этой задачи такова: на плоскость наудачу бросается прямоугольный параллелепипед с ребрами, находящимися в отношении $a : b : c$. Найти отношение шансов выпадения параллелепипеда гранями ab , bc и ca .

К концу XVII в. завершался длительный период накопления первичных сведений о случайных событиях, точно поставленных задач и подходов к их решению. Многие выдающиеся умы занимались этими вопросами и с разных позиций подходили к количественной оценке возможности наступления случайного события. Ферма фактически пользовался понятием математического ожидания, применение которого при решении разнообразных задач было широко развито Гюйгенсом; Паскаль, Ферма и Гюйгенс использовали представления о теоремах сложения и умножения вероятностей и подошли вплотную к понятию вероятности, однако они его не ввели. Казалось бы, что этот шаг — переход от рассмотрения числа возможных исходов, благоприятствующих наступлению события, к рассмотрению отношения этого числа к числу всех возможных исходов был естествен. Однако никто этого шага не сделал. Рассуждения, благодаря этому, были сложны и формулировки задач не очень точны. Тем не менее этого в XVII в. не произошло, и введение в науку классического понятия вероятности было сделано лишь в XVIII в. Однако оно было хорошо подготовлено исследованиями XVII в. Период предистории завершился и начинался период собственно истории теории вероятностей. Для этого уже был создан достаточно прочный фундамент.

§ 6. О первых исследованиях по демографии

В следующей главе мы узнаем, что одним из толчков для развития основных понятий теории вероятностей были исследования Джона Граунта (1620—1675) и Вильяма Петти (1623—1687) по демографии или, как тогда говорили, по политической арифметике. Их работы наглядно продемонстрировали, каким мощным орудием могут служить для изучения массовых явлений статистические наблюдения, если их соответствующим образом обработать. Их книги получили широкое распространение, старательно изучались учеными самых разнообразных направлений, в том числе и математиками.

Первой работой, с которой начинается история статистики как области научного знания, следует назвать книгу Д. Граунта, опубликованную в 1662 г. под названием «Естественные и политические наблюдения, перечисленные в прилагаемом оглавлении и сделанные над бюллетенями смертности. По отношению к управлению, религии, торговле, росту, воздуху, болезням и разным изменениям означенного города».

Основная задача, которая интересовала Граунта, состояла в указании метода, который позволял бы установить с достаточной точностью возрастной состав населения города в результате наблюдений за возрастом умерших. С этой целью им были проанализированы результаты 229 250 регистраций смерти в Лондоне, происшедших за 20 лет. Среди этих умерших было 71 124 детей возрастом от 0 до 6 лет. Причины смерти были тщательно перечислены Граунтом. Он специально отметил, что отношение числа умерших детей от 0 до 6 лет к общему числу умерших за тот же период времени составляло $71\ 224/229\ 250$, т. е. приблизительно $1/3$. Иными словами Граунт ввел представление о частоте события. Для развития

теории вероятностей это обстоятельство сыграло огромную роль, как, впрочем, и его замечание: «... мы хотели бы отметить, что некоторые из случайностей имеют постоянное отношение к числу всех похорон» (цитированная книга, с. 32). Здесь Граунт вплотную подошел к представлению о статистической устойчивости частот.

Он установил, что для Лондона число рождений мальчиков к числу рождений девочек относится как 14 : 13, что в среднем на каждые 11 семейств ежегодно умирают 3 их члена, что одна из 2000 женщин умирает от родов, что в среднем на каждые 63 покойника приходится 52 новорожденных. Тем самым численность населения Лондона пополняется систематически за счет провинции. Он установил на основании таблицы смертности, что в Лондоне на каждые 100 мужчин 34 имеют возраст от 16 до 56 лет. Так что по его данным в ту пору из 199 112 жителей мужского пола 67 694 имели возраст от 16 до 56 лет.

Им была составлена первая таблица смертностей, которую мы теперь приведем: из каждых 100 новорожденных доживает до

6 лет	64	36 лет	16	66 лет	3
16 лет	40	46 лет	10	76 лет	1
26 лет	25	56 лет	6	86 лет	0

В этой таблице поражает огромная детская и юношеская смертность: 64 % в ту пору доживали только до 6 лет и только 40 % — до 16 лет.

Граунт прекрасно понимал, что точность его выводов тем больше, чем больше наблюдений имеется для обработки.

Понятие частоты оказалось полезным и его сразу подхватили другие авторы. Так, в небольшой книге В. Петти «Два очерка по политической арифметике, относящиеся к людям, зданиям, больницам в Лондоне, Париже», вышедшей в 1682 г. в Лондоне, а через два года во французском переводе в Париже, были даны сравнительные данные о смертности в госпиталях шарите¹ Парижа и Лондона.

Несомненно, что работы Граунта, Петти и ряда из последователей представляют собой не что иное, как первые шаги в области математической статистики.

Непосредственным продолжателем исследований, начатых Граунтом и Петти, был знаменитый английский астроном Э. Галлей (1656—1742). В 1693 г. Галлей опубликовал в изданиях Лондонского королевского общества две статьи: «Оценка степеней смертности человечества, выведенная на основании любопытных таблиц рождений и погребений города Бреславля, с попыткой установить цену пожизненных рент» и «Несколько дальнейших замечаний по поводу Бреславльских бюллетеней смертности». В основу этих статей были положены данные о движении населения Бреславля за 1687—1691 гг., присланные по просьбе секретаря общества Генриха Жюстелля пастором Каспаром Нейманом. В дальнейшем Галлей к этим вопросам не возвращался.

Одна из причин интереса Галлея к таблицам смертности состоит в том, что сами Граунт и Петти признавали недостаточную обоснованность своих выводов, поскольку у них отсутствовали численность населения и возраст умерших (за частую). Кроме того, в городах, которые они изучали (Лондон и Дублин), был большой приток населения извне. Это обстоятельство делает указанные города «неподходящими в качестве стандарта для этой цели, которая требует, если это возможно, чтобы население, с которым имеют дело, было совершенно закрытым, т. е. таким, где все умирают там, где они родились, где нет никаких эмигрантов и иммигрантов» (Галлей, первый мемуар). По словам Галлея, Бреславльские материалы не имеют указанных дефектов.

На основании известных ему данных Галлей составил таблицу смертности, которую он рассматривал одновременно и как таблицу доживающих по возрасту лиц и как распределение населения по возрасту. Он ввел в науку понятие о вероятной продолжительности жизни как о возрасте, которого одинаково можно достигнуть и не достигнуть. На современном языке это медиана длительности жизни.

¹ La charité — милосердие. Так назывались больницы, организованные церковью для бедняков.

В вычислениях Галлея можно заметить использование им принципов, лежащих в основе теорем сложения и умножения вероятностей, а также рассуждения, близкие к формулировке закона больших чисел.

Работы Галлея имели очень большое значение для развития науки и применений статистических исследований о народонаселении к вопросам страхования.

Глава 2

ПЕРИОД ФОРМИРОВАНИЯ ОСНОВ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 1. Возникновение классического определения вероятности

Образование основных математических понятий представляет важные этапы в процессе математического развития. Мы видели, что до конца XVII в. наука так и не подошла к введению классического определения вероятности, а продолжала оперировать только с числом шансов, благоприятствующих тому или иному интересующему исследователя событию. Отдельные попытки, которые нами были отмечены у Кардано и у позднейших исследователей, не привели к ясному пониманию значения этого нововведения и остались инородным телом в завершенных работах. Однако в 30-х гг. XVIII в. классическое понятие вероятности стало общепотребительным и никто из ученых этих лет не мог бы ограничиться только подсчетом числа благоприятствующих событию шансов. Кто же ввел это понятие и настолько ясно показал его необходимость, чтобы в дальнейшем уже не возникло сомнений в его целесообразности для развития науки? Мы должны заметить, что введение классического определения вероятности произошло не в результате однократного действия, а заняло длительный промежуток времени, на протяжении которого происходило непрерывное совершенствование формулировки, переход от частных задач к общему случаю.

Внимательное изучение показывает, что в книге Гюйгенса «О расчетах в азартных играх» (1657) еще нет понятия вероятности как числа, заключенного между 0 и 1 и равного отношению числа благоприятствующих событию шансов, к числу всех возможных. А в трактате Я. Бернулли (1713) понятие это введено, хотя и в далеко несовершенной форме, но, что особенно важно, широко использовалось. Что же произошло за время между публикациями этих книг? Что заставило Я. Бернулли ввести в научный обиход классическое понятие вероятности?

Несомненно, что формулировка закона больших чисел, осуществленная Я. Бернулли, сама по себе является достаточным основанием для этого. Однако имеется и другое соображение, которое, несомненно, оказало сильное влияние на ход мыслей ряда исследователей, в том числе и Я. Бернулли. Речь идет о работах Граунта и Петти, о которых было сказано в гл. 1. Эти произведения решающим образом воздействовали на лучшие умы того времени, и не было ни одного мало-мальски крупного математика, который не изучал их и не находился под их воздействием. Этого влияния не избежал и Я. Бернулли. Произведения Граунта и Петти убедительно показали преимущества понятия частоты перед понятием численности. Отсюда оставался лишь один шаг до введения понятия классической вероятности. Заметим, что выводы Граунта и Петти относительно устойчивости частоты некоторых событий подготовили почву для формулировки закона больших чисел.

В несовершенной форме классическое определение вероятности появилось в первой главе четвертой части «*Argi conjectandi*» Я. Бернулли, где он отметил: «Вероятность есть степень достоверности и отличается от нее, как часть от целого». Далее было дано пояснение сказанного на примере, который отчетливо показывает, что Я. Бернулли в данную им формулировку фактически вкладывал тот же самый смысл, какой мы вкладываем в классическое определение вероятности. Вот это пояснение: «Именно, если полная и безусловная достоверность, обозначимая нами буквой α или единицей 1, будет для примера, предположена состоящей из пяти вероятностей, как бы частей, из которых три благоприятствуют существованию или осуществлению какого-либо события, остальные же не благоприятствуют, то будет сказано, что это событие имеет $\frac{3}{5}$ α или $\frac{3}{5}$ достоверности».

При формулировке главного предположения в пятой главе четвертой части Я. Бернулли вновь писал об отношении числа благоприятствующих случаев к числу всех возможных. Но при этом он не оговаривал, а предполагал само собой разумеющимся, что эти случаи должны быть равновероятными.

Интересны рассуждения четвертой главы четвертой части сочинения Бернулли. Он задал вопрос: как определить вероятность случайного события, если у нас нет возможности подсчитать числа всех возможных и благоприятствующих ему шансов? Ответ им был сформулирован следующим образом: «Но здесь нам открывается другая дорога для достижения искомого. И то, что не дано вывести *à priori*, то, по крайней мере, можно получить *à posteriori*, т. е. из многократного наблюдения результатов в подобных примерах... Ибо, если, например, при наблюдениях, сделанных некогда над тремя сотнями людей того же возраста и сложения, в каких находится теперь Тит, было замечено, что из них двести до истечения десяти лет умерли, а остальные остались в живых и дальше, то можно заключить с достаточным основанием, что имеется вдвое больше случаев Титу умереть в течение ближайшего десятилетия, чем остаться в живых по истечении этого срока... Этот опытный способ определения числа случаев по наблюдениям не нов и не необычен».

Теперь нужно подчеркнуть, что в высказанных отрывках достаточно четко прослеживается мысль о статистическом определении вероятности. Наверняка при этом Я. Бернулли основывался и на работах Граунта и Петти.

Таким образом, в трактате Я. Бернулли присутствуют обе концепции вероятности — классическая и статистическая. Обе они изложены не очень четко, но существенно то, что они уже введены в рассмотрение и использованы. Этим был сделан принципиальный шаг в науке о случае — введено в рассмотрение понятие вероятности случайного события как числа, заключенного между 0 и 1. Достоверному событию при этом приписывается максимально возможное значение вероятности единица, а невозможному — минимальное нуль. Кроме того, было отмечено, что это число может быть определено двумя различными способами: 1) путем подсчета числа равновозможных случаев, которые благоприятствуют событию и всех возможных случаев, и вычисления их отношения; 2) путем проведения большого числа независимых испытаний и вычисления частоты события. Можно считать, что теория вероятностей с этого момента начала свою историю. До этого была прелюдия, которая подготавливала почву для формирования основных понятий и задач теории вероятностей.

Я. Бернулли обдумывал свое «Искусство предположений» долгие годы — по его словам, по меньшей мере двадцать лет. Но свет оно увидело лишь в 1713 г., восемь лет спустя после смерти автора. Однако это произведение многие годы до его публикации было известно научной общественности по рукописи.

Муавр воспринял классическое определение вероятности, данное Бернулли, и вероятность события определил почти в точности так, как это делаем мы теперь. Он писал: «Следовательно, мы строим дробь, числитель которой будет число случаев появления события, а знаменатель — число всех случаев, при которых оно может появиться или не появиться, дробь будет выражать действительную вероятность его появления». Обратим внимание на то, что Муавр, как и Я. Бернулли, не отнял от обстоятельство, что шансы должны быть равновероятными. Это замечание было впервые введено в определение классической вероятности лишь Лапласом в его «Аналитической теории вероятностей».

§ 2. О формировании понятия геометрической вероятности

Уже в первой половине XVIII в. выяснилось, что классическое понятие вероятности имеет ограниченную область применения и возникают ситуации, когда оно не действует, а потому необходимо какое-то естественное его расширение. Обычно считают, что таким толчком послужили работы французского естествоиспытателя Ж. Бюффона (1707—1788), в которых он сформулировал знаменитую задачу о бросании иглы на разграфленную плоскость и предложил ее решение. Это утверждение требует поправки, поскольку исторически оно неверно. Дело в том, что задолго до рождения Бюффона появилась работа, в которой фактически уже был поставлен вопрос о нахождении геометрической вероятности. Правда, в ту пору еще не было и определения вероятности,

В 1962 г. в Лондоне был опубликован английский перевод книги Гюйгенса «О расчетах в азартных играх», выполненный Арбутнотом. В конце первой части переводчик добавил несколько задач, среди которых была сформулирована задача совсем иной природы по сравнению с теми, которые были рассмотрены великим автором. Он назвал эту задачу трудной и поместил ее в дополнении «для того, чтобы она была решена теми, кто считает такого рода проблемы достойными внимания». Задача, предложенная Арбутнотом, состоит в следующем: на плоскость наудачу бросается прямоугольный параллелепипед с ребрами, равными a , b , c . Спрашивается, как часто параллелепипед будет выпадать гранью ab .

Сам Арбутнот не сделал даже попытки решения придуманной им задачи. Это было осуществлено значительно позднее Т. Симпсоном (1710—1761) в книге «Природа и законы случая» (1740), где задача была приведена под номером XXVII.

Бюффон дважды публиковал работы, посвященные геометрическим вероятностям. Первая его публикация на эту тему относится к 1733 г., когда он сделал в Парижской академии наук доклад, напечатанный под названием «Мемуар об игре франк-карро» (мемуар об игре прямо в клетку). В 1777 г. этот мемуар был целиком включен в «Опыт нравственной арифметики», являющийся дополнением к VI тому его «Естественной истории».

Цель, которую ставил перед собой Бюффон, состояла в том, чтобы показать, что «геометрия может быть использована в качестве аналитического инструмента в области теории вероятностей».

Одна из задач, сформулированных и решенных Бюффоном, состояла в следующем: плоскость разграфлена равноотстоящими параллельными прямыми. На плоскость наудачу бросается игла. Один игрок утверждает, что игла пересечет одну из параллельных прямых, другой — что не пересечет. Определить вероятность выигрыша каждого из игроков.

Менее известна задача Бюффона об игре, когда игла бросается на плоскость, разграфленную на квадраты. В решении этой задачи Бюффон допустил ошибку, позднее исправленную Лапласом.

После Бюффона задачи на геометрические вероятности стали систематически включаться в трактаты и учебники по теории вероятностей. Так, в знаменитую книгу Лапласа «Аналитическая теория вероятностей» были включены и подробно рассмотрены все задачи Бюффона. Но Лаплас не счел нужным отметить, откуда они были заимствованы и кто автор этих задач.

Следует отметить, что терминология Лапласа далека от совершенства. Так, например, он писал, что « $8r$ равняется сумме всех случаев, в которых игла пересекает одну или другую параллельные линии» и что $2al$ равно «числу всех возможных комбинаций». Здесь $2l$ означает длину иглы, а a — расстояние между параллельными прямыми.

Серьезный шаг в развитии геометрических вероятностей связан с именами Г. Ламе (1795—1870), Барбье, Дж. Сильвестра (1814—1897), М. Крофтона, которые не только поставили новые задачи, но и привлекли к их решению понятие меры множества (пусть еще интуитивно). На базе их рассмотрений позднее возникла новая ветвь геометрии, получившая наименование интегральной геометрии.

В XIX в. на развитие проблематики геометрических вероятностей особое влияние оказал Крофтон. Он начал изучать пересечение случайными прямыми заданных выпуклых контуров. Эти его результаты вошли в курсы интегральной геометрии и монографии по геометрическим вероятностям.

На необходимость совершенствования понятия геометрической вероятности несомненное влияние оказала книга Ж. Бертрана (1822—1900) «*Calcul de probabilité*» (Paris, 1899), в которой на хорошо подобранных примерах было показано, что логические понятия геометрической вероятности не выдерживает критики. Играв на неопределенности терминологии, казалось бы, для одной и той же задачи, ему удалось получить несколько различных ответов. В качестве основной модели им была избрана известная задача о проведении наудачу хорды внутри круга. Само собой разумеется, что критика Бертрана привлекла внимание математиков к общим вопросам логического обоснования теории вероятностей.

В XX в. интерес к геометрическим вероятностям не ослабел, а вырос, поскольку, помимо чисто математического интереса, они имеют большое прикладное значение в физике, биологии, медицине, инженерном деле и др.

§ 3. Основные теоремы теории вероятностей

Мы обратимся теперь к следующему естественному вопросу: когда и кто выделил в теории вероятностей основные ее теоремы — сложения и умножения и полной вероятности? В конечном счете на этих простых результатах покоятся вся теория вероятностей и ее многочисленные применения. Зачатки этих теорем можно проследить буквально с первых шагов теории вероятностей как математической науки.

Так, в работах Паскаля можно усмотреть, что он отчетливо понимал, как следует подсчитывать число благоприятных шансов для события A , если нам известны шансы для несовместных событий A_j , составляющих событие A . Это, конечно, еще не теорема сложения, но важный шаг на пути ее формулировки. При решении задачи о разделе ставки Паскаль рассуждал следующим образом: пусть игроку A для выигрыша игры недостает трех партий, а игроку B — четырех. Тогда для завершения игры достаточно шести партий. Игрок A выигрывает, если из этих шести партий он выигрывает все шесть партий, пять, четыре или три партии. Таким образом, число благоприятных шансов для выигрыша игрока A оказывается равным

$$C_6^6 + C_6^5 + C_6^4 + C_6^3 = 1 + 6 + 15 + 20 = 42.$$

Это рассуждение Паскаль предложил и в общем случае, когда для окончания игры игроку A недостает m партий, а игроку B — n партий.

В работах Я. Бернулли и Н. Бернулли (1687—1759) дается отчетливая формулировка правила вычисления вероятности противоположного события, если известна вероятность прямого события.

При выводе формул, получивших наименование формул Бернулли, Я. Бернулли сознательно использовал правила сложения и умножения вероятностей, но самих правил он не сформулировал. Они в его рассуждениях присутствуют как бы неявно. Я. Бернулли фактически находился рядом с предложением, которое мы теперь записываем в виде

$$P(A + B) = P(A) + P(B) - P(AB),$$

но явно его не сформулировал.

Впрочем, если с современных позиций рассматривать работу Кардано «Книга об игре в кости», то в главе XIV «О соединении очков» можно в частном примере усмотреть этот же результат, но не для вероятностей, а для числа шансов. Тем не менее, нет оснований приписывать ни Кардано, ни Паскалю и Ферма, ни Я. Бернулли формулировку теоремы сложения вероятностей как важнейшего положения теории вероятностей.

Первая четкая и окончательная формулировка теоремы сложения вероятностей находится в работе Т. Байеса (1702—1761), зачитанной на заседании Лондонского королевского общества 27 декабря 1763 г., спустя два года после смерти автора. В этой работе содержится определение несовместных событий. Байес употребляет другой термин — «неплотные события» (inconsistent). Согласно Байесу, «несколько событий являются неплотными, если наступление одного из них исключает наступление других». Формулировка же теоремы сложения дается Байесом в предложении, состоящем в следующем: «Если несколько событий являются неплотными, то вероятность того, что наступит какое-то из них, равна сумме вероятностей каждого из них».

Точно так же теорема умножения вероятностей длительный период формировалась на рассмотрении частных примеров и на подсчете числа шансов, благоприятствующих наступлению произведения двух или нескольких событий. Такого рода подсчеты встречаются практически у всех предшественников Я. Бернулли. Я. Бернулли широко использует эти правила при выводе своих знаменитых формул. Широко использовал правила сложения и умножения вероятностей Монмор. Однако формулировка теоремы умножения ни у кого из них не встречается. Четкое выделение теоремы умножения было осуществлено лишь А. Муавром (1667—1754). Во введении к «Доктрине шансов» в § 8 он определил важное понятие независимости случайных событий. Именно, он формулирует следующее положение: «Мы скажем, что два события независимы, когда каждое из них не имеет никакого отношения к другому, а появление одного из них не оказывает никакого влияния на появление другого». Еще более определенно им дано определение зависимых

событий. А именно: «Два события зависимы, когда они связаны друг с другом и когда вероятность появления одного из них изменяется при появлении другого».

Приведем следующую формулировку Муавра: «...вероятность появления двух зависимых событий равна произведению вероятности появления одного из них на вероятность того, что другое должно появиться, если первое из них уже появилось. Это правило может быть обобщено на случай нескольких событий».

Таким образом, формулировка теоремы умножения вероятностей и введение понятий условной вероятности удалось осуществить только Муавру. Это им было сделано уже в 1718 г. в первом издании его «Доктрины шансов».

В упомянутой ранее работе Байеса содержится формулировка теоремы умножения вероятностей для зависимых событий. Мы уже видели, что это предложение было четко сформулировано Муавром и, поскольку произведение Муавра было широко известно, Байес, несомненно, заимствовал его у своего знаменитого предшественника. Единственно в чем Байес пошел дальше Муавра — это в формулировке следствия в вычислении вероятности $P(B|A)$ по вероятностям $P(AB)$ и $P(A)$. Это предложение дало основание приписывать Байесу формулы, называемые его именем. В действительности у него их нет, поскольку он не знал формулы полной вероятности.

Результат, приписываемый Байесу, по-видимому, впервые получил современную формулировку в работе Лапласа «Опыт философии теории вероятностей». В главе «Общие принципы теории вероятностей» Лаплас сформулировал принцип VI, который относится к вероятности гипотез или, как он писал, вероятности причин. Вот формулировка, данная Лапласом: «Вероятность существования какой-либо из этих причин равна, следовательно, дроби, числитель которой есть вероятность события, вытекающая из этой причины, а знаменатель есть сумма подобных вероятностей, относящихся ко всем причинам; если эти различные причины, рассматриваемые а priori, не одинаково вероятны, то вместо вероятности события, вытекающей из каждой причины, следует взять произведение этой вероятности на вероятность самой причины». Таким образом, Лаплас словесно сформулировал известное «правило Байеса». Более того, этот принцип Лапласа содержит и формулу полной вероятности, которой с начала XVIII в. широко пользовались в своих работах многочисленные математики, работающие в области теории вероятностей. Они понимали, как использовать принцип, заложенный в формуле полной вероятности, но его не формулировали.

Мы видим, таким образом, что основные принципы действия с вероятностями вычленились длительным путем. Их многократно использовали при решении отдельных задач и использовали правильно, но не формулировали в качестве особых предложений. И потребовалось почти целое столетие, чтобы после введения в науку понятия вероятности сформулировать для этого понятия систему правил действия с ним. Как постоянно происходит в истории науки, такие правила широко использовались фактически, но потребности в их формулировке не ощущалось. Попутно при этом вводились и дополнительные понятия, которые позволяли глубже вникать в природу вещей. В нашем случае этими понятиями являются понятия несовместности и независимости случайных событий.

§ 4. Задача о разорении игрока

Несомненно, что задача о разорении игрока в развитии теории вероятностей играла серьезную роль — она позволяла оттачивать методы решения сложных вопросов и в какой-то мере являлась исходным пунктом развития теории случайных процессов. Действительно, именно в этой задаче впервые начали изучать состояние системы в зависимости от времени. Точнее — положение игроков после заданного числа партий. Задача о разорении игрока была впервые сформулирована Гюйгеном в книге «О расчетах в азартных играх» (см. § 5 первой главы настоящего очерка). Этой задачей занимались многие выдающиеся математики прошлого — Я. Бернулли, Н. Бернулли, А. Муавр, П. Лаплас и др.

Первые подходы к решению задачи о разорении игрока почти одновременно были предложены тремя математиками — П. Монмором (1678—1719), А. Муавром и Н. Бернулли. Их результаты относились к 1710—1711 гг. Задача Гюйгенса в их формулировке слегка преобразовалась и приобрела привычный для нас вид: иг-

роки A и B имеют соответственно a и b франков и при каждой партии некоторой игры один из них выигрывает у другого 1 франк. Вероятность выигрыша игрока A для каждой партии равна p , для игрока B вероятность выигрыша равна $q = 1 - p$. Спрашивается, чему равны вероятности p_a и p_b того, что игрок A (соответственно игрок B) выигрывает игру (т. е. игрок A выигрывает все деньги у игрока B раньше, чем B выиграет их у A).

Муавр опубликовал свои результаты в журнале *Philosophical Transactions* за 1711 г. Он нашел, что

$$p_a = \frac{\left(\frac{q}{p}\right)^a - 1}{\left(\frac{q}{p}\right)^{a+b} - 1}, \quad p_b = \frac{\left(\frac{p}{q}\right)^b - 1}{\left(\frac{p}{q}\right)^{a+b} - 1}$$

и что математическое ожидание числа N необходимых для завершения игры партий равно

$$\frac{bp_a - ap_b}{p - q}.$$

Ему же удалось найти вероятности p_{an} и p_{bn} того, что игрок A выиграет игру за n партий (соответственно выиграет игру за n партий игрок B).

Вдобавок им был подробно рассмотрен случай, когда $a = \infty$.

В 1710 г. формулы для p_{an} , p_{bn} в случае $p = q$ нашел Монмор. Свои соображения он переслал Иоганну Бернулли, который передал письмо своему племяннику Николаю. Ответное письмо Николая Бернулли от 26.II.1711 г. содержало решение и для случая $p \neq q$. Это письмо Монмор опубликовал в 1713 г. в трактате «Опыт анализа азартных игр».

Я. Бернулли рассматривал задачу о разорении игрока как в частном (для $a = b = 2$), так и в общем случаях. При ее решении он следовал методу Гюйгенса и получил довольно далеко идущие результаты (для вероятностей p_a и p_b).

Рассмотрение решений, предложенных Я. Бернулли, Н. Бернулли, Монмором и Муавром, ясно показывает, что все они владели приемами оперирования с вероятностями сложных событий. Практически они безукоризненно точно использовали теоремы сложения и умножения вероятностей, а также формулу полной вероятности, хотя в ту пору они еще не получили четкой формулировки. Происходило накопление опыта и выделение тех правил, которые постоянно необходимы при подсчете вероятностей сложных событий.

§ 5. Возникновение предельных теорем теории вероятностей

На последующее развитие теории вероятностей огромное воздействие оказала идея, впервые высказанная и осуществленная Я. Бернулли — рассматривать не только точные решения задач теории вероятностей, но и их асимптотические постановки при неограниченном увеличении некоторого параметра. Конечно, прежде всего следует указать в этом плане на закон больших чисел в форме Бернулли. Именно он послужил источником для различного рода уточнений как в XVIII, так и в последующие столетия.

Я. Бернулли дал формулировку своей теоремы в отличном от принятого теперь виде. Мы приведем его формулировку ниже. Сейчас же отметим, что и принятая им терминология отлична от современной и связана с демографией. Так, Я. Бернулли использовал для обозначения испытаний, при которых интересующее нас событие происходит, слова «плодовитый», «фертильный», а для противоположных исходов — слово «стерильный». Теперь мы можем перейти к оригинальной формулировке теоремы Я. Бернулли, которую он ценил и вынашивал, по его словам, свыше двадцати лет.

«Пусть число фертильных случаев к числу стерильных случаев относится точно или приближенно как $\frac{r}{s}$, или же это число относится к числу всех случаев как

$\frac{r}{r+s}$, или же как $\frac{r}{t}$. Последнее отношение находится, следовательно, между $\frac{r-1}{t}$ и $\frac{r+1}{t}$. Нужно доказать, что можно произвести столь большое число опытов, что отношение числа появившихся фертильных наблюдений к числу всех опытов будет больше, чем $\frac{r-1}{t}$, и меньше, чем $\frac{r+1}{t}$.

Мы уже говорили, что книга «Искусство предположений» Я. Бернулли была широко известна многим математикам задолго до ее публикации. В частности, она была тщательно изучена его племянником Н. Бернулли, который в 1709 г. защитил диссертацию для получения ученой степени лиценциата прав под названием «О применении искусства предположений в вопросах прав». Во второй главе «О способе установления вероятности человеческой жизни», исходя из таблиц Граунта, он изучил вопрос о вероятности дожития до определенного возраста. Интересно, что он отметил факт, подмеченный из изучения долголетних регистраций рождений, что мальчиков рождается больше, чем девочек. При этом отношение числа рождений мальчиков к числу рождений девочек оказывается, как он считал, равным 18 : 17. Подробное изучение содержания этой главы показывает, что Н. Бернулли принимал вероятность рождения мальчика равной $p_m = 18 : 35 \approx 0,514$ и соответственно вероятность рождения девочки равной $p_d = 17 : 35 \approx 0,486$.

Далее он рассмотрел пример, когда имеется 14 000 рождений. Тогда, согласно формулам Я. Бернулли, имеет место равенство (μ — означает фактическое число рождений мальчиков)

$$P(|\mu - 7200| \leq 163) = P(7037 \leq \mu \leq 7363) = \sum_{i=7037}^{7363} C_{14\ 000}^i p^i q^{14\ 000-i}.$$

Фактическое число рождений мальчиков зависит от случая. Приведенная формула позволяет вычислять вероятность того, что число рождений мальчиков будет заключено в указанных границах. Однако вычисления, которые при этом необходимо произвести, сложны.

В двух последних изданиях книги Муавра «Доктрина шансов» был помещен перевод его статьи 1733 г. «*Approximatio ad Summum terminum Binomii $(a+b)^n$ in seriem expensis*». Согласно словам самого автора «Я помещаю здесь перевод моей работы, написанной 12 ноября 1733 года и сообщенной некоторым друзьям, но никогда не публиковавшейся» («Доктрина шансов», 1756, с. 242). В кратком введении Муавр отметил, что для решения ряда задач теории вероятностей необходимо подсчитывать суммы $\sum_{m=l}^k P_n(m)$ членов биномиального распределения и что

вычисления становятся громоздкими при больших значениях числа испытаний n . В результате перед Муавром возник вопрос о разыскании асимптотической формулы. Эта задача была им успешно решена. Основная трудность, которая при этом возникла, состояла в оценке факториала $m!$ при больших значениях m . Муавру удалось доказать, что имеет место асимптотическое равенство $m! \approx B \sqrt{m} e^{-m} m^m$, где B — постоянное. При этом оказалось, что

$$\ln B = 1 - 1/12 + 1/360 - 1/1260 + 1/1680 - \dots$$

Муавр нашел, что $B \approx 2,5074$, однако это его не удовлетворило и ему хотелось связать эту константу с ранее введенными в математику. Он обратился со своей проблемой к Д. Стирлингу (1692—1770). Стирлинг с успехом разрешил этот вопрос и показал, что $B = \sqrt{2\pi} \approx 2,506628\dots$. В связи с этим хотелось бы отметить, что известную формулу Стирлинга для приближенного вычисления фактори-

¹ Я. Бернулли счел излишним говорить, что $t = r + s$. Заметим также, что r, s и t не фиксированы, а могут принимать любое значение, лишь бы отношение $\frac{r}{t}$ имело заданное значение. Отсюда следует, что отношение $\frac{1}{t}$ может быть сделано как угодно малым.

ала в случае больших чисел следовало бы называть точнее формулой Муавра или самое меньшее — формулой Муавра — Стирлинга. Заметим, что Муавр впервые вычислил и опубликовал таблицу функций $\ln n!$ для значений n от 10 до 900.

Используя найденную им формулу «Стирлинга», Муавр первоначально выяснил, что в случае $p = q = 0,5$ средний член бинома $(\frac{1}{2} + \frac{1}{2})^n$ асимптотически равен $1/\sqrt{2\pi npq}$, а затем доказал локальную теорему, носящую теперь его имя («Доктрина шансов», с. 243—244). То, что Муавр начал со случая $p = 0,5$, вполне естественно, поскольку именно этот случай играет значительную роль в простейших задачах демографии. Далее Муавр получил локальную теорему для $p \neq 0,5$ фактически в принятом теперь виде.

Имея в руках локальную теорему, Муавр без затруднений сформулировал и интегральную теорему, правда, только для симметричных границ. Впрочем, интегральная теорема, доказанная для симметричных границ, без труда распространяется и на общий случай. Он оценил важность выражения \sqrt{npq} для теории и предложил для него специальное наименование — модуль.

Используя метод приближенного интегрирования Ньютона—Котса, Муавр вычислил для случая $p = q = 0,5$ предел вероятности

$$P\left(\frac{1}{2}n - \sqrt{n} < \mu < \frac{1}{2}n + \sqrt{n}\right).$$

Согласно его подсчетам, он оказался равным 0,95428. Теперь, используя таблицы, несложно проверить его расчеты и убедиться в том, что допущенная им ошибка невелика, только в четвертой значащей цифре (табличное значение равно 0,95450).

Муавр отметил, что интегральную теорему можно использовать и для оценки неизвестной вероятности p , т. е. для решения обратной задачи — задачи математической статистики.

§ 6. Контроль качества продукции

В связи с переходом промышленности на массовое изготовление изделий за последние пятьдесят — шестьдесят лет резко увеличился интерес к вопросам проверки качества изделий, входящих в принимаемую партию. Появилась глубокая по содержанию и значительная по своим практическим применениям теория статистических методов приемочного контроля, основанная на широком использовании теории вероятностей.

Первым шагом, относящимся к этому кругу идей, по-видимому, следует считать одну из задач, рассмотренных Т. Симпсоном в книге «Природа и законы случая» (1740). Вот формулировка этой задачи: имеется данное число вещей различного сорта — n_1 вещей первого, n_2 — второго, ... Наудачу берутся m вещей. Найти вероятность того, что при этом будет взято m_1 вещей первого сорта, m_2 вещей второго и т. д.

Спустя сто с небольшим лет к этой задаче вновь вернулся Остроградский в работе «Об одном вопросе, касающемся вероятностей» (1846). В математическом отношении это произведение Остроградского не представляет большого интереса, но глубокое понимание самой практической задачи заслуживает нашего внимания. По-видимому, в этом отношении он имеет приоритет перед всеми исследователями. Во всяком случае Симпсон практических следствий своих подсчетов не делал, а Остроградский вычислил и таблицы, необходимые для практических применений.

Статистические методы приемочного контроля получили особенно бурное развитие в годы второй мировой войны, поскольку было необходимо принимать огромные партии однородной продукции, а проверять ее сплошь не было возможности по ряду причин, из которых укажем лишь на следующие: для некоторых видов продукции проверка равносильна уничтожению рабочих свойств изделия (фотобумага, взрыватели и т. д.); для сплошной проверки требовалось такое количество рабочей силы и рабочего времени, что ни то, ни другое не могло быть обеспечено в условиях военного времени. Нет возможности здесь перечислить даже основные этапы развития теории статистических методов приемочного контроля. Большинство исследователей работали над различными проблемами этой тематики и внесли в ее развитие значительный вклад. Из отечественных ученых заслуживают

быть отмеченными А. Н. Колмогоров (1903—1987), В. И. Романовский (1879—1954), С. Х. Сираждинов (1921—1988), Ю. К. Беляев (р. 1932) и др.

В период Великой Отечественной войны 1941—1945 гг. огромное внимание было уделено разработке методов управления качеством продукции в процессе производства. Это весьма важное направление работы, поскольку оно должно не просто разбраковывать изготовленную продукцию, но своевременно вмешиваться в производственный процесс и не допускать изготовления некачественной продукции. Именно к этому и должно стремиться любое производство. Итак, для управления качеством нужно разработать такие методы, которые позволяли бы установить тот момент, когда бракованная продукция еще не производится, но возникает опасность начала ее производства.

Идея статистического метода управления качеством состоит в том, чтобы время от времени проверять небольшие партии продукции (5—10 штук), только что сошедших со станка. По результатам таких проверок судят о качестве наладки. Эти проверки осуществляются не очень часто, чтобы не лихорадить переналадками оборотования производственный процесс, и не очень редко, чтобы не пропустить момент его разладки. Далее результаты наблюдений наносятся на так называемые контрольные карты, которые позволяют судить, что следует предпринимать после каждой серии таких наблюдений — прекращать работу для переналадки оборудования или продолжать производственный процесс.

Методы приемочного контроля и статистические методы управления качеством оказались весьма эффективными средствами упорядочения производства и экономии стачного времени, материалов, рабочей силы. Экономический эффект от использования этих методов исчисляется миллиардами рублей (долларов, марок и т. д.).

Широта применения теории вероятностей в начале двадцатого века настойчиво требовала глубокого осмысления самих основ этой науки. Наивные представления о случайном событии и его вероятности уже перестали удовлетворять запросы науки. Вот почему ряд ученых — Э. Борель, С. Н. Бернштейн, Р. Миэс и другие предприняли серьезные усилия для просмотра основ теории вероятностей. Наиболее плодотворный подход был предложен А. Н. Колмогоровым в статье 1927 г. и нашел завершение в его монографии «Основные понятия теории вероятностей» (1933).

Глава 3

К ИСТОРИИ ФОРМИРОВАНИЯ ПОНЯТИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

§ 1. Развитие теории ошибок наблюдений

Мы уже упоминали в гл. 1 о том, что Галилео Галилей заложил основы теории ошибок измерений и ввел в рассмотрение ряд важных понятий, которые сохранили значение и в наши дни.

Позднее под влиянием прежде всего астрономических и геодезических наблюдений интерес к ошибкам измерений заметно возрос. Знаменитый астроном-наблюдатель Тихо Браге (1546—1601) обратил внимание на то, что каждое отдельное измерение несет в себе возможную ошибку и точность измерений значительно повышается, если произвести несколько измерений и взять из них среднее арифметическое. Впрочем рекомендация пользоваться средним арифметическим для уточнения размеров была дана задолго до Тихо Браге. Так, согласно литературным данным, в одном древнеиндийском математическом трактате рекомендовалось при подсчете объема земляных работ делать измерения в нескольких местах и затем оперировать со средними арифметическими (см. Л. Е. Майстров, «Развитие понятия вероятности», М., с. 74).

Казалось бы от И. Кеплера (1571—1630), сделавшего так много для формирования законов движения планет, следовало ожидать повышенного внимания к методам обработки результатов наблюдений. Но эти вопросы фактически остались в стороне от его интересов и он заметил только то, что хороший наблюдатель приводит измерения с ошибками ограниченной величины.

Первые попытки построить математическую теорию ошибок измерений принадлежат Р. Котсу (1682—1716), Т. Симпсону и Д. Бернулли (1700—1782).

Следует отметить работы И. Ламберта (1728—1777) 1760 г. и 1765 г., в которых излагаются цели теории ошибок измерений (кстати, ему принадлежит и сам этот термин), свойства ошибок, оценку точности наблюдений и правила подбора кривых по наблюдаемым точкам, содержащим случайные ошибки. Позднее появилась работа Ж. Лагранжа (1736—1813), посвященная выяснению роли среднего арифметического при оценке истинного значения измеряемой величины.

Для П. Лапласа (1749—1827) теория вероятностей была не столько математической, сколько естественнонаучной дисциплиной. Поскольку он занимался астрономией, то неизбежно должен был прийти к вопросам теории ошибок наблюдений и вместе с тем заинтересоваться теорией вероятностей. Лаплас получил ряд важных результатов в теории ошибок измерений, которые вошли в практику обработки данных наблюдений. Особый интерес представляют две идеи Лапласа. Первая из них вызвала к жизни значительное увеличение интереса к предельным теоремам для сумм независимых случайных величин. Именно, согласно Лапласу, наблюдаемые ошибки являются результатом суммирования очень большого числа элементарных ошибок. Если эти ошибки равномерно малы, то Лаплас предположил, что их распределение должно быть близко к нормальному.

Вторая идея касается оценки измеряемой величины по результатам измерений x_1, x_2, \dots, x_n . В качестве оценки неизвестного значения a измеряемой величины Лаплас предложил брать то значение $\hat{a} = \hat{a}(x_1, x_2, \dots, x_n)$, при котором обращается в минимум сумма $\sum_{i=1}^n |x_i - \hat{a}|$.

Оказывается, что \hat{a} равняется эмпирической медиане, т. е. тому значению аргумента, слева и справа от которого расположено одинаковое число наблюдённых значений. Этот метод не получил в ту пору распространения, поскольку вскоре был предложен другой метод, приводящий к более простым результатам, разработка которого принадлежит К. Гауссу (1777—1855), А. Лежандру (1752—1833) и американскому математику Р. Эдрейну (1775—1843). Их работы составили в теории ошибок наблюдений настоящую эпоху. Гауссом и Лежандром был предложен и разработан метод наименьших квадратов. Гаусс предложил его во второй книге большого трактата «Теория движения небесных тел, вращающихся вокруг Солнца по коническим сечениям» (1809). Лежандр же изложил свои идеи в работе «Новые методы для определения орбит комет» (1806), к которой было сделано специальное дополнение «О методе наименьших квадратов». Гаусс неоднократно писал, что он пользовался этим методом начиная с 1795 г. Гауссом и Эдрейном было показано, что при некоторых, весьма широких условиях плотность ошибок измерений имеет вид

$$\frac{h}{\sqrt{\pi}} \exp \{-h^2 x^2\}.$$

Необходимо отметить, что влияние Гаусса, Лежандра и Эдреяна на развитие науки оказалось весьма различным. Статья Эдреяна, опубликованная в мало распространенном американском журнале, прошла практически незамеченной. Работы же Гаусса и Лежандра почти мгновенно стали известны научному миру и ученые восприняли предложенный ими метод и начали им пользоваться в практической работе.

Большой вклад в дальнейшее развитие этой теории внес С. Пуассон (1781—1840). В частности, Пуассон задался вопросом: всегда ли среднее арифметическое дает лучший результат по сравнению с отдельным наблюдением? Ответ оказался отрицательным. Именно ему удалось указать распределение, для которого это правило ошибочно. Плотность этого распределения

$$p(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

Пуассон обнаружил, что сумма двух независимых случайных величин с такой плотностью распределения имеет с точностью до масштаба такое же распределение.

Далее он показал, что среднее арифметическое из независимых наблюдений над такой случайной величиной имеет в точности такое же распределение. Через двадцать лет (в 1853 г.) О. Коши (1789—1857) повторил этот результат, после чего указанное распределение получило название распределения Коши, а работа Пуассона была забыта.

Позднее теория ошибок измерений привлекла внимание практически всех выдающихся специалистов в области теории вероятностей. П. Л. Чебышев (1821—1894), А. А. Марков (1856—1922) и многие другие уделяли внимание как методу наименьших квадратов, так и другим вопросам теории ошибок. Теория ошибок оказала серьезное влияние на постановки задач и разработку методов математической статистики. Теперь же теория ошибок включается в качестве естественной части в математическую статистику.

§ 2. Формирование понятия случайной величины

Мы неоднократно отмечали, что формирование научных понятий проходит длительный и сложный путь, прежде чем войти во всеобщее употребление. Как правило, необходимое понятие еще не введено в научный обиход, а фактически им уже пользуются как при решении практических задач, так и при выводе общетеоретических закономерностей. Этот путь характерен и для случайной величины — основного понятия теории вероятностей и современного естествознания. Введение этого понятия связано с именами многих ученых, которые хотя и не использовали, но фактически исследовали отдельные его свойства.

Действительно, мы уже знаем, что, начиная с Котса, Симпсона, Д. Бернулли, в XVIII в. начала развиваться теория ошибок наблюдений, возникшая прежде всего под влиянием астрономии. Ошибка измерения в зависимости от случая может принимать различные значения. Эта позиция была высказана Галилеем задолго до работ выше упомянутых ученых. Он ввел в обиход термины «случайная» и «систематическая ошибка» измерения. Вторая — тесно связана с качеством изготовления прибора, мастерством наблюдателя, условиями наблюдений. Первая — зависит от многочисленных причин, влияние которых невозможно учесть и которые изменяются от наблюдения к наблюдению, от измерения к измерению. Теперь мы ясно видим, что ошибка измерения представляет собой случайную величину с каким-то неизвестным нам распределением вероятностей.

Но с понятием случайной величины встречались уже Я. Бернулли, Н. Бернулли, Монмор, Муавр. В самом деле, Я. Бернулли рассмотрел число появлений интересующего его события A в n независимых испытаниях. Для нас теперь это случайная величина, способная принимать значения $0, 1, 2, \dots, n$ с вероятностями, задаваемыми формулами Бернулли. Н. Бернулли, Монмор и Муавр, исследуя задачу о разорении игрока, также имели дело со случайной величиной — числом партий, которые необходимы для разорения одного из них. Муавр пошел еще дальше — он ввел в рассмотрение нормальное распределение вероятностей. Однако никто из перечисленных ученых не заметил, что в науку властно постучалась необходимость введения нового понятия — случайной величины. Первый из них оставался на уровне схемы последовательности случайных событий, остальные ограничились той частной задачей, которая перед ними стояла. Для Муавра нормальное распределение было лишь аппроксимирующей функцией, дающей хорошее приближение к точному значению искомых вероятностей.

Первоначально считалось, что возможные значения ошибок измерений составляют арифметическую прогрессию с неопределенной, но очень малой разностью. Затем постепенно от этого предположения отказались и стали представлять себе, что возможные значения, принимаемые ошибками наблюдений, заполняют целый отрезок. Вероятности возможных значений определялись посредством плотности распределения. И если Д. Бернулли в отношении плотности распределения вероятностей допускал еще определенные вольности, то у Лапласа, Гаусса, Лежандра с плотностью распределения было уже все в порядке. Это была неотрицательная функция, интеграл от которой по всей прямой равен единице, а вероятность попадания в тот или иной отрезок равнялся интегралу от плотности, взятому по этому отрезку. Лапласу уже была известна формула для нахождения плотности распределения суммы по плотностям распределения слагаемых. В знаменитой

книге «Аналитическая теория вероятностей» Лаплас умело оперирует с плотностями распределения, ставит и решает ряд интересных задач, но нигде не вводит понятия случайной величины. Он либо использует язык теории ошибок измерений, либо язык математического анализа и не ощущает потребности в новом понятии теории вероятностей.

Первая половина XIX в. принесла новые задачи, которые нуждаются в понятии случайной величины. Прежде всего — это исследования бельгийского естествоиспытателя Л. А. Кетле (1796—1874), заметившего, что размеры органов животных определенного возраста подчиняются нормальному распределению. Изучение уклонений снаряда от цели явилось предметом исследования многих ученых; они также привели к выводу о нормальном распределении этой величины. С середины XIX в. начались замечательные исследования Д. К. Максвелла (1831—1879) и ряда других ученых по математической теории молекулярной физики газов. И здесь снова нормальное распределение завоевало почетное место.

Заслуживает внимания постановка еще одной задачи Гауссом. Он сформулировал ее 25 октября 1800 г. Именно за этот день в его дневнике под № 113 сделана соответствующая запись. Через двенадцать лет он так сформулировал ее в письме к Лапласу от 30 января 1812 г.: «Пусть M — неизвестная величина, заключенная между пределами 0 и 1, для которой все значения или одинаково вероятны, или же более или менее следуют данному закону: предположено, что она разложена в непрерывную дробь $M = 1/a^{(1)} + 1/a^{(2)} + \dots$. Чему равна вероятность того, что, отбросив в разложении конечное число членов до $a^{(n)}$, следующая дробь $1/a^{(n)} + 1/a^{(n+1)} + \dots$ будет заключена в пределах от 0 до x ? Я обозначаю ее через $P(n, x)$ и предполагаю, что для M все значения одинаково вероятны: $P(0, x) = x$ ».

Гипотеза Гаусса состояла в том, что

$$\lim_{n \rightarrow \infty} P(n, x) = \log(1+x)/\log 2.$$

Эта задача относится к интересному, начавшему развиваться лишь в нашем веке разделу математики — метрической теории чисел; одновременно она имеет самое непосредственное отношение к изучению равномерно распределенных случайных величин. Решение этой задачи появилось лишь в 1928 г., его дал Р. О. Кузьмин (1891—1949). Через год П. Леви (1886—1971) дал для этой задачи чисто вероятностное решение, получив для быстроты сходимости к пределу лучшую, чем у Кузьмина, оценку. Позднее было доказано, что результат сохраняется для любой случайной величины M , для которой $P(0, x)$ имеет ограниченную производную. Это замечание делает более ясными неопределенные слова Гаусса о том, что для величины M «все значения или одинаково вероятны, или же более или менее следуют данному закону».

Заслуживает внимания то обстоятельство, что функция $P(0, x)$, так же, как и $P(n, x)$, представляет собой функцию распределения.

Мы видим, что многочисленные исследования многих великих математиков достаточно подготовили почву для введения понятия случайной величины. Казалось бы для этого все условия уже созрели. По-видимому, этот шаг впервые был сделан Пуассоном в мемуаре 1832 г. «О вероятности средних результатов наблюдений». Этот факт мне сообщил О. Б. Шейнин. Термина «случайная величина» у Пуассона еще нет, но он пишет о «некоторой вещи», которая способна принять значения $a_1, a_2, \dots, a_\lambda$ соответственно с вероятностями $p_1, p_2, \dots, p_\lambda$. Он рассмотрел также непрерывные случайные величины и их плотности распределения.

Итак, Пуассоном был сделан важный шаг в науке — он ввел в научный обиход новое понятие — случайную величину. Его первоначальный термин «вещь» не прижился и вскоре перестал употребляться. Чебышев в своих мемуарах по теории вероятностей уже использует термин «величина» и автоматически считает все случайные величины, с которыми имеет дело, независимыми. В работах А. М. Ляпунова (1857—1918) по теории вероятностей систематически используется термин «случайная величина» и всюду, где это необходимо, оговаривается, что автор имеет дело с независимыми случайными величинами.

Отметим еще одно обстоятельство. Ляпунов в своей работе «Об одном предложении теории вероятностей» в самом начале § 4 определил функцию распреде-

ления точно так же, как мы делаем это теперь. Он привел в этой работе широко используемую формулу

$$P(a \leq \xi < b) = F(b) - F(a).$$

Следует отметить, что в трактатах по теории вероятностей Ж. Пуанкаре (1854—1912) «Исчисление вероятностей», Бёртрана «Исчисление вероятностей», Чубера «Теория вероятностей и математическая статистика» понятие функции распределения не вводилось.

Определение случайной величины, данное Пуассоном, теперь уже не может считаться математическим. Это скорее всего описание реального объекта изучения, обращение к интуиции, полученной в результате житейского и научного опыта. Это описание широко используется и в наши дни на начальных ступенях педагогического процесса, связанного с изложением основ теории вероятностей. Даже несложный логический анализ этого определения показывает, что из него совсем не следуют правила для действий над случайными величинами — сложения, вычитания, умножения и пр. Для того чтобы случайная величина приобрела статус полноценного математического понятия, ей необходимо дать строго формализованное определение. Это было сделано в конце 20-х гг. Колмогоровым в упомянутой выше работе. Подход Колмогорова стал теперь общепринятым, поскольку он полноценно включил теорию вероятностей в общий стиль современного изложения, принятый в математике.

§ 3. Закон больших чисел

Знаменитая теорема Я. Бернулли о сближении при увеличении числа наблюдений вероятности события A с частотой его появления получила первое обобщение лишь в 1837 г. в работе Пуассона «Исследование о вероятностях в решении судебных дел уголовных и гражданских». В этом мемуаре он ввел термин «закон больших чисел». Существенный сдвиг в этом направлении связан с работой Чебышева «О средних величинах» (1867), опубликованной одновременно на русском и французском языках. В этой работе он перешел от рассмотрения случайных событий к случайным величинам и тем самым перенес центр тяжести интересов теории вероятностей к изучению случайных величин. Нужно отметить, что Чебышев не упоминал, что он интересуется только независимыми случайными величинами, а, согласно традициям того времени, считал, что других величин не рассматривают.

Закон больших чисел в форме Чебышева

Доказательство закона больших чисел для попарно независимых случайных величин с ограниченными дисперсиями основано на простом, но исключительно плодотворном неравенстве Чебышева

$$P(|\xi - M\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2},$$

которое неоднократно обобщалось многими авторами.

Отметим, что закон больших чисел в форме Чебышева содержит в качестве простейших частных случаев теоремы Бернулли и Пуассона. На это обстоятельство указал еще сам Чебышев. Рассуждения Чебышева приводят к простым необходимым и достаточным условиям для закона больших чисел, причем исходные случайные величины могут быть как угодно зависимы между собой. Вот формулировка этого результата:

Чтобы для последовательности случайных величин ξ_1, ξ_2, \dots (как угодно зависимых) с математическими ожиданиями a_1, a_2, \dots выполнялось соотношение

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{n} |s_n - A_n| < \varepsilon\right) = 1$$

при любом $\varepsilon > 0$, в котором $s_n = \sum_{k=1}^n \xi_k$, $A_n = \sum_{k=1}^n a_k$, необходимо и достаточно выполнение условия

$$M \frac{(s_n - A_n)^2}{n^2 + (s_n - A_n)^2} \xrightarrow{n \rightarrow \infty} 0.$$

Эта теорема в несколько различных формулировках была последовательно доказана С. Н. Бернштейном, Е. Е. Слуцким, Б. В. Гнеденко. Оригинальное обобщение теоремы Чебышева для одинаково распределенных независимых случайных величин было дано А. Я. Хинчиным. Он показал, что существование конечного математического ожидания достаточно для выполнения закона больших чисел.

В 1909 г. Э. Борель для $p = 0,5$ показал, что в случае схемы Бернулли имеет место более сильное предложение, чем закон больших чисел. Именно он доказал (в 1917 г. это предложение на произвольное p распространил итальянский математик Кантелли), что

$$P\left(\lim_{n \rightarrow \infty} \frac{\mu_n}{n} = p\right) = 1.$$

Это предложение получило наименование усиленного закона больших чисел.

В 1930 г. Колмогоровым была доказана следующая замечательная теорема: *если независимые случайные величины ξ_1, ξ_2, \dots имеют математические ожидания a_1, a_2, \dots , то*

$$P\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (\xi_k - a_k) = 0\right) = 1$$

в предположении, что сходится ряд $\sum_{k=1}^n \frac{D\xi_k}{k^2}$.

Позднее для случая одинаково распределенных независимых случайных величин Колмогоров доказал, что усиленный закон больших чисел имеет место при единственном условии существования у слагаемых конечного математического ожидания.

При доказательстве этих теорем Колмогоров опирался на замечательное неравенство, являющееся значительным обобщением неравенства Чебышева, а именно: если взаимно независимые случайные величины $\xi_1, \xi_2, \dots, \xi_n$ имеют конечные дисперсии, то вероятность совместного осуществления неравенств

$$\left| \sum_{s=1}^k (\xi_s - a_s) \right| < \varepsilon, \quad k = 1, 2, \dots, n,$$

не меньше, чем $1 - \frac{1}{\varepsilon^2} \sum_{k=1}^n D\xi_k$.

В 1935 г. Хинчин ввел новое понятие относительной устойчивости сумм, которое должно было дать максимально общую форму закона больших чисел для положительных случайных величин. Пусть ξ_1, ξ_2, \dots — последовательность неотрицательных случайных величин. Про суммы $s_n = \xi_1 + \xi_2 + \dots + \xi_n$ говорят, что они относительно устойчивы, если можно найти такие положительные константы A_n , что при любом $\varepsilon > 0$ и $n \rightarrow \infty$ выполняется соотношение

$$P\left(\left|\frac{s_n}{A_n} - 1\right| > \varepsilon\right) \rightarrow 0.$$

В случае одинаково распределенных величин ξ_k Хинчину удалось найти необходимое и достаточное условие для относительной устойчивости сумм s_n . Ученик Хинчина А. А. Бобров (1912—1989) распространил этот результат на случай поразному распределенных слагаемых.

Закон больших чисел рассматривался вплоть до 1939 г. как особая предельная теорема, причем обособленно от остальных предельных теорем для сумм независимых случайных величин. В работе Гнеденко, о которой речь пойдет в § 5, закон больших чисел был включен в общую теорию предельных теорем, когда предельное распределение имеет единственную точку роста в нуле. Точно так же теоремы об относительной устойчивости сумм являются предельными для того случая, когда предельное распределение имеет единственную точку роста при $x = 1$.

Существенное расширение проблематики, связанной с законом больших чисел, было осуществлено В. И. Гливленко в работах, относящихся к 1929—1933 гг., когда он начал рассматривать предельные теоремы для случайных величин со значениями в функциональных пространствах. Пожалуй, вершиной его результатов является замечательная теорема о сходимости эмпирических распределений к истинной функции распределения наблюдаемой случайной величины. Теорема Гливленко сразу же после ее опубликования была названа Кантелли основной теоремой математической статистики.

Пусть над случайной величиной ξ с функцией распределения $F(x)$ производятся n испытаний, в результате которых получена эмпирическая функция распределения $F_n(x)$. Согласно теореме Гливленко, имеет место равенство

$$P(\sup |F_n(x) - F(x)| \xrightarrow{n \rightarrow \infty} 0) = 1.$$

Теорема Гливленко неоднократно обобщалась. В этом направлении работало большое число исследователей, среди которых отметим лишь французских ученых Р. Форте (род. 1912) и Э. Мурье.

§ 4. Центральная предельная теорема

Локальная теорема Муавра долгое время служила образцом для последующих обобщений. Пожалуй, первое обобщение принадлежит Лапласу и уже формулируется как предельная теорема для сумм независимых случайных величин ξ_k , каждая из которых равномерно распределена в интервале $(-h; h)$. Это было сделано Лапласом в 1809 г. Он рассматривал дискретные случайные величины с увеличивающимся числом возможных значений. Этим самым давалась аппроксимация непрерывного распределения дискретным.

Заслуживает внимания тот факт, что Лаплас при доказательстве асимптотической нормальности суммы равномерно распределенных слагаемых использовал метод характеристических функций, который, естественно, так тогда еще не назывался.

Существенное продвижение исследований по предельной теореме связано с именем Пуассона. В знаменитом мемуаре 1837 г. «*Récherches sur la probabilité...*» он рассмотрел схему последовательности независимых испытаний с разными вероятностями появления события A в каждом из испытаний. Пусть вероятность наступления события A при k -м испытании равна p_k . Пуассон доказал для этого случая локальную теорему: *если ряд $\sum p_k (1 - p_k)$ расходится, то вероятность того, что в n испытаниях событие A появится m раз, равна*

$$P(m = np - \theta c \sqrt{n}) = \frac{1}{c \sqrt{\pi n}} e^{-\theta^2} - \frac{h(\theta)}{2c^4 n \sqrt{\pi}} (3 + 2\theta^2) e^{-\theta^2},$$

где

$$c^2 = 2 \sum p_k (1 - p_k), \quad h = \frac{4}{3n} \sum_{k=1}^n (2p_k - 1) p_k (1 - p_k).$$

В той же книге Пуассон дал ошибочное обобщение этой теоремы на сумму произвольных независимых случайных величин, имеющих конечные дисперсии, при условии их центрирования суммами математических ожиданий и нормирования корнем квадратным из суммы квадратов дисперсий слагаемых.

Следует сказать, что в частном случае одинаковой распределенности слагаемых эта теорема верна, однако строгое ее доказательство пришло значительно позднее и связано с именами наших современников — Линдберга, Феллера и Хинчина.

Отметим, что как работы Лапласа, так и работы Пуассона и всех последующих исследователей, занимающихся центральной предельной теоремой, были непосредственно связаны с теорией ошибок измерений. И во всех работах говорилось не о сложении абстрактных случайных величин, а о сложении ошибок. По-видимому, впервые от этой традиции отошел Чебышев.

Интерес к нормальному распределению в начале XIX в. возрос в связи с появлением знаменитых исследований Лежандра и Гаусса по формулировке и обоснованию метода наименьших квадратов. Бессель еще в 1818 г. заметил, что наблюдения гринвичского астронома Брэдли прекрасно укладываются в схему нормального распределения. Объяснение, которое он предложил этому факту, совпадает с идеей, которую перед этим в течение тридцати лет вынашивал Лаплас: результирующая большого числа случайных воздействий, каждое из которых оказывает малое влияние по сравнению с суммой остальных, подчиняется общему закону и этот общий закон должен быть нормальным. Эту мысль в совершенно отчетливой форме повторил Бессель также в работе 1838 г. Следует сказать, что Бессель обратил внимание на то, что это правило не является всеобщим и могут у ошибок наблюдений встречаться другие, отличные от нормального распределения. Так, если при измерении углов один источник ошибок превалирует над всеми остальными, то плотность распределения результирующей ошибки

$$f(x) = \frac{1}{\pi \sqrt{a^2 - x^2}}.$$

Та же концепция обоснования нормального закона, как закона распределения ошибок, дважды встречается в известном учебнике Пуанкаре «Calcul des probabilités» (Paris, 1912).

Следует отметить, что все эти идеи носят лишь качественный характер и нуждаются в математическом оформлении и последующем доказательстве строгих теорем. Вторым толчком, который вызвал дополнительный интерес к предельным теоремам теории вероятностей, была статистическая физика, начала которой были построены в середине XIX в.

Первый общий результат в этом направлении был сформулирован в 1887 г. в работе Чебышева «О двух теоремах относительно вероятностей» Теорема Чебышева имеет следующую формулировку.

Если математические ожидания величин u_1, u_2, \dots, u_n равны 0, а математические ожидания всех их степеней имеют числовую величину ниже какого-либо предела, то вероятность того, что сумма

$$u_1 + u_2 + \dots + u_n,$$

деленная на квадратный корень из удвоенной суммы математических ожиданий их квадратов, заключается между двумя какими-нибудь пределами t и t' , с возрастанием n до ∞ стремится к интегралу

$$\frac{1}{\sqrt{\pi}} \int_t^{t'} e^{-z^2} dz.$$

Для доказательства этого предложения Чебышевым был разработан метод, получивший наименование метода моментов и являющийся одним из крупнейших достижений науки того времени. Однако в формулировке теоремы и в ее доказательстве был допущен ряд неточностей, которые сразу же взялся устранить ученик Чебышева Марков. Критика мемуара Чебышева содержится в письмах Маркова к профессору Казанского университета А. В. Васильеву, который считал необходимым опубликовать в 1898 г. «Закон больших чисел и метод наименьших квадратов». В этих письмах Марковым была доказана совершенно строго несколько исправленная теорема Чебышева.

Если s_n — последовательность сумм $u_1 + u_2 + \dots + u_n$ и $\Phi_n(x)$ — их функции распределения, то из предположения, что при любом целом положительном k имеют место соотношения

$$\int_{-\infty}^{\infty} x^k d\Phi_n(x) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-\frac{x^2}{2}} dx,$$

вытекает, что при любых a и b

$$P(a \leq s_n < b) \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx,$$

Метод моментов, который разработал Чебышев, торжествовал победу. Была доказана сильная, и казалось бы, окончательная теорема. Некоторую неудовлетворенность только приносило то, что для простого результата требовалось выполнение счетного множества условий. Неожиданно в нескольких публикациях Ляпунова было обнаружено на протяжении 1900 и 1901 гг., что окончательный результат имеет место при выполнении только одного очень простого условия, которое, вдобавок, выясняло смысл предположений, приводящих к сходимости распределений нормированных и центрированных сумм к нормальному распределению.

Сначала Ляпунов показал, что если величины имеют конечные третьи моменты $c_k = M |\xi_k - a_k|^3$, $C_n = \sum_{k=1}^n c_k$, $B_n^2 = \sum_{k=1}^n D\xi_k$ и соотношение C_n/B_n^3 при $n \rightarrow \infty$ стремится к нулю, то имеет место сходимость функций распределения сумм s_n к нормальному распределению.

На следующий год Ляпунов обнаружил, что для окончательного результата не обязательно требуется существование третьих моментов слагаемых. Достаточно, чтобы существовали моменты некоторого порядка $2 + \delta$, где $\delta > 0$. Ляпунов показал, что для сходимости нормированных корней из дисперсии суммы независимых слагаемых к нормальному распределению достаточно выполнения следующего условия: отношение $\sum_{k=1}^n M |\xi_k - a_k|^{2+\delta}/B_n^{2+\delta}$ должно с ростом n стремиться к 0.

Ляпунов сделал несколько большее; он оценил скорость сходимости к предельному распределению функций распределения сумм. Порядок этой оценки оказался равным $n^{-\frac{1}{2} \ln n}$.

Точно так же в упомянутой статье Чебышева, помимо предложения о сходимости к нормальному распределению, было дано асимптотическое разложение по степеням $n^{-\frac{1}{2}}$.

Общность результатов Ляпунова произвела огромное впечатление на современников. По-видимому, именно в ту пору и появился термин «центральная предельная теорема» для обозначения условий сходимости функций распределения нормированных и центрированных математическими ожиданиями сумм к нормальному распределению. Марков подошел к результатам Ляпунова с иных позиций. В связи с этим полезно привести подлинные слова Маркова: «Общность выводов в последней работе Ляпунова далеко превосходила ту, которая была достигнута методом математических ожиданий. Достигнуть столь общих выводов методом математических ожиданий казалось даже невозможным, ибо он основан на рассмотрении таких математических ожиданий в неограниченном числе, существование которых в случаях А. М. Ляпунова не предполагается».

Для восстановления поколебленного таким образом значения метода математических ожиданий необходимо было выяснить, что вышеупомянутыми работами он не исчерпан до конца. Марков в 1908 г. выступил с замечательной идеей — урезания случайных величин. Этот прием дал возможность доказать предельную теорему в условиях Ляпунова методом моментов или, как говорил Марков, методом математических ожиданий. Идея урезания прочно вошла в жизнь теории вероятностей.

Дальнейшая история центральной предельной теоремы такова: в 1922 г. финскому математику Линдбергу удалось пойти дальше Ляпунова и отказаться от предположения существования даже каких-либо моментов кроме вторых. А именно, он доказал, что если при любом $\tau > 0$ имеет место соотношение

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-a_k| > \tau B_n} (x - a_k)^2 dF_k(x) = 0,$$

то функции распределения сумм, независимых, центрированных математически-

ми ожиданиями и нормированных корнем квадратным из суммы квадратов дисперсий слагаемых, сходятся к стандартному нормальному распределению.

Через 12 лет В. Феллер показал, что условие Линдберга является и необходимым в предположении, что слагаемые равномерно малы.

В работе 1927 г. Бернштейн рассмотрел несколько более общую задачу: имеется последовательность независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$, относительно которых не предполагается ни существования дисперсий, ни существования математических ожиданий. Спрашивается, при каких условиях можно подыскать такие постоянные $B_n > 0$ и A_n , чтобы распределение сумм $(s_n - A_n)/B_n$ сходилась к стандартному нормальному распределению.

Достаточные условия для этой задачи были найдены Бернштейном в той же работе 1927 г.; через восемь лет Феллер показал, что эти условия не только достаточны, но и необходимы в предположении, что слагаемые равномерно малы в смысле теории вероятностей.

В 1935 г. независимо один от другого Хинчин и Леви в постановке Бернштейна нашли необходимое и достаточное условие сходимости к нормальному распределению функций распределения сумм независимых, одинаково распределенных случайных величин.

Еще в 1926 г. в специальном курсе по предельным теоремам Хинчин задал следующий вопрос: имеется ли связь между законом больших чисел и центральной предельной теоремой? Ответ был найден Д. А. Райковым и А. А. Бобровым, которые доказали следующую теорему:

чтобы функции распределения сумм

$$\frac{\xi_1 + \xi_2 + \dots + \xi_n - A_n}{B_n}$$

при надлежащем выборе действительных постоянных $B_n > 0$ и A_n сходились к нормальному распределению, необходима и достаточна устойчивость сумм

$$\sum_{k=0}^n (\xi_k - a_k)^2,$$

где $a_n = \int_{\varepsilon}^{\varepsilon} x dF_n(x)$, $\varepsilon > 0$ — произвольно.

Исследование вопросов сходимости функций распределения к нормальному закону не окончились и в наши дни, но теперь исследуются другие вопросы: быстра сходимости к предельному распределению, сходимости случайного числа случайных слагаемых, суммирование неравномерно малых случайных величин.

§ 5. Общие предельные распределения для сумм

Естественный вопрос о том, какие распределения вообще возможны в качестве предельных для сумм независимых случайных величин при условии, что они примерно одинаковы по величине, возник только в двадцатые — тридцатые годы нашего столетия. Раньше во всей общности этот вопрос не возникал, хотя частные результаты по этому поводу и появлялись. Мы уже упоминали мемуар Пуассона, где было выведено и исследовано распределение, получившее впоследствии название распределения Коши.

На тридцать лет позднее, в 1853 г., в мемуаре «О средних результатах наблюдений той же природы и о результатах наиболее вероятных» Коши получил характеристическую функцию для всех тех распределений, для которых функция распределения суммы только на множитель при аргументе (коэффициент растяжения) отличается от распределения отдельных слагаемых. Коши нашел, что все такие функции имеют вид $f(t) = \exp\{-t^\mu\}$, где μ — положительное число. Позднее выяснилось, что $f(t)$ тогда и только тогда является характеристической функцией, когда $0 < \mu \leq 2$.

П. Леви в книге «Calcul des probabilités» (Paris, 1925) в VI главе «Экспоненциальные распределения» построил первую теорию устойчивых распределений.

Эта теория естественно продолжала исследования Коши, уйдя от них далеко вперед. Пусть $F(x)$ — функция распределения и $f(t)$ — ее характеристическая функция. Распределение $F(x)$ называется устойчивым, если при любых положительных постоянных a_1 и a_2 найдется такое положительное постоянное a , что выполняется равенство

$$f(a_1 t) f(a_2 t) = f(at).$$

Леви указал, что для устойчивых распределений функция $f(t)$ имеет вид $\exp \left\{ -c \left(1 + i\beta \frac{t}{|t|} \right) |t|^\alpha \right\}$, где $c > 0$, $-1 \leq \beta \leq 1$, $0 < \alpha \leq 2$. Леви также ввел понятие области притяжения устойчивого закона: множество всех тех распределений $F(x)$, для которых функции распределения независимых и распределенных по этому закону случайных величин при соответствующем нормировании сходятся к данному устойчивому распределению.

В 1935 г. Хинчин пополнил понятие устойчивого распределения, введенного Леви, а именно, он предложил называть устойчивыми те распределения, для которых линейная форма $a_1 \xi_1 + a_2 \xi_2$ при произвольных положительных постоянных a_1 и a_2 имеет такое же распределение, как $a \xi_1 + b$, где a — некоторое положительное, а b — действительное постоянное. Класс устойчивых в смысле Хинчина распределений оказался несколько шире класса Леви.

В 1939 г. независимо друг от друга Гнеденко и Деблин нашли области притяжения устойчивых распределений. Условия принадлежности области притяжения устойчивого закона очень простые и сводятся к поведению «хвостов» распределений — поведению исходного распределения при больших значениях аргумента.

Основной результат, принадлежащий Леви и Хинчину, можно сформулировать так: если ξ_1, ξ_2, \dots — последовательность одинаково распределенных независимых случайных величин, то суммы

$$s_n = \frac{\xi_1 + \dots + \xi_n - A_n}{B_n} \quad (1)$$

при надлежащем выборе постоянных $B_n > 0$ и действительных A_n могут сходить только к устойчивым законам распределения. Каждый устойчивый закон является предельным для функций распределения сумм (1).

В заметке 1936 г. Леви и Хинчин дали окончательное представление устойчивых распределений через логарифмы характеристической функции.

Этим полностью завершены исследования, которые были начаты Пуассоном и Коши.

Естественный вопрос о классе предельных распределений для сумм (1), когда слагаемые могут быть распределены неодинаково, был поставлен Хинчиным в письме к Леви. Вскоре ответ был предложен. По предложению Хинчина этот класс распределений получил наименование класса \mathcal{L} . На слагаемые суммы ξ_k/B_n при этом естественно наложить требование: каждое из слагаемых оказывает на сумму незначительное влияние. Это требование можно представить так: величины ξ_k/B_n предельно постоянны, т. е. для них можно найти такую последовательность постоянных m_{kn} , что равномерно относительно k ($1 \leq k \leq n$) для любого $\varepsilon > 0$ выполняется соотношение

$$P \left(\left| \frac{\xi_k}{B_n} - m_{kn} \right| > \varepsilon \right) \rightarrow 0 \quad \text{при } n \rightarrow \infty.$$

Характеристическое свойство законов класса \mathcal{L} , найденное Леви, состоит в следующем: чтобы функция $\varphi(t)$ была характеристической функцией закона класса \mathcal{L} , необходимо и достаточно выполнение следующего условия: при каждом α ($0 < \alpha \leq 1$) имеет место равенство $\varphi(t) = \varphi(\alpha t) f_\alpha(t)$, где $f_\alpha(t)$ — некоторая характеристическая функция.

На этом история вопроса не завершилась, поскольку оставалось для полного завершения проблематики ответить еще на один вопрос, поставленный Гнеденко. Каковы классы возможных предельных распределений, если случайные величины

$\xi_1, \xi_2, \dots, \xi_n, \dots$ могут быть распределены только по k ($1 \leq k \leq \infty$) различным законам распределения $F_1(x), F_2(x), \dots, F_k(x)$? Полное решение этой задачи было дано в 1971 г. Зингером.

В рассмотренном круге вопросов была изучена еще одна задача: а что будет, если рассматривать суммы (1) одинаково распределенных независимых слагаемых не по всем значениям n , а только по некоторой подпоследовательности? Какие предельные распределения при этом могут встретиться? Этот вопрос был поставлен Хинчиным; он же и дал на него ответ: класс возможных предельных распределений в только что указанном смысле совпадает с классом безгранично-делимых распределений, незадолго до 1936 г. введенного итальянским математиком Бруно де Финетти и подробно исследованным Колмогоровым, Леви и Хинчиным.

В 1933 г. Колмогоров высказал гипотезу, что если суммируются примерно равноправные независимые случайные величины, то при увеличении числа слагаемых их распределения будут приближаться к безгранично-делимым законам и, следовательно, если распределения последовательных сумм будут сходиться к предельному, то этот предельный закон обязательно должен быть безгранично-делимым. В предположении, что слагаемые имеют конечные дисперсии, а дисперсии последовательных сумм ограничены, эту гипотезу доказал ученик Колмогорова Г. М. Бавли (1908—1941) в 1934 г. В полном объеме эта гипотеза была доказана Хинчиным с привлечением довольно громоздких аналитических средств через три года. Отправляясь от этой работы, Гнеденко построил теорию суммирования независимых случайных величин, основанную на сравнительно легко доказываемом факте: если суммируются предельно постоянные независимые слагаемые и функции распределения соответствующих централизованных сумм сходятся к некоторому предельному, то можно построить последовательность безгранично-делимых случайных величин, функции распределения которых сближаются с функциями распределения сумм. Эти безгранично-делимые величины получили название сопровождающих. Из этого предположения в качестве частных случаев получились теоремы Бавли и Хинчина. Кроме того, этот подход давал возможность совершенно прозрачно найти условия существования предельных распределений и условия сходимости функций распределения сумм к любому возможному предельному распределению. В частности, были найдены необходимые и достаточные условия для закона больших чисел, для сходимости к нормальному распределению, распределению Пуассона, устойчивым распределениям. Весь круг этих вопросов нашел отражение в монографии Гнеденко и Колмогорова «Предельные распределения для сумм независимых случайных величин» (1949).

В последние годы большое число исследователей приступило к изучению предельного поведения сумм независимых случайных слагаемых в случайном числе. Первоначально усилия были сосредоточены только на условиях сходимости к нормальному распределению и выполнимости закона больших чисел. Позднее были поставлены вопросы о классе предельных распределений и об условиях существования предельного распределения. Эту задачу удалось решить в условиях одинаковой распределенности и независимости слагаемых, а также независимости числа слагаемых от самих слагаемых. Заметим также, что к самой постановке этих задач привели вопросы теории надежности и физики. Основная теорема, относящаяся к названной проблематике, получила наименование теоремы переноса. Вот ее содержание.

Пусть $\{\xi_{nk}\}$ — последовательность взаимно независимых случайных величин, для которых $P(\xi_{nk} < x) = F_n(x)$; $\{k_n\}$ — возрастающая последовательность положительных целых чисел и $\{v_n\}$ — последовательность целочисленных случайных величин, независимых от ξ_{nk} . Предположим, что выполнены условия

$$A) \quad P\left(\sum_{k=1}^{k_n} \xi_{nk} < x\right) \xrightarrow{n \rightarrow \infty} \Phi(x),$$

$$B) \quad P\left(\frac{v_n}{k_n} < x\right) \xrightarrow{n \rightarrow \infty} A(x), \quad A(+0) = 0,$$

где $\Phi(x)$ и $A(x)$ — функции распределения, тогда

$$C) \quad P\left(\sum_{k=1}^{v_n} \xi_{nk} < x\right) \xrightarrow{n \rightarrow \infty} \Psi(x).$$

Предельная функция $\Psi(x)$ определяется через свою характеристическую функцию, определяемую формулой

$$\psi(t) = \int_0^{\infty} (\varphi(t))^z dA(z).$$

Теорема переноса позволяет получить ряд важных следствий. В частности, оказывается, имеет место такой результат: предельное распределение для случайного числа случайных слагаемых может оказаться нормальным тогда и только тогда, когда $\Phi(x)$ нормально и $A(x)$ имеет единственную точку роста при $x = 1$.

§ 6. Закон повторного логарифма

От закона больших чисел взяла начало новая предельная закономерность, получившая наименование закона повторного логарифма. Эта теорема не ставит перед собой цели разыскания предельного распределения, но зато переводит задачу рассмотрения последовательных сумм совсем в новую область, а именно изучает поведение этих сумм всех вместе. Мы вначале рассмотрим эту задачу для простейшего случая — для схемы Бернулли. Это вполне естественно; тем более, что это соответствует историческому ходу исследований.

Обозначим через μ_n число появлений события A в n независимых испытаниях и рассмотрим разности $s_n = \mu_n - np$. В 1909 г. Борель дал обобщенную формулировку закона больших чисел, показав, что имеет место более сильное утверждение, а именно

$$P\left(\frac{s_n}{n} \rightarrow 0\right) = 1.$$

Через четыре года Ф. Хаусдорф (1868—1942) доказал, что имеет место еще более сильное утверждение, а именно, что при любом $\varepsilon > 0$

$$P\left(\frac{|s_n|}{\sqrt{n^{1+\varepsilon}}} \rightarrow 0\right) = 1.$$

Год спустя Г. Харди (1877—1947) и Дж. Литлвуд (1885—1977) обнаружили еще более сильное предложение, согласно которому с вероятностью единица отношение $s_n/\sqrt{n \ln n}$ остается ограниченным. В 1922 г. Хинчин дал для роста сумм s_n оценку $s_n = O(\sqrt{n \ln \ln n})$. Через два года Хинчин нашел окончательный результат. Оказалось, что

$$P\left(\overline{\lim}_{n \rightarrow \infty} \frac{|s_n|}{\sqrt{2npq \ln n}} = 1\right) = 1.$$

В 1926 г. Хинчину удалось распространить этот результат на случай схемы Пуассона, т. е. на случай последовательных испытаний с переменной вероятностью появления события A .

Работа Колмогорова 1929 г. значительно перекрывала результаты Хинчина, которые являлись для нее простыми следствиями. Этими словами мы не хотим преуменьшить значения работ Хинчина, поскольку открытие новой закономерности даже на простом случае заслуживает самой высокой оценки.

Пусть имеется последовательность ξ_1, ξ_2, \dots взаимно независимых случайных величин, имеющих математические ожидания $a_k = M\xi_k$ и дисперсии

$$b_k = D\xi_k; \quad B_n = \sum_{k=1}^n b_k, \quad s_n = \sum_{k=1}^n (\xi_k - a_k).$$

Если последовательность $\{\xi_k\}$ удовлетворяет еще двум условиям: при $n \rightarrow \infty$ $B_n \rightarrow \infty$ и $|\xi_k| < m_n = o((B_n / \ln \ln B_n)^{1/2})$, то она удовлетворяет закону повторного логарифма, т. е. для нее выполняется соотношение

$$P\left(\lim_{n \rightarrow \infty} \frac{|\xi_n|}{\sqrt{2B_n \ln \ln B_n}} = 1\right) = 1.$$

Позднее законом повторного логарифма занимались многочисленные исследователи: Леви, Феллер, Зигмунд и Марцинкевич, Хартман, Сарымсаков, Петров, Гнеденко и др. Среди многих прекрасных результатов мы выделим один: *если случайные величины ξ_k имеют конечную дисперсию (конечно, отличную от нуля), то это условие достаточно для выполнения закона повторного логарифма.*

Аналогичная задача была поставлена для устойчивых распределений, отличных от нормального. При этом выяснилось (Б. В. Гнеденко), что для любой не убывающей функции $u(n)$ и для любого устойчивого закона с показателем α ($0 < \alpha < 2$) с вероятностью 1 $\lim_{n \rightarrow \infty} \frac{|\xi_n|}{u(n)}$ равен нулю или бесконечности.

§ 7. Формирование понятий математического ожидания и дисперсии

Понятие математического ожидания в самых начальных его элементах было введено в теории вероятностей очень рано, впервые оно появилось в известной переписке Паскаля с Ферма. В более явной форме оно было введено Гюйгенсом, фактически определившим математическое ожидание для случайной величины, способной принимать два или три значения. Сам термин «ожидание» был предложен Схоутоном — учителем Гюйгенса. Этот термин прижился и сохранился до нашего времени. Но в ту пору этому термину придавался смысл ожидания той средней цены, которую можно дать за приобретение случайной величины, приносящей выигрыш x_1 с вероятностью p_1 , выигрыш x_2 с вероятностью p_2 , ..., выигрыш x_n с вероятностью p_n .

Эта мысль красной строкой проходит и в книге Н. Бернулли «О применении искусства предположений в вопросах права». Заслуживает внимания не только то, что Н. Бернулли рассмотрел ожидание для случайных величин, принимающих не только два или три значения, но и большее число значений, но и нечто совсем новое, в частности сравнение формулы для вычисления математического ожидания с правилом вычисления координат центра тяжести системы материальных точек.

Для XVIII в. обращение к математическому ожиданию было не характерным. Все внимание привлекло понятие вероятности случайного события. В энциклопедии науки о вероятностях, знаменитой книге Лапласа «Аналитическая теория вероятностей» нет определения математического ожидания и тем более правил действий с ним. Возможно, это связано с тем, что Лаплас не рассматривал и понятия случайной величины, вместо этого он изучал ошибки наблюдений, плотность их распределений и даже вывел и использовал формулу для плотности сумм двух независимых ошибок.

Казалось бы, создание и развитие теории ошибок наблюдений должно было стимулировать развитие числовых характеристик случайных величин (которые в ту пору еще назывались ошибками измерений). Однако этого не случилось. Впрочем, для нормального распределения были введены понятия истинного значения и точности наблюдений: было известно, как их вычислять по плотности распределения. Таким образом, для этого частного случая уже была известна формула для вычисления математического ожидания и дисперсии.

Обратим внимание на то, что в начале XIX в. нормальное распределение затмило собой все остальные, поскольку с ним столкнулись в теории ошибок наблюдений и, казалось, доказано в работах Гаусса и Лежандра, что распределение ошибок наблюдений должно быть нормальным. С ним же столкнулись в теории стрельбы. Бельгийский биолог Кетле давал многочисленные свидетельства того, что и в биологии нормальное распределение играет центральную роль. К остальным распределениям потеряли интерес, о них попросту не думали. Несомненно, в связи с этим никто и не помышлял о доказательстве теорем относительно математи-

ских ожиданий и дисперсий, поскольку для нормального распределения все уже было известно. В связи со сказанным интересно заметить, что в книге Чебышева «Опыт элементарного анализа теории вероятностей» (1845) понятия случайной величины, математического ожидания и дисперсии даже не упоминаются. Однако в курсе лекций по теории вероятностей, который систематически читал в Петербургском университете Чебышев, говорится о величинах (имея в виду случайные величины) с математическим ожиданием и дисперсией. Более того, в этих лекциях (записанных Ляпуновым, переписанных у него Крыловым и изданных в 1936 г. в издательстве АН СССР) было отмечено, что оно (понятие математического ожидания) имеет большее значение на практике, чем сама вероятность, потому что «на основании ее у нас составляется суждение о том, что мы можем ожидать перед появлением известного события» (с. 159). Само это утверждение не очень понятно, но, несомненно, Чебышев имел в виду какое-то определенное замечательное свойство математического ожидания. По-видимому, свою роль сыграла и формулировка закона больших чисел в форме Чебышева.

Заслуживает пристального внимания то обстоятельство, что в этих записках лекций имеется доказательство и формулировка теорем о математическом ожидании и дисперсии суммы случайных величин. Там же он привел и вывод своего знаменитого неравенства. При этом он предполагал как нечто самоочевидное, что речь идет о независимых величинах. Следует отметить, что сам факт о том, что дисперсия суммы равна сумме дисперсий, имеется и использован Чебышевым в статье «О средних величинах». Там же впервые встречается и неравенство Чебышева. Следует также отметить, что в распространенных учебниках Пуанкаре и Бертрана вообще нет теорем о математическом ожидании и дисперсии.

Естественно спросить себя: когда же стал известен факт, что математическое ожидание суммы равно сумме математических ожиданий всегда, а не только при независимых слагаемых? Пока на это можно ответить, что в учебнике Чубера (1908) и в переиздании книги Пуанкаре (1912) такой теоремы нет. А в знаменитом для своего времени учебнике «Исчисление вероятностей» (1913, 1924) строго доказываются и теоремы о математическом ожидании произведения и математическом ожидании суммы со специальным упоминанием о том, что она верна не только для независимых величин.

Таким образом, основы понятия математического ожидания возникли одновременно с понятием вероятности, но выделены основные его свойства были очень поздно — только во второй половине прошлого — в начале нынешнего столетия.

Глава 4

К ИСТОРИИ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 1. Общие представления

Понятие случайного процесса связано с именами Колмогорова, Хинчина, Слуцкого, Н. Винера (1894—1965). Это понятие в наши дни является одним из центральных не только в теории вероятностей, но также в естествознании, инженерном деле, экономике, организации производства, теории связи. Теория случайных процессов принадлежит к категории наиболее быстро развивающихся математических дисциплин. Несомненно, что это обстоятельство в значительной мере определяется ее глубокими связями с практикой.

XX в. не мог удовлетвориться тем идейным наследием, которое было получено им от прошлого. Действительно, в то время как физика, биолога, инженера интересовал процесс, т. е. изменение изучаемого явления во времени, теория вероятностей предлагала им в качестве математического аппарата лишь средства, изучавшие стационарные состояния. Для исследования изменения во времени теория вероятностей конца прошлого — начала нынешнего века не имела ни разработанных частных схем, ни тем более общих приемов. А необходимость их создания буквально стучала в окна и двери математической науки. Изучение броуновского движения в физике подвело математику к порогу создания теории случайных процессов. В исследованиях датского ученого А. К. Эрланга была начата новая важная область поисков, связанных с изучением загрузки телефонных сетей. Число вызо-

вов абонентов изменяется во времени случайно, а длительность каждого разговора обладает большой индивидуальностью. И вот в этих-то условиях двойной случайности следует производить расчет пропускной способности телефонных сетей, коммутационной аппаратуры и управляющих связью систем. Несомненно, что работы Эрланга оказали значительное влияние не только на решение чисто телефонных задач, но и на формирование элементов теории случайных процессов, в частности процессов гибели и размножения.

Во втором десятилетии ХХ в. начались исследования динамики биологических популяций. Итальянский математик Вито Вольтерра разработал математическую теорию этого процесса на базе чисто детерминистских соображений. Позднее ряд биологов и математиков развивали его идеи уже на основе стохастических представлений. Первоначально и в этой теории применялись исключительно идеи процессов гибели и размножения. Собственно именно от задач биологии и пошло наименование этого очень частного типа случайных процессов.

Представим себе, что мы задались целью проследить за движением какой-нибудь молекулы газа или жидкости. Эта молекула в случайные моменты времени сталкивается с другими молекулами и меняет при этом направление движения и скорость. Положение молекулы, таким образом, подвержено случайным изменениям и представляет собой не что иное, как случайный процесс. Этот процесс определяется шестью параметрами — тремя координатами и тремя компонентами скорости. Многие физические явления для своего изучения требуют умения вычислять вероятности того, что определенная доля молекул успеет за заданный промежуток времени перейти из одной области пространства в другую. Так, например, если приведены в соприкосновение две жидкости, то начинается взаимное проникновение молекул одной жидкости в другую. Происходит диффузия. По каким законам происходит процесс диффузии? На этот вопрос дает ответ статистическая теория диффузии, базирующаяся на использовании теории случайных процессов. Очевидно, что подобные же задачи возникают в химии, когда приступают к изучению процессов химических реакций.

Весьма важный круг явлений протекает по принципу радиоактивного распада. Суть его состоит в том, что атомы радиоактивного вещества распадаются, превращаясь в атомы другого элемента. Распад каждого атома происходит мгновенно, подобно взрыву, с выделением некоторого количества энергии. Многочисленные наблюдения показывают, что распад отдельных атомов происходит в случайно взятые моменты времени и расположение этих моментов, если количество распадающегося вещества не превосходит некоторого определенного критического предела, не зависят друг от друга. Для изучения процесса радиоактивного распада весьма важно определить вероятность того, что за определенный промежуток времени распадется то или иное число атомов. Формально, если задаться целью выяснения только математической стороны явлений, аналогично происходят многие другие процессы: обрывы нитей в прядильной машине, число броуновских частиц, оказавшихся в данный момент в определенной области пространства, вызовы от абонентов, поступающие на телефонную станцию, и т. д.

Теория броуновского движения, исходящая из теоретико-вероятностных предположений, была разработана в 1905 г. двумя известными физиками М. Смолуховским (1872—1917) и А. Эйнштейном (1879—1955). Позднее высказанные ими идеи использовались неоднократно как при изучении физических явлений, так и в различных инженерных задачах. В частности, именно с их работ, как и с работ Эрланга, начался широкий интерес к процессу Пуассона. Впрочем, сам Пуассон ввел в рассмотрение только распределение, носящее его имя, а о процессе даже не мечтал, но он заслужил того, чтобы его имя произносилось и при рассмотрении случайных процессов, связанных с его распределением. Это не единственный случай, когда в честь исследователя новым понятиям присваиваются их имена, хотя до этих понятий они и не доходили. Теперь широко распространены гауссовские случайные процессы, хотя сам Гаусс о них не имел никакого представления, да и само исходное распределение задолго до его рождения было получено Муавром, Лапласом и др. В теории же ошибок измерений одновременно с Гауссом к нему пришел также Лежандр.

Попытка изучения средствами теории вероятностей явления диффузии была предпринята в 1914 г. двумя известными физиками — М. Планком (1858—1947) и А. Фоккером,

Винер в середине двадцатых годов при изучении броуновского движения ввел в рассмотрение процесс, получивший название винеровского процесса (процесса броуновского движения).

Мы должны упомянуть еще о двух группах исследований, начатых в разное время и по разным поводам. Во-первых, это работа Маркова по изучению цепных зависимостей. Во-вторых, работы Е. Е. Слуцкого (1880—1948) по теории случайных функций. Оба эти направления играли очень существенную роль в формировании общей теории случайных процессов. Для этой цели уже был накоплен значительный исходный материал и необходимость построения теории как бы носилась в воздухе. Оставалось осуществить глубокий анализ уже имеющихся работ, высказанных в них идей и результатов и на его базе осуществить необходимый синтез.

В 1931 г. была опубликована большая статья Колмогорова «Об аналитических методах в теории вероятностей», а через три года работа Хинчина «Теория корреляции стационарных стохастических процессов», которые следует считать началом построения общей теории случайных процессов. В первой из этих работ были заложены основы теории марковских процессов, а во второй — основы стационарных процессов. Они были источником огромного числа последующих исследований, среди которых следует отметить статью Феллера «К теории стохастических процессов» (1936), давшую интегро-дифференциальные уравнения для скачкообразных марковских процессов.

Обе только что упомянутые основополагающие работы содержат не только математические результаты, но и глубокий философский анализ причин, послуживших исходным пунктом для построения основ теории случайных процессов. В работе Колмогорова были заложены основы теории случайных процессов без последствия и получены дифференциальные уравнения (прямые и обратные), которые управляют вероятностями перехода. В этой же работе был дан набросок теории скачкообразных процессов без последствия, подробное развитие которой позднее дано Феллером и Дубровским.

В настоящее время теория марковских процессов превратилась в большую и разветвленную главу математической науки, которая получила огромное число различных применений в физике, инженерном деле, геофизике, химии и ряде других областей знания.

Построение основ другого класса случайных процессов на базе физических задач было осуществлено Хинчиным в упомянутой нами работе. Он ввел понятие стационарного процесса в широком и узком смысле и получил знаменитое спектральное разложение корреляционной функции. Эта работа послужила основанием для последующих исследований Крамера, Вальда, Колмогорова и многих других ученых.

В процессе развития теории случайных процессов произошло разделение близких понятий. Если случайная величина $\xi(t)$ или вектор $(\xi_1(t), \dots, \xi_n(t))$ со значениями на числовой прямой зависят от одного действительного параметра t , то принято говорить о случайном процессе $\xi(t)$. При этом, как правило, параметр t носит название времени. Если время принимает дискретную последовательность значений — t_1, t_2, \dots , то говорят не о случайном процессе, а о случайной последовательности. Если же случайная величина ξ (или вектор) зависит не от одного, а от нескольких параметров, то ее называют случайным полем.

Со случайными полями столкнулись раньше всего в биологии и геофизике, а затем оказалось, что практически все области знания приводят к необходимости рассмотрения наряду со случайными процессами и случайных полей.

В истории каждой науки постоянно приходится сталкиваться с такими ситуациями, когда эта наука еще не создана, а исследователи рассматривают отдельные задачи, которые относятся к ее компетенции. Так было с арифметикой и геометрией, алгеброй и теорией чисел. С таким же положением мы сталкиваемся и в теории случайных процессов. Этой теории еще не было, не было и свойственных ей понятий, не было даже идеи рассмотрения изменения случайной величины во времени, а отдельные задачи в этом направлении уже изучались. Для примера еще Н. Бернулли, Монмор и Муавр занимались задачами о разорении игрока и состоянии игроков после n партий. Это типичная задача теории случайных процессов, в которой число сыгранных партий играет роль времени. Такая же ситуация складывается и с задачей Лапласа перекладывания шаров из урны и урну и подсчета содержания урны после n перекладываний. Всегда новое рождается в недрах ста-

рого и со временем вырастает из становящихся тесными рамок уже установившихся представлений и понятий. В результате появляется необходимость выделения специальной области научных исследований. Первоначально же отдельные новые задачи решаются в рамках старых представлений, как правило специальными приемами, создаваемыми для каждой задачи. Но время еще не созрело для выделения соответствующей новой ветви научного знания. Требуется иногда длительный срок, чтобы первоначальная идея и отдельные задачи сформировались и дали начало новой теории со своими постановками проблем и методами исследования, позволяющими продвинуться по пути познания явлений окружающего нас мира.

Теория вероятностей имеет богатую и поучительную историю. Она наглядно показывает, как возникали ее основные понятия и развивались методы из задач, с которыми сталкивался общественный прогресс. История теории вероятностей еще далека от совершенства и требуется систематическая работа для того, чтобы восстановить пройденный путь и воздать должное ее создателям. При этом мы увидим, как человечество переходило от первичных догадок к более полному и совершенному знанию, как создание теории вероятностей позволяло переходить от строгих детерминистических представлений к более широким стохастическим концепциям, тем самым открывая новые возможности для более глубоких заключений о природе вещей.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Адомиан Дж. Стохастические системы.— М. : Мир, 1987.— 376 с.
2. Анисимов В. В. Случайные процессы с дискретной компонентой. Пределные теоремы.— К. : Вища шк. Изд-во при Киев. ун-те, 1988.— 184 с.
3. Беллев Ю. К., Чепурин Е. В. Основы математической статистики: В 2 ч.— М.: Изд-во Моск. ун-та, 1982.— 1983.— Ч. 1—2.
4. Бендат Дж., Пирсол А. Применения корреляционного и спектрального анализа.— М. : Мир, 1983.— 312 с.
5. Боровков А. А. Теория вероятностей.— М. : Наука, 1976.— 352 с.
6. Вентцель Е. С. Исследование операций: задачи, принципы, методология.— М. : Наука, 1980.— 208 с.
7. Вентцель Е. С., Овчаров Л. А. Теория вероятностей.— М. : Наука, 1973.— 368 с.
8. Гихман И. И., Скороход А. В. Введение в теорию случайных процессов.— М. : Наука, 1977.— 568 с.
9. Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика.— К. : Вища шк. Головное изд-во, 1988.— 439 с.
10. Гнеденко Б. В. Курс теории вероятностей.— 5-е изд.— М. : Наука, 1988.— 448 с.
11. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания.— М. : Наука, 1987.— 336 с.
12. Гнеденко Б. В., Хинчин А. Я. Элементарное введение в теорию вероятностей.— М. : Наука, 1976.— 144 с.
13. Дуб Дж. Л. Вероятностные процессы.— М. : Изд-во иностр. лит., 1956.— 606 с.
14. Дынкин Е. Б., Юшкевич А. А. Теоремы и задачи о процессах Маркова.— М. : Наука, 1967.— 232 с.
15. Ежов И. И., Скороход А. В., Ядренко М. И. Элементы комбинаторики.— М. : Наука, 1977.— 80 с.
16. Ермаков С. М., Михайлов Г. А. Курс статистического моделирования.— М. : Наука, 1976.— 319 с.
17. Захаров В. К., Севастьянов Б. А., Чистяков В. П. Теория вероятностей.— М. : Наука, 1983.— 160 с.
18. Золотарев В. М. Современная теория суммирования независимых случайных величин.— М. : Наука, 1986.— 416 с.
19. Ивченко Г. И., Каштанов В. А., Коваленко И. Н. Теория массового обслуживания.— М. : Вышш. шк., 1982.— 256 с.

20. *Ивченко Г. И., Медведев Ю. И.* Математическая статистика.— М. : Высш. шк., 1984.— 248 с.
21. *Кемени Дж., Снелл Дж.* Конечные цепи Маркова.— М. : Наука, 1970.— 272 с.
22. *Климов Г. П.* Теория вероятностей и математическая статистика.— М. : Изд-во Моск. ун-та, 1983.— 328 с.
23. *Коваленко И. Н., Кузнецов Н. Ю.* Методы расчета высоконадежных систем.— М. : Радио и связь, 1988.— 176 с.
24. *Коваленко И. Н., Сарманов О. В.* Краткий курс теории случайных процессов.— К. : Выща шк. Головное изд-во, 1978.— 262 с.
25. *Коваленко И. Н., Филиппова А. А.* Теория вероятностей.— М. : Высш. шк., 1982.— 256 с.
26. *Колчин В. Ф., Севастьянов Б. А., Чистяков В. П.* Случайные размещения.— М. : Наука, 1976.— 224 с.
27. *Королюк В. С., Турбин А. Ф.* Полумарковские процессы и их приложения.— К. : Наук. думка, 1976.— 184 с.
28. *Крэйн М., Лемуан О.* Введение в регенеративный метод анализа моделей.— М. : Наука, 1982.— 104 с.
29. *Левин Б. Р.* Теоретические основы статистической радиотехники: В 2 кн.— М. : Сов. радио, 1966.—1968.— Кн. 1—2.
30. *Мешалкин Л. Д.* Сборник задач по теории вероятностей.— М. : Изд-во Моск. ун-та, 1963.— 151 с.
31. *Прохоров Ю. В., Розанов Ю. Л.* Теория вероятностей.— М. : Наука, 1973.— 494 с.
32. *Пугачев В. С.* Теория случайных функций и ее применение к задачам автоматического управления.— М. : Физматгиз, 1962.— 883 с.
33. *Розанов Ю. А.* Теория вероятностей, случайные процессы и математическая статистика.— М. : Наука, 1985.— 320 с.
34. *Севастьянов Б. А., Чистяков В. П., Зубков А. М.* Сборник задач по теории вероятностей.— М. : Наука, 1980.— 224 с.
35. *Соловьев А. Д.* Аналитические методы расчета и оценки надежности // Вопросы математической теории надежности.— М. : Радио и связь, 1983.— С. 9—112.
36. *Феллер В.* Введение в теорию вероятностей и ее приложения: В 2 т.— М. : Мир, 1964—1967.— Т. 1—2.
37. *Тутубалин В. Н.* Теория вероятностей.— М. : Изд-во Моск. ун-та, 1972.— 230 с.
38. *Хан Г., Шапиро С.* Статистические модели в инженерных задачах.— М. : Мир, 1969.— 396 с.
39. *Ширяев А. Н.* Вероятность.— М. : Наука, 1980.— 576 с.
40. *Чистяков В. П.* Курс теории вероятностей.— М. : Наука, 1982.— 256 с.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аксиома непрерывности** 10
— сложения расширенная (аксиома счетной аддитивности) 10, 283
Аппроксимация нормальная 31
— пуассоновская 27
Ассоциативность операций над событиями 8
- Блуждание случайное неограниченное** 45
— — ограниченное 49
- Величина асимптотически пуассоновская** 41
— вырожденная нормальная 86
— очереди 244
— пуассоновская (распределенная по закону Пуассона) 27
— распределенная по закону Релея 100
— случайная 11, 57
— асимптотически нормальная 170, 171
— дискретная 11
— комплексная 90
— — независимая 91
— комплексно сопряженная 90
— некоррелированная 74
— непрерывная 65
— одномерная нормальная (гауссова) 80
— равномерно распределенная 64
— центрированная 69
— экспоненциальная (показательная) 98
— m -мерная нормальная невырожденная 83
— — целочисленная 45
- Величины случайные бесконечно малые** 185
— — независимые 20, 45, 61
— — одинаково распределенные 45
- Вероятности априорные** 18, 260
— предельные 118
— стационарные 118
— уклонений 28
- Вероятность вырождения** 149
— перехода цепи Маркова 138
— потери требования 251
— события 5, 10
— апостериорная 18
— безусловная 16
— статистическая 6
— условная 16
— состояния финальная 116
- Время ожидания среднее** 251
— пребывания требования среднее 251
- Гамма-распределение** 99
Гипотеза альтернативная 263
— основная 263
— простая 263
— сложная 263
— статистическая 263
- Граф переходов** 126, 144
Группа событий полная 17
- Дисперсия случайной величины** 70
- Дистрибутивность операций над событиями** 8
Длина кодового слова средняя 37
Допуски упреждающие 162
Доска Шеппарда 13
- Задача интервальной оценки** 257
— оценки параметра 257
— статистического выбора 257
— — решения 256
- Закон больших чисел** 29, 48, 154
— — — в форме Бернулли 29
— — — — Хинчина 161
— — — — Чебышева 76, 308
— — — усиленный 163
- Значение случайного процесса** 134, 189 7
- Инвариантность предельного перехода** 219
Индикатор множества 285
— события 8
- Интеграл Лебега абстрактный** 285
— Лебега — Стильеса 287
— Римана — Стильеса 287
— Стильеса 60
— стохастический 191, 192, 204
- Интенсивность выхода цепи Маркова** 139
— перехода цепи Маркова 139
— потока 237
— стационарного 245
- Интервал доверительный** 278
— — односторонний 279
— — симметричный 279
- Испытания** 5
— независимые 21
— с несколькими исходами 35
- Исходы** 6
- Классификация Кендалла** 244
Класс сообщающихся состояний 126
— — замкнутый 126
- Ковариация** 70
- Коммутативность операций над событиями** 8
- Компоненты в показательной форме** 90
— — тригонометрической форме 90
- Коэффициент готовности** 55, 56, 242
— — системы 120
— корреляции 70
— надежности 242
- Коэффициенты сноса и диффузии локальные** 228
- Кривая Гаусса** 79
— — интегральная 80
- Критерии непараметрические** 265
Критерий возвратности 128
— гипотезы 263
— допустимый 263
— Колмогорова 266
— Неймана — Пирсона (критерий отношения правдоподобия) 261
— положительности 129
— равномерно наиболее мощный 263
— согласия 265
— эргодичности 130

Матрица интенсивностей 139

- корреляционная 72
- перехода 110
- — цепи Маркова 113, 139
- случайная 72
- стохастическая 110

Медiana выборочная 282

Мера 283

- абсолютно непрерывная 288
- вероятностная (распределение вероятностей) 10
- дискретная 283
- конечная 284
- нормированная 284
- Радона 237

Места для ожидания 244

Метод наименьших квадратов 276

Множества измеримые 284

Множество борелевское 59

Модель статистическая 256

Момент биномиальный 40

- начальный 69
- факториальный 40
- центральный 69
- эмпирический 277
- n -го восстановления 229

Монотонность 285

Независимость статистическая 19

Неравенство Бернштейна 123, 125

- Колмогорова 162
- Рао — Крамера 272
- Чебышева 75

Несмещенность 271

Область гипотезы критическая 258

Объединение событий 6, 29

Ожидание математическое 11, 48, 58, 90, 189

- — случайной матрицы 72
- — условное 16, 106
- — частичное 16

Оператор 208

- разностный 210
- скользящего среднего 211

Определение вероятности аксиоматическое 9

Определенность неотрицательная 190

Ординарность потока 238

Отклонение вероятное 81

— среднеквадратическое 70

Отношение правдоподобия 260

Оценка асимптотически эффективная 274

- — нормальная 275
- максимального правдоподобия 275
- моментная 277
- несмещенная 271
- состоятельная 270

— «хвоста» нормального распределения 82

— эффективная 271

Оценки асимптотические 184

— точные 184

Ошибка второго рода 259

— первого рода 258

Параметр мешающий 279

— потока 237

— стационарного 245

Переход цепи Маркова 109

Плотность вероятности (плотность распределения) 65

- восстановления 231
- меры 288
- перехода 227
- спектральная 197, 199
- условная 105

Полнодоступность 244

Полнота пространства 154

Положение частицы 47, 109

Понятие вероятности классическое 12

Последовательность бернуллиева 22

- испытаний Бернулли 22
- независимых случайных величин 45, 61
- стационарная 166
- сходящаяся по вероятности 153
- — в среднем 153
- — — квадратическом 153
- эргодическая 168

Поток однородных событий простейший 133

— — ординарный 245

— Пальма 237

Правила де-Моргана 8

Предел среднеквадратический 153

Преобразование Лапласа — Стильеса 230

— линейное 208

Принципы инвариантности 219

Произведение событий 6

Производная среднеквадратическая 193

Пространство вероятностное 10

— измеримое 284

— с мерой 284

— элементарных событий 6

Процесс ветвящийся 147

— винеровский 213

— стандартный 214

— восстановления 229

— обрывающийся 243

— общий 236

— стационарный 236

— гауссовский 223

— диффузионный 228

— дробового эффекта 136

— марковский 227

— однородный 138

— полумарковский 238

— общий 239

— стационарный 240

— Пуассона 134

— пуассоновский обобщенный 221

— размножения и гибели 145

— с дискретным спектром 198

— случайный 134, 189

— действительный 189

— комплексный 189

— марковский 138

— многомерный 189

— некоррелированных 212

— непрерывный в среднем квадратическом 190

— одномерный 189

— с дискретным временем 190

— — непрерывным временем 190

— стационарный в широком смысле 194

— стохастически непрерывный 223

— узкополосный 206

— центрированный 189

— эргодический 201

— с независимыми приращениями 134, 220

— непрерывным спектром 198

— ортогональными приращениями 204

Процессы стационарно связанные 194

Равенство Парсеваля 161

Разложение корреляционной функции

— спектральное 197

— Эджворта 185

Разность событий 7

Распределение биномиальное 24

— вырожденное 245

— нормальное 86

— Коши 103

— полиномиальное 36

— предельное 185

— безгранично делимое 185

— случайной величины 11, 45

— стандартное одномерное нормальное (гауссово) 79

— стационарное 117

— Стюдента 103

— суммы случайных величин 87

- теоретическое 265
- условное 104
- «хи-квадрат» с n степенями свободы 99
- Цепи Маркова начальное 110, 138
- — — предельное 116
- — — эргодическое 116, 144
- Эрланга 99
- Расстояние Хемминга 18
- Регрессия 107
- Режим стационарный 248
- Решение нерандомизированное 258
- рандомизированное 258
- статистическое байесовское 260
- Ряд распределения 11

- Свертка 88
- Свойства дисперсии 73
- непрерывности интеграла 286
- функции распределения 61
- Свойство марковости (марковское) 109, 138, 227
- экспоненциального распределения основное 132
- Семейство распределений экспоненциальных 264
- Система обратных дифференциальных уравнений Колмогорова 143
- обслуживания 243
- одноканальная (двухканальная и т. п. 244
- прямых дифференциальных уравнений Колмогорова 142
- с ограниченной очередью 244
- — ожиданием 244
- — потерями 244
- Следствие события 7
- Событие достоверное 6
- невозможное 6
- противоположное 6
- рекуррентное 51, 52
- случайное 6
- События благоприятствующие 6
- независимые 17, 19
- несовместные 7
- попарно независимые 19
- — несовместные 7
- равные 7
- случайные 5
- элементарные 6
- Совокупность генеральная 257
- Сообщаемость состояний 144
- Состояние аperiodическое 128
- возвратное 128
- — нулевое 128
- достижимое 126
- невозвратное 128
- периодическое 128
- положительное 128
- — периодическое 128 —
- цепи Маркова 109
- эргодическое 128
- Состояния несущественные 127
- сообщающиеся 126
- Спектр энергетический 198
- Статистика 257
- критерия 265
- хи-квадрат 267
- Сходимость среднеквадратическая 155

- Теорема Биркгофа — Хинчина эргодическая 167
- Бореля (усиленный закон больших чисел) 30
- восстановления 53
- — узловая 235
- — элементарная 234
- Гливленко — Кантелли 267
- интегральная предельная 37
- Колмогорова предельная 266
- Корольюка 247
- Лапласа интегральная 34
- Лебега 285

- Линдберга 180
- локальная предельная 37
- Ляпунова 181
- Муавра — Лапласа 31
- о среднем 285
- предельная для ограниченных слагаемых 179
- Пуассона (закон редких событий) 25
- — обобщенная 25
- Севастьянова 41
- Хинчина 198, 223, 246
- центральная предельная 173, 184
- Шеннона — Котельникова 207
- элементарная эргодическая 118
- Теоремы непрерывности 160
- Теория массового обслуживания 243
- Траектория случайного процесса 189
- Требование 243

- Уклонение нормированное 31, 172
- Уравнение восстановления 231, 232
- Колмогорова 228
- Чепмена — Колмогорова 138, 228
- Уравнения Маркова 113
- Уровень значимости 263
- Условие большой загрузки 250
- Линдберга 180
- Ляпунова 181
- ординарности 139

- Формула Байеса (формула гипотез) 18
- Бернулли 22
- включения и исключения 12
- Леви — Хинчина 187
- Лобачевского 101
- обращения 158, 159
- Поллачека — Хинчина 51, 253
- полного математического ожидания 17
- Райса 225
- Хинчина 197, 199
- Эрланга 251
- Функции распределения частные (маргинальные) 63
- Функционал линейный 208
- Функция восстановления 229
- Гаусса 34
- — интегральная 34
- измеримая 284
- Колмогорова 266
- корреляционная 168, 189
- — взаимная 189
- марковская переходная 227
- мощности 263
- производящая 50, 120
- распределения 58
- — совместная 63
- — условная 104
- — эмпирическая 265
- спектральная 198, 199
- ступенчатая 285
- характеристическая 46, 92

- Характеристика частотная 209, 210
- Характеристики второго порядка 189
- Хи-квадрат Пирсона 267

- Цепи Маркова с запретами 114
- Цепь Маркова 109
- — вложенная 252
- — неприводимая 130
- — однородная 109
- — с непрерывным временем 138
- — эргодическая 116, 144

Частота 5

Шум белый 201

- Энтропия сообщения 37
- Эффективность оценки 274
- — асимптотическая 274

ОГЛАВЛЕНИЕ

Предисловие	3
Глава 1. Основные понятия теории вероятностей	
§ 1. Случайные события. Статистическая вероятность	5
§ 2. Пространство элементарных событий. Действия над событиями	6
§ 3. Аксиоматическое определение вероятности	9
§ 4. Классическое понятие вероятности	12
§ 5. Условные вероятности. Независимость. Формула Байеса	15
Глава 2. Последовательность независимых испытаний	
§ 1. Независимые испытания с конечным числом исходов	21
§ 2. Последовательность испытаний Бернулли	22
§ 3. Теорема Пуассона (закон редких событий)	24
§ 4. Исследование поведения биномиальных вероятностей	27
§ 5. Вероятности уклонений. Закон больших чисел	28
§ 6. Теорема Муавра — Лапласа	31
§ 7. Интегральная теорема Лапласа	33
§ 8. Испытания с несколькими исходами	35
§ 9. Применение вероятностей уклонений к теории информации	37
§ 10. События, связанные с исходами независимых испытаний	39
§ 11. Неограниченное случайное блуждание	45
§ 12. Ограниченное случайное блуждание. Производящие функции	49
§ 13. Рекуррентные события	51
§ 14. Теорема восстановления	53
Глава 3. Случайные величины	
§ 1. Случайная величина. Функция распределения. Математическое ожидание	57
§ 2. Свойства функции распределения	61
§ 3. Непрерывные распределения	65
§ 4. Дисперсия. Ковариация. Моменты распределения	69
§ 5. Свойства дисперсии. Неравенство Чебышева	73
§ 6. Линейное преобразование случайной величины и нормальное распределение	77
§ 7. Распределение суммы случайных величин	87
§ 8. Комплексные случайные величины	90
§ 9. Характеристические функции	92
§ 10. Распределение некоторых сумм	98
§ 11. Преобразования случайных величин	102
§ 12. Условные распределения	104
Глава 4. Цепи Маркова	
§ 1. Цепь Маркова с дискретным временем	109
§ 2. Рекуррентные формулы	112
§ 3. Финальные вероятности	115
§ 4. Предельное поведение вероятностей	120
§ 5. Экспоненциальное распределение и процесс Пуассона	132
§ 6. Цепь Маркова с непрерывным временем	137
§ 7. Эргодическое распределение	144
§ 8. Ветвящийся процесс	147

Глава 5. Последовательности случайных величин

§ 1. Пространства и сходимость	152
§ 2. Теорема обращения. Теоремы непрерывности. Закон больших чисел	158
§ 3. Усиленный закон больших чисел	162
§ 4. Стационарные и эргодические последовательности	166

Глава 6. Центральная предельная теорема

§ 1. Асимптотическая нормальность	170
§ 2. Предельная теорема для ограниченных слагаемых	179
§ 3. Теорема Линдберга	180
§ 4. Теорема Ляпунова	181
§ 5. Отклонение от нормального распределения	184
§ 6. Другие предельные законы	185

Глава 7. Корреляционный анализ случайных процессов

§ 1. Случайный процесс, заданный характеристиками второго порядка	189
§ 2. Операции над процессами	191
§ 3. Случайный процесс, стационарный в широком смысле	194
§ 4. Анализ корреляционной функции. Эргодичность	196
§ 5. Спектральное разложение стационарного процесса	202
§ 6. Анализ узкополосного случайного процесса	206
§ 7. Линейные преобразования случайных процессов	208

Глава 8. Другие важные классы случайных процессов

§ 1. Винеровский процесс	213
§ 2. Сходимость к винеровскому процессу	215
§ 3. Процессы с независимыми приращениями	220
§ 4. Гауссовские процессы	223
§ 5. Марковские процессы	226

Глава 9. Теория восстановления и полумарковские процессы

§ 1. Процесс восстановления	229
§ 2. Теоремы восстановления	234
§ 3. Стационарный процесс восстановления	235
§ 4. Полумарковский процесс	238

Глава 10. Теория массового обслуживания

§ 1. Системы обслуживания	243
§ 2. Стационарный поток однородных событий	245
§ 3. Системы обслуживания $M M m r$	248
§ 4. Основные стационарные характеристики	249
§ 5. Метод вложенных цепей Маркова	252

Глава 11. Элементы математической статистики

§ 1. Задача статистического решения	255
§ 2. Выбор между двумя гипотезами	258
§ 3. Проверка статистической гипотезы	263
§ 4. Непараметрические критерии проверки гипотез	264
§ 5. Статистическая оценка параметров	270
§ 6. Асимптотические свойства статистических оценок	274
§ 7. Метод наименьших квадратов	276
§ 8. Метод моментов	277
§ 9. Доверительные интервалы	278

Дополнение

§ 1. Мера	283
§ 2. Интеграл Лебега	284
§ 3. Основные свойства интеграла	285
§ 4. Интеграл и мера в \mathbb{R}^m	287

Очерк истории теории вероятностей

Глава 1. Предыстория понятия вероятности и случайного события . . .	289
§ 1. Первоначальные данные	289
§ 2. Исследования Дж. Кардано и Н. Тарталья	290
§ 3. Исследования Галилео Галилея	291
§ 4. Вклад Б. Паскаля и П. Ферма в развитие теории вероятностей	292
§ 5. Работа Х. Гюйгенса	293
§ 6. О первых исследованиях по демографии	294
Глава 2. Период формирования основ теории вероятностей	296
§ 1. Возникновение классического определения вероятности	296
§ 2. О формировании понятия геометрической вероятности	297
§ 3. Основные теоремы теории вероятностей	299
§ 4. Задача о разорении игрока	300
§ 5. Возникновение предельных теорем теории вероятностей	301
§ 6. Контроль качества продукции	303
Глава 3. К истории формирования понятия случайной величины	304
§ 1. Развитие теории ошибок наблюдений	304
§ 2. Формирование понятия случайной величины	306
§ 3. Закон больших чисел	308
§ 4. Центральная предельная теорема	310
§ 5. Общие предельные распределения для сумм	313
§ 6. Закон повторного логарифма	316
§ 7. Формирование понятий математического ожидания и дисперсии	317
Глава 4. К истории теории случайных процессов	318
§ 1. Общие представления	318
Список рекомендуемой литературы	321
Предметный указатель	323

Учебник

Коваленко Игорь Николаевич
Гнеденко Борис Владимирович

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Переплет художника Н. Т. Кормыло
Художественный редактор С. В. Анненков
Технический редактор Л. Ф. Волкова
Корректор С. Я. Кахетелидзе

ИБ № 13314

Сдано в набор 29.09.89. Подписано в печать 08.06.90.
Формат 60×90/16. Бум. тип. № 2. Гарнитура литера-
турная. Высокая печать Усл. печ. л. 20,5. Усл.
кр.-отт. 20,5. Уч.-изд. л. 22,46. Тираж 8000 экз. Изд.
№ 8795. Заказ № 214. Цена 95 к.

Издательство «Выща школа», 252054, Киев-54, ул. Го-
голевская, 7.

Отпечатано с матриц Головного предприятия респуб-
ликанского производственного объединения «Поли-
графкнига». 252057, Киев-57, ул. Довженко, 3 на Бело-
церковской книжной фабрике, 256400, г. Белая Цер-
ковь, ул. К. Маркса, 4.

