КУРС ВЫСШЕЙ МАТЕМАТИКИ

ЛЕКЦИИ И ПРАКТИКУМ

Под общей редакцией
И. М. ПЕТРУШКО

Издание третье, стереотипное

ЛАНЬ®

САНКТ-ПЕТЕРБУРГ • МОСКВА • КРАСНОДАР
2008

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

ДОПУЩЕНО Министерством образования РФ в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлениям: «Технические науки», «Техника и технологии»

ЛАНЬ®

САНКТ-ПЕТЕРБУРГ • МОСКВА • КРАСНОДАР
2008

КУРС ЛЕКЦИЙ

Лекция 1

АКСИОМАТИКА ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 1.1. ПРЕДМЕТ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Теория вероятностей — математическая дисциплина, изучающая *случайные явления* или *эксперименты.* Случайные эксперименты отличаются от детерминированных тем, что при выполнении всех условий проведения эксперимента невозможно добиться заранее определенного результата.

Так, если эксперимент состоит в выборе наудачу трех карт из тщательно перетасованной новой колоды, то результат может быть совсем иным, чем мы предполагали. При стрельбе по мишени даже у первоклассного стрелка, находящегося в хорошей форме, возможны промахи. При покупке того или иного товара в магазине, которому мы вполне доверяем, возможно приобретение некачественной или поддельной продукции.

С каждым случайным экспериментом можно связать различные *случайные события,* которые могут осуществляться или нет в этом эксперименте. Случайное событие, по сути, представляет собой некое описание результата эксперимента. Случайные события принято обозначать большими латинскими буквами.

Приведем примеры случайных событий при выборе наудачу трех карт из колоды:

$A = \{$все три карты принадлежат одной масти$\}$,

$B = \{$среди трех карт есть по крайней мере один туз$\}$,

$C = \{$все три карты одного цвета$\}$.

Главный для теории вероятностей вопрос состоит в том, насколько часто осуществляется то или иное случайное событие при воспроизведении опыта. При этом предполагается, что опыты в неизменных условиях можно воспроизводить какое угодно число раз. Пусть опыт был воспроизведен n раз и при этом случайное событие A осуществилось $n(A)$ раз. *Частотой* события A называется число

$$\nu_n(A) = n(A)/n.$$

Подмечено, что с ростом n проявляется свойство *устойчивости частоты.* Более строго, это означает существование $\lim\limits_{n\to\infty} \nu_n(A)$. Значение этого предела называется *вероятностью* события A и обозначается $P(A)$ (от английского слова probability). Такое определение вероятности случайного события называется *статистическим.* Устойчивость частот является важнейшей предпосылкой для создания содержательной математической теории о случайных явлениях. Тем не менее, построение математической теории на основе статистического определения вероятности весьма затруднительно.

Приступим к изложению аксиоматики теории вероятностей, созданной нашим соотечественником, выдающимся математиком А. Н. Колмогоровым в тридцатые годы двадцатого столетия.

§ 1.2. ПРОСТРАНСТВО ЭЛЕМЕНТАРНЫХ ИСХОДОВ

Среди всевозможных случайных событий, связанных с данным случайным экспериментом, можно выделить так называемые *элементарные события* или *исходы,* которые характеризуются тем, что

а) любая реализация случайного эксперимента завершается одним и только одним элементарным исходом;

б) любое случайное событие можно представить как совокупность некоторых элементарных исходов.

ПРИМЕР 1.1. Пусть бросается симметричный игральный кубик. Нас интересует число выпавших очков на верхней грани кубика. Элементарными исходами будут следующие случайные события:

$$\omega_i = \{\text{выпало } i \text{ очков}\}, \quad i = 1, \ldots, 6.$$

Тогда, например, случайное событие

$$A = \{\text{выпало четное число очков}\}$$

представляет собою совокупность $\{\omega_2, \omega_4, \omega_6\}$.

ПРИМЕР 1.2. Пусть бросаются два игральных кубика. Нас интересует число выпавших очков на верхних гранях кубиков. Если даже кубики неразличимы, то разумно их пронумеровать. Пусть n_1 — число выпавших очков на первом кубике, n_2 — на втором. Элементарными исходами будут векторы (n_1, n_2), где $n_1, n_2 = 1, ..., 6$. Удобно представить все элементарные исходы в виде таблицы:

(1,1)	(1,2)	...	(1,6)
(2,1)	(2,2)	...	(2,6)
...	...	...	...
(6,1)	(6,2)	...	(6,6)

Хорошо видно, что общее число элементарных исходов равно 36. Рассмотрим, например, случайное событие

$$A = \{\text{сумма выпавших очков не меньше } 10\}.$$

Перебирая все элементарные исходы из представленной таблицы, видим, что

$$A = \{(4,6), (5,5), (5,6), (6,4), (6,5), (6,6)\}.$$

Замечание 1.1. Если кубики неразличимы, то в качестве элементарных исходов можно рассматривать векторы (n_1, n_2), где n_1 — наименьшее из выпавших очков, а n_2 — наибольшее. Тогда общее число элементарных исходов равно 21.

Совокупность всех элементарных исходов образует *пространство элементарных исходов,* которое обозначается Ω. Мы пришли к понятию пространства элементарных исходов, рассмотрев последовательно понятия случайного эксперимента, случайного события, элементарного исхода.

В аксиоматике Колмогорова понятие пространства элементарных исходов является первичным. Предполагается, что изначально задано некоторое множество произвольной природы Ω, которое называется *пространством элементарных исходов,* а элементы этого множества называются *элементарными исходами* и обозначаются ω. При этом *случайными событиями* объявляются некоторые подмножества Ω, но не все. Разберемся с этим подробнее.

§ 1.3. АЛГЕБРА СЛУЧАЙНЫХ СОБЫТИЙ

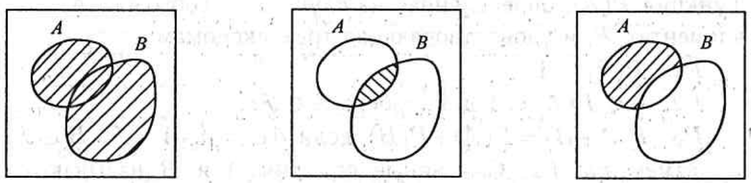
Пусть проводится некоторый случайный эксперимент. Над случайными событиями, связанными с этим экспериментом, можно производить определенные действия.

1) *Объединением* случайных событий A и B называется такое событие, которое происходит, если происходит хотя бы одно из событий A, B. Объединение обозначается $A + B$ и состоит из тех элементарных исходов, которые принадлежат или A, или B.

2) *Пересечением* случайных событий A и B называется такое событие, которое происходит, когда происходят оба события A и B. Пересечение обозначается AB и состоит из тех элементарных исходов, которые принадлежат и A, и B.

3) *Разностью* случайных событий A и B называется такое событие, которое происходит, если событие A происходит, а B нет. Разность обозначается $A \backslash B$ и состоит из тех элементарных исходов, которые принадлежат A, но не принадлежат B.

ПРИМЕР 1.3. Пусть в квадрате K наудачу выбирается точка. Элементарные исходы — всевозможные точки, составляющие K. Случайные события — подмножества K. Тогда указанные действия над случайными событиями легко продемонстрировать графически (рис. 1.1).



а) $A + B$ *б)* AB *в)* $A \backslash B$

Рис. 1.1.

Среди случайных событий выделяют достоверное и невозможное события. *Достоверное* событие осуществляется в каждой реализации случайного эксперимента. Оно состоит из всех элементарных исходов, т. е. равно Ω. *Невозможное* событие не осуществляется ни в одной реализации случайного эксперимента. Оно не содержит ни одного элементарного исхода и обозначается ϕ.

В аксиоматике Колмогорова, как мы помним, сначала задается пространство элементарных исходов Ω. *Случайными событиями* называются некоторые подмножества Ω (не обязательно все), образующие класс, замкнутый относительно операций объединения, пересечения и разности и включающий в себя Ω. Этот класс подмножеств называется *алгеброй случайных событий,* которая обычно обозначается $\mathcal{F}$. Таким образом, по определению, алгебра случайных событий $\mathcal{F}$ — это совокупность некоторых подмножеств Ω, включающая Ω и обладающая свойством: если A и B принадлежат $\mathcal{F}$, то $A+B$, AB, $A\backslash B$ также принадлежат $\mathcal{F}$. Заметим, что одному и тому же пространству элементарных исходов могут соответствовать различные алгебры случайных событий. Поэтому в аксиоматике Колмогорова наряду с пространством элементарных исходов Ω задается алгебра случайных событий $\mathcal{F}$. Именно элементы $\mathcal{F}$ называются случайными событиями, а остальные подмножества Ω нет.

§ 1.4. ВЕРОЯТНОСТЬ

Пусть заданы $(\Omega, \mathcal{F})$ — пространство элементарных исходов и алгебра случайных событий.

Определение 1.1. Вероятностью называется числовая функция $P(A)$, определенная на случайных событиях A, т. е. элементах $\mathcal{F}$, и удовлетворяющая трем аксиомам:

$P1$. $P(\Omega) = 1$;

$P2$. $0 \leqslant P(A) \leqslant 1$ для любого $A \in \mathcal{F}$;

$P3$. $P(A+B) = P(A)+P(B)$, если $AB = \phi$, $A \in \mathcal{F}$, $B \in \mathcal{F}$.

Замечание 1.2. Случайные события A и B называются *несовместными*, если они не могут осуществляться одновременно (в одной реализации случайного эксперимента). Иными словами, не существует ни одного общего элементарного исхода у A и B. Это означает, что $AB = \phi$. Таким образом, в аксиоме $P3$ речь идет о несовместных случайных событиях.

Поясним указанные аксиомы, исходя из частотного представления о вероятности. Вспомним, что частота $\nu_n(A)$ случайного события A в n реализациях случайного эксперимента равна $n(A)/n$, где n — общее число реализаций, а $n(A)$ — число тех из них, в которых произошло событие A. Поэтому $0 \leqslant n(A) \leqslant n$. После деления на n получаем, что

$$0 \leqslant \nu_n(A) \leqslant 1.$$

Это объясняет появление аксиомы 2. Если случайные события A и B несовместны, то

$$n(A+B) = n(A) + n(B).$$

Поэтому

$$\nu_n(A+B) = \nu_n(A) + \nu_n(B).$$

Это объясняет появление аксиомы 3. Самостоятельно поясните аксиому 1.

Определение 1.2. Вероятностным пространством называется тройка объектов $(\Omega, \mathcal{F}, P)$, где Ω — пространство элементарных исходов, $\mathcal{F}$ — алгебра случайных событий, P — вероятность, заданная на элементах $\mathcal{F}$.

Вероятностное пространство служит математической моделью случайного явления. Именно оно является основным объектом исследования теории вероятностей.

§ 1.5. СЛЕДСТВИЯ ИЗ АКСИОМ ВЕРОЯТНОСТИ

Определение 1.3. Случайное событие $\bar{A}$ называется *противоположным* случайному событию A, если оно происходит тогда, когда не происходит A, т. е. $\bar{A}$ состоит их всех тех элементарных исходов, которые не принадлежат A.

Легко видеть, что $A + \bar{A} = \Omega$, $A\bar{A} = \phi$.

Следствие 1.1. $P(\bar{A}) = 1 - P(A)$ *для любого случайного события* A, в частности $P(\phi) = 0$.

Доказательство. Поскольку $A+\bar{A} = \Omega$ и A и $\bar{A}$ — несовместные события, то по аксиомам $P1$ и $P3$:

$$1 = P(\Omega) = P(A+\bar{A}) = P(A) + P(\bar{A}).$$

Сравнивая крайние выражения, убеждаемся в справедливости следствия.

Замечание 1.3. Операции объединения и пересечения могут быть определены не только для двух, но и для любого конечного числа случайных событий $A_1, A_2, ..., A_n$. Так, например, случайное событие $A_1 + A_2 + ... + A_n$ осуществляется, если происходит хотя бы одно из событий $A_1, A_2, ..., A_n$.

Следствие 1.2. Если $A_1, A_2, ..., A_n$ — попарно несовместные случайные события (т. е. $A_i A_j = \phi$ при $i \neq j$, $i, j = 1, ..., n$), то

$$P(A_1 + A_2 + ... + A_n) = P(A_1) + P(A_2) + ... + P(A_n).$$

Доказательство. Очевидно, что

$$A_1 + A_2 + ... + A_n = A_1 + (A_2 + ... + A_n),$$

причем события A_1 и $A_2 + ... + A_n$ несовместны, поэтому по аксиоме $P3$

$$P(A_1 + A_2 + ... + A_n) = P(A_1 + (A_2 + ... + A_n)) = \\ = P(A_1) + P(A_2 + ... + A_n).$$

Последовательно продолжая эту процедуру, получаем требуемое утверждение.

Контрольные вопросы

1. В чем состоит отличие понятий случайного эксперимента и случайного события?
2. Почему в теории вероятностей изучаются только массовые случайные явления?
3. Опишите пространство элементарных исходов при выборе наудачу трех карт из колоды, состоящей из 36 карт.
4. Докажите для случайных событий A и B, что $A \backslash B = A\bar{B}$.
5. Сравните статистическое и аксиоматическое определения вероятности. В чем состоит преимущество, а в чем — неудобство статистического определения вероятности?
6. Будем говорить, что случайное событие A влечет случайное событие B ($A \subset B$), если из наступления события A следует наступление события B, т. е. все элементарные исходы, составляющие A, содержатся в B. Докажите, что $P(A) \leqslant P(B)$.

Ответы

3. Пронумеруем все карты из колоды. Пусть n_1 — номер первой взятой карты, n_2 и n_3 — второй и третьей соответственно. Элементарные исходы — векторы (n_1, n_2, n_3), где $n_1, n_2, n_3 = 1, ..., 36$ и n_1, n_2, n_3 различны.

4. Случайное событие $A \backslash B$ осуществляется, если произошло событие A и не произошло событие B. А это означает, что произошли и событие A, и событие $\bar{B}$, т. е. произошло событие $A\bar{B}$.

6. Очевидно, $B = A + (B \backslash A)$, причем A и $B \backslash A$ — несовместные случайные события, поэтому $P(B) = P(A) + P(B \backslash A) \geqslant P(A)$.

Лекция 2

ПРИМЕРЫ ВЕРОЯТНОСТНЫХ ПРОСТРАНСТВ

§ 2.1. КЛАССИЧЕСКАЯ ВЕРОЯТНОСТНАЯ МОДЕЛЬ

Пусть пространство элементарных исходов Ω состоит из конечного числа элементов $\omega_1, \omega_2, ..., \omega_n$ и алгебра случайных событий $\mathcal{F}$ — это совокупность всех подмножеств Ω. Если при этом вероятность всех элементарных исходов одна и та же (элементарные исходы равновозможны), то получаем так называемую *классическую вероятностную модель*.

Теорема 2.1. *Если случайное событие A (в классической вероятностной модели) состоит из k элементарных исходов, то*

$$P(A) = \frac{k}{n},$$

где n — общее число элементарных исходов.

(Элементарные исходы, составляющие случайное событие, называются *благоприятными*.)

Доказательство. Поскольку

$$\Omega = \omega_1 + \omega_2 + ... + \omega_n$$

и элементарные исходы по определению классической вероятностной модели являются случайными событиями, причем попарно несовместными, то по аксиоме $P1$ и следствию 1.2 (см. лекцию № 1)

$$1 = P(\Omega) = P(\omega_1) + P(\omega_2) + ... + P(\omega_n).$$

Так как при этом

$$P(\omega_1) = P(\omega_2) = ... = P(\omega_n),$$

то получаем, что

$$1 = nP(\omega_1) \;\Rightarrow\; P(\omega_1) = \frac{1}{n} \;\Rightarrow\; P(\omega_1) = P(\omega_2) = ... = P(\omega_n) = \frac{1}{n}.$$

Пусть случайное событие A состоит из k элементарных исходов $\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_k}$. Тогда

$$A = \omega_{i_1} + \omega_{i_2} + ... + \omega_{i_k},$$

причем случайные события справа попарно несовместны. По следствию 1.2

$$P(A) = P(\omega_{i_1}) + P(\omega_{i_2}) + ... + P(\omega_{i_k}) = \frac{1}{n} + \frac{1}{n} + ... + \frac{1}{n} = \frac{k}{n}.$$

ПРИМЕР 2.1. Пусть бросаются два игральных кубика. Найти вероятность того, что суммарное число выпавших очков не меньше 10.

Элементарный исход — это пара чисел (n_1, n_2), где n_1 — число очков на первом кубике, а n_2 — на втором. Рассмотрим случайное событие $A = \{$суммарное число очков не меньше 10$\}$. Перебирая все элементарные исходы (см. пример 1.2), получаем, что

$$A = \{(4,6), (5,5), (5,6), (6,4), (6,5), (6,6)\},$$

т. е. случайное событие A состоит из шести благоприятных исходов. Из соображений симметрии ясно, что все 36 элементарных исходов в данном примере равновозможны. Поэтому по теореме 2.1

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

Замечание 2.1. Если кубики неразличимы, то в качестве элементарных исходов можно рассматривать пары чисел (n_1, n_2), где n_1 — наименьшее из выпавших очков, а n_2 — наибольшее. Тогда

$$A = \{(5,5), (5,6), (4,6), (6,6)\}.$$

Поскольку всего элементарных исходов 21, то для вероятности события A получаем значение $4/21$, что не равно $1/6$. Этот парадокс объясняется тем, что рассмотренные элементарные исходы не равновозможны и поэтому нельзя пользоваться теоремой 2.1. Нетрудно убедиться на опыте, что элементарный исход (5,6) встречается примерно в два раза чаще, чем (5,5).

Введем некоторые комбинаторные понятия, тесно связанные с классической вероятностной моделью.

Размещение — упорядоченный выбор k элементов из совокупности n различных элементов (первому выбранному элементу присваивается номер один, второму — номер два и т. д.). Число размещений обозначается A_n^k. На первое место выбирается любой из n элементов, на второе — любой из $(n-1)$ оставшихся, на третье — любой из $(n-2)$ оставшихся и т. д. Поэтому

$$A_n^k = n(n-1)(n-2)\cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}, \tag{2.1}$$

где $n! = n(n-1)(n-2)\cdot \ldots \cdot 2\cdot 1$.

Перестановка — частный случай размещения, когда осуществляется упорядоченный выбор n элементов из совокупности n различных элементов. По сути перестановка — способ упорядочить n элементов. Число перестановок равно $n!$

Сочетание — неупорядоченный выбор k элементов из совокупности n различных элементов (нас интересует именно совокупность выбранных элементов, а порядок, в котором они выбирались, — нет). Число сочетаний обозначается C_n^k. Оказывается,

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{2.2}$$

Действительно, из каждой неупорядоченной совокупности из k элементов можно $k!$ способами сделать упорядоченную совокупность, поэтому

$$A_n^k = k!C_n^k,$$

откуда с учетом (2.1) получается требуемый результат (2.2).

Замечание 2.2. Напомним, что число сочетаний встречается в известной формуле для бинома Ньютона

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k},$$

где по определению $C_n^0 = 1$.

ПРИМЕР 2.2. Имеется n ящиков и r шаров, причем $r \leqslant n$. Шары наудачу размещаются по ящикам (каждый ящик может вместить сразу все шары). Найти вероятность того, что все шары окажутся в разных ящиках.

Пронумеруем все ящики и все шары. Пусть i_1 — номер ящика, в котором оказался первый шар; i_2 — номер ящика, в котором оказался второй шар, и т. д. Элементарный исход — это вектор $(i_1, i_2, \ldots, i_r)$. Очевидно, что каждое из чисел $i_1, i_2, \ldots, i_r$ может принимать любое натуральное значение от 1 до n. Поэтому число элементарных исходов равно n^r. Благоприятный исход $(i_1, i_2, \ldots, i_r)$ характеризуется тем, что все числа $i_1, i_2, \ldots, i_r$ различны. В качестве i_1 можно взять любое из n чисел от 1 до n, в качестве i_2 можно взять любое из $(n-1)$ оставшихся чисел и т. д. Поэтому число благоприятных исходов равно $n(n-1)\cdot \ldots \cdot (n-r+1)$. Ясно, что число благоприятных исходов совпадает с числом размещений A_n^r (для благоприятного исхода необходимо осуществить упорядоченный выбор r ящиков из совокупности n ящиков и в первый из выбранных ящиков положить первый шар, во второй из выбранных ящиков положить второй шар и т. д.). Из соображений симметрии все элементарные исходы равновозможны, поэтому вероятность того, что все шары окажутся в разных ящиках, равна $\dfrac{A_n^r}{n^r}$.

ПРИМЕР 2.3. Из студенческой группы, в которой 10 студентов и 12 студенток, для анкетирования произвольным образом отбирают пять человек. Найти вероятность того, что среди них будет три студентки.

Очевидно, нас не интересует порядок, в котором выбираются студенты (все было бы иначе, если бы первого выбранного назначали старостой, второго — профоргом и т. д.). Поэтому элементарный исход — это сочетание пяти элементов из совокупности 22 элементов. Очевидно, что все элементарные исходы равновозможны. Их общее число равно C_{22}^5. Благоприятные исходы — это сочетания, в которых будет три студентки и два студента. Чтобы получить такое сочетание, надо выбрать трех студенток из общего числа 12 (число способов равно C_{12}^3) и двух студентов из общего числа 10 (число способов равно C_{10}^2).

Таким образом, число благоприятных исходов равно $C_{12}^3 \times C_{10}^2$. Поэтому искомая вероятность равна $\dfrac{C_{12}^3\cdot C_{10}^2}{C_{22}^5}$.

§ 2.2. ГЕОМЕТРИЧЕСКИЕ ВЕРОЯТНОСТИ

Рассмотрим квадрат K. Выберем наудачу точку в этом квадрате. Опишем вероятностную модель, соответствующую такому выбору. Элементарные исходы — это точки квадрата K. Случайные события — это любые квадрируемые подмножества K (т. е. такие, для которых определено понятие площади). Будем предполагать, что все элементарные исходы равновозможны в следующем смысле. Вероятность того, что наудачу выбранная точка принадлежит произвольному маленькому прямоугольнику из K будет одна и та же независимо от того, где этот прямоугольник находится. Пусть $S(A)$ означает площадь области A.

Теорема 2.2. *Вероятность того, что наудачу выбранная в квадрате K точка принадлежит подобласти A этого квадрата, равна отношению* $S(A)/S(K)$.

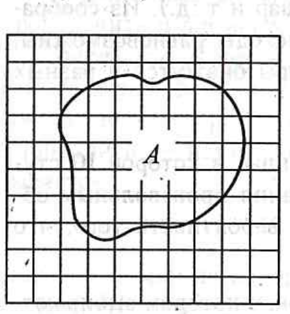


Рис. 2.1

Доказательство. Разобьем квадрат K на маленькие одинаковые прямоугольники (рис. 2.1).

Поскольку вероятность того, что наудачу выбранная точка M окажется в определенном прямоугольнике, одна и та же для всех прямоугольников, то вероятность того, что эта точка окажется в области A практически пропорциональна числу прямоугольников, находящихся внутри области A, а значит, пропорциональна площади области A (и это будет тем точнее, чем меньше выбраны прямоугольники). Итак, можно считать, что

$$P(M \in A) = cS(A), \tag{2.3}$$

где c — коэффициент пропорциональности. Рассматривая в качестве A сам квадрат K, получаем, что

$$1 = P(M \in K) = cS(K) \Rightarrow c = \frac{1}{S(K)}.$$

Подставляя найденное значение c в (2.3), получаем утверждение теоремы 2.2.

Замечание 2.3. Вместо квадрата K можно использовать любую квадрируемую область D на плоскости и считать, что наудачу выбранная в D точка окажется в подобласти A области D с вероятностью $S(A)/S(D)$. Вместо плоскости можно рассматривать пространство любой конечной размерности и выбирать наудачу точку в некоторой области этого пространства. Тогда в трехмерном пространстве при расчете вероятностей вместо понятия площади надо использовать понятие объема.

Замечание 2.4. Известно, что не все подобласти квадрата квадрируемы. Это объясняет, почему не все подмножества пространства элементарных исходов можно считать случайными событиями.

ПРИМЕР 2.4 (задача о встрече). Двое сотрудников А и Б решили встретиться в определенном месте во время обеденного перерыва, продолжающегося один час. Они договорились, что каждый приходит на место встречи, когда ему удобно, и ждет другого не более десяти минут. Найти вероятность того, что встреча произойдет.

Рассмотрим декартову систему координат (рис. 2.2). На оси Ox будем отмечать время прихода сотрудника А, а оси Oy — время прихода сотрудника Б.

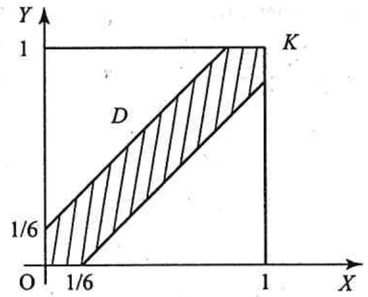


Рис. 2.2

Точка с указанными координатами представляет собой элементарный исход. Все элементарные исходы заполняют указанный на рис. 2.2 квадрат K. Очевидно, что все элементарные исходы равновозможны в указанном ранее смысле. Благоприятные исходы (x, y) составляют область $D = \{(x, y) : |x - y| \leqslant 1/6\}$ (десять минут — это 1/6 часа). На рис. 2.2 эта область заштрихована. Площадь области D легко найти, вычитая из площади квадрата K площади двух равнобедренных прямоугольных треугольников с катетами длины 5/6. Итак, $S(D) = 1 - (5/6)^2 = 11/36$.

Поэтому вероятность того, что встреча произойдет равна

$$\frac{S(D)}{S(K)} = \frac{11}{36} : 1 = \frac{11}{36}.$$

Контрольные вопросы

1. Что общего у двух рассмотренных вероятностных моделей, а в чем состоит их различие?
2. Все ли вероятностные пространства с конечным числом элементарных исходов являются классическими?
3. Подойдет ли рассмотренная модель геометрических вероятностей при распределении земельных участков в строящемся городе?
4. Какое понятие вместо понятия площади надо использовать при расчете геометрических вероятностей в случае одномерного пространства?
5. Найти вероятность того, что наудачу выбранная в квадрате точка окажется на одной из диагоналей этого квадрата.

Ответы

3. Не подойдет, поскольку в городе есть как более благоприятные для жизни районы, так и менее благоприятные.
4. Понятие длины.
5. 0, поскольку площадь области, состоящей из диагоналей квадрата, равна 0.

Лекция 3

ОСНОВНЫЕ ФОРМУЛЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 3.1. ТЕОРЕМА СЛОЖЕНИЯ ВЕРОЯТНОСТЕЙ

Известно из аксиомы $P3$, что если A и B — несовместные случайные события, то $P(A+B) = P(A) + P(B)$. Что будет в случае, когда A и B — совместные события?

Теорема 3.1 (сложения вероятностей). *Если A и B — произвольные случайные события, то*

$$P(A+B) = P(A) + P(B) - P(AB).$$

Доказательство. Событие $A+B$ можно представить как объединение попарно несовместных событий $A\backslash B$, AB, $B\backslash A$ (в случае выбора наудачу точки в квадрате K это разбиение $A+B$ на части показано на рис. 3.1).

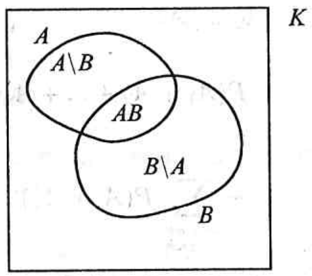


Рис. 3.1

Тогда по аксиоме $P3$

$$P(A+B) = P(A\backslash B) + P(AB) + P(B\backslash A). \quad (3.1)$$

Точно также $A = (A\backslash B) + AB$, $B = (B\backslash A) + AB$ и, следовательно,

$$P(A) = P(A\backslash B) + P(AB), \quad (3.2)$$

$$P(B) = P(B\backslash A) + P(AB). \quad (3.3)$$

Складывая формулы (3.2) и (3.3), получаем, что

$$P(A) + P(B) = P(A\backslash B) + 2P(AB) + P(B\backslash A). \quad (3.4)$$

Сравнивая формулы (3.1) и (3.4), видим, что

$$P(A)+P(B)=P(A+B)+P(AB),$$

откуда следует утверждение теоремы.

Замечание 3.1. Доказанная теорема не противоречит аксиоме $P3$, поскольку если A и B — несовместные события, то $P(AB)=0$.

Замечание 3.2. Если точка наудачу выбирается в квадрате со стороной длины 1, то для любого случайного события A

$$P(A)=S(A),$$

где $S(A)$ — площадь области A. Для пересекающихся областей A и B легко видеть, что

$$S(A+B)=S(A)+S(B)-S(AB),$$

что согласуется с теоремой 3.1.

Для большего, чем два, числа событий справедливо следующее обобщение теоремы 3.1.

Теорема 3.2. *Для произвольных случайных событий* $A_1, A_2, \dots, A_n$

$$P(A_1+A_2+\dots+A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{\substack{(i,j):\\ i<j}}P(A_iA_j)+$$
$$+\sum_{\substack{(i,j,k):\\ i<j<k}}P(A_iA_jA_k)-\dots+(-1)^{n+1}P(A_1A_2\cdot\ldots\cdot A_n)$$

(во всех суммах индексы $i, j, k, \dots$ принимают натуральные значения от 1 до n).

§ 3.2. ТЕОРЕМА УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

Введем условную вероятность $P(A|B)$. Это вероятность случайного события A при условии, что случайное событие B произошло. Как находить такую вероятность? Пусть случайный эксперимент реализован n раз. Нас интересуют только те реализации, при которых произошло событие B. Пусть их

число $n(B)$. Из этих реализаций необходимо отобрать те, которые привели к событию A. Нетрудно понять, что их число совпадает с числом реализаций, приведших к событию AB, т. е. с $n(AB)$. Разумно считать, исходя из свойства устойчивости частот, что при больших n

$$P(A|B)\approx\frac{n(AB)}{n(B)}=\frac{n(AB)/n}{n(B)/n}.$$

Но числитель последней дроби близок к $P(AB)$, а знаменатель — к $P(B)$.

Итак,

$$P(A|B)\approx P(AB)/P(B), \tag{3.5}$$

и это равенство тем точнее, чем больше n. Может показаться, что условная вероятность вычислена. Но надо заметить, что пока еще не доказано свойство устойчивости частот исходя из аксиом теории вероятностей. А вот в качестве определения соотношение (3.5) можно использовать.

Определение 3.1. Пусть $P(B)\neq 0$. *Условной вероятностью* случайного события A при условии, что случайное событие B произошло, называется следующее число:

$$P(A|B)=\frac{P(AB)}{P(B)}.$$

ПРИМЕР 3.1. Из колоды карт выбирается наудачу одна карта. Рассмотрим случайные события

$$A=\{\text{выбран туз}\},\quad B=\{\text{выбрана фигура}\}.$$

Найти $P(A|B)$.

Всего имеется 36 элементарных исходов. Событию B благоприятны 16 элементарных исходов. Событие AB совпадает с событием A, которому благоприятны четыре элементарных исхода.

Следовательно, $P(B)=16/36, P(AB)=4/36$ и

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{4}{16}=\frac{1}{4}.$$

Замечание 3.3. В случае классической вероятной модели при расчете условной вероятности $P(A|B)$ можно рассматривать только те элементарные исходы, которые приводят к событию B, и считать их равновозможными. Из этих исходов надо отобрать те, которые приводят к событию A, и их число разделить на число элементарных исходов, благоприятных событию B. Это и будет $P(A|B)$. В рассмотренном примере 3.1 из 16 элементарных исходов, благоприятных событию B, четыре элементарных исхода благоприятны событию A. Следовательно,

$$P(A|B) = \frac{4}{16} = \frac{1}{4}.$$

Теорема 3.3 (умножения вероятностей). *Пусть A_1, $A_2, \ldots, A_n$ — произвольные случайные события, тогда*

$$P(A_1 A_2 \cdot \ldots \cdot A_n) = P(A_1)P(A_2|A_1) \times \\ \times P(A_3|A_1A_2) \cdot \ldots \cdot P(A_n|A_1A_2 \cdot \ldots \cdot A_{n-1}).$$

Доказательство (при $n = 3$). По определению условной вероятности

$$P(A_1)P(A_2|A_1)P(A_3|A_1A_2) = \\ = P(A_1)\frac{P(A_1A_2)}{P(A_1)}\frac{P(A_1A_2A_3)}{P(A_1A_2)} = P(A_1A_2A_3),$$

что требовалось доказать.

ПРИМЕР 3.2 (задача о совпадениях). Пусть имеется n различных писем, и они наудачу раскладываются по подписанным конвертам. Какова вероятность того, что хотя бы одно письмо попадет в нужный конверт?

Пронумеруем письма и введем случайные события

$A = \{$хотя бы одно письмо попало в нужный конверт$\}$,

$A_i = \{i$ – е письмо попало в свой конверт$\}, i = 1, 2, \ldots, n$.

Ясно, что

$$A = A_1 + A_2 + \ldots + A_n. \tag{3.6}$$

Вероятность того, что первое письмо попадет в свой конверт, равна $1/n$ (один благоприятный исход из n равновозможных). Поэтому

$$P(A_1) = \frac{1}{n}.$$

Точно также для любого i:

$$P(A_i) = \frac{1}{n}. \tag{3.7}$$

Вероятность того, что второе письмо попадет в свой конверт при условии, что первое попало в свой конверт, равна $1/(n - 1)$ (один благоприятный исход из $(n-1)$ равновозможных). Итак,

$$P(A_2|A_1) = \frac{1}{n-1}.$$

По теореме умножения вероятностей

$$P(A_1A_2) = P(A_1)P(A_2|A_1) = \frac{1}{n}\cdot\frac{1}{n-1} = \frac{1}{n(n-1)}.$$

Точно так же при любых различных i, j

$$P(A_iA_j) = \frac{1}{n(n-1)}. \tag{3.8}$$

Вероятность того, что третье письмо попадет в свой конверт при условии, что и первое, и второе письма попали в свои конверты, равна $1/(n-2)$. Поэтому по теореме умножения вероятностей

$$P(A_1A_2A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) = \frac{1}{n(n-1)(n-2)}.$$

Точно так же при любых различных i, j, k:

$$P(A_iA_jA_k) = \frac{1}{n(n-1)(n-2)} \tag{3.9}$$

и т. д. Ввиду (3.6) по теореме сложения вероятностей

$$P(A) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \\ + \sum_{i<j<k} P(A_iA_jA_k) - \ldots + (-1)^{n+1}P(A_1A_2 \cdot \ldots \cdot A_n).$$

В первой сумме n одинаковых слагаемых (3.7), во второй сумме C_n^2 одинаковых слагаемых (3.8), в третьей сумме C_n^3 одинаковых слагаемых (3.9) и т. д. Поэтому

$$P(A) = n\frac{1}{n} - C_n^2\frac{1}{n(n-1)} + $$
$$+ C_n^3\frac{1}{n(n-1)(n-2)} - \ldots + (-1)^{n+1}\frac{1}{n(n-1)\cdot\ldots\cdot 2\cdot 1}.$$

Поскольку $C_n^2 = n(n-1)/2!, C_n^3 = n(n-1)(n-2)/3!$ и т. д., то

$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n+1}\frac{1}{n!}.$$

Вспомним, что при любом x

$$e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots$$

Поэтому при больших n

$$P(A) \approx 1 - e^{-1}.$$

§ 3.3. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ

Определение 3.2. Случайные события $A_1, A_2, \ldots, A_n$ составляют *полную группу событий*, если

1) объединение событий $A_1, A_2, \ldots, A_n$ совпадает с пространством элементарных исходов Ω;

2) события $A_1, A_2, \ldots, A_n$ попарно несовместны.

В случае выбора наудачу точки в квадрате K полная группа событий $A_1, \ldots, A_6$ показана на рис. 3.2.

Теорема 3.4. *Если случайные события $A_1, A_2, \ldots, A_n$ составляют полную группу событий, то для любого случайного события A справедлива формула полной вероятности*

$$P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + \ldots + P(A_n)P(A|A_n).$$

Доказательство. Из первого свойства полной группы событий следует, что

$$A = AA_1 + AA_2 + \ldots + AA_n,$$

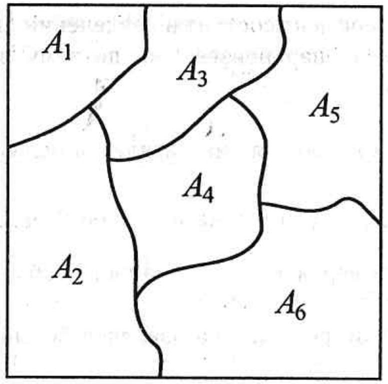


Рис. 3.2

а из второго — следует, что события $AA_1, AA_2, \ldots, AA_n$ попарно несовместны, поэтому по следствию 1.2

$$P(A) = P(AA_1) + P(AA_2) + \ldots + P(AA_n).$$

По теореме умножения вероятностей первая вероятность в правой части равна $P(A_1)P(A|A_1)$, вторая — $P(A_2)P(A|A_2), \ldots,$ последняя — $P(A_n)P(A|A_n)$, откуда следует утверждение теоремы.

Формула полной вероятности находит широкое применение в теории вероятностей. Пусть проводится такой двухступенчатый случайный эксперимент, что для проведения второго эксперимента нужно знать результат первого эксперимента. Если надо рассчитать вероятность случайного события, связанного со вторым экспериментом, то следует ввести полную группу событий $A_1, A_2, \ldots, A_n$, представляющих собою всевозможные различные исходы первого эксперимента, и воспользоваться формулой полной вероятности.

ПРИМЕР 3.3. Пусть в первом ящике имеется n_1 белых и m_1 черных шаров, а во втором — n_2 белых и m_2 черных шаров. Из первого ящика наудачу извлекается один шар и перекладывается во второй, а затем из второго ящика наудачу извлекается один шар. Найти вероятность того, что этот последний шар белый.

Первый эксперимент состоит в извлечении шара из первого ящика. Какой это шар неизвестно, поэтому введем полную группу событий:

$$A_1 = \{\text{из первого ящика извлечен белый шар}\},$$

$$A_2 = \{\text{из первого ящика извлечен черный шар}\}.$$

Требуется найти вероятность случайного события

$$A = \{\text{из второго ящика извлечен белый шар}\},$$

связанного со вторым экспериментом — извлечением шара из второго ящика. По формуле полной вероятности

$$P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2). \tag{3.10}$$

Очевидно,

$$P(A_1) = \frac{n_1}{n_1+m_1}, P(A_2) = \frac{m_1}{n_1+m_1}. \tag{3.11}$$

Если произошло событие A_1, то во втором ящике имеется (n_2+1) белых и m_2 черных шаров, следовательно,

$$P(A|A_1) = \frac{n_2+1}{n_2+1+m_2}. \tag{3.12}$$

Аналогично

$$P(A|A_2) = \frac{n_2}{n_2+m_2+1}. \tag{3.13}$$

Подставляя найденные выражения для вероятностей (3.11)–(3.13) в формулу (3.10), получаем, что

$$P(A) = \frac{n_1}{n_1+m_1}\cdot\frac{n_2+1}{n_2+m_2+1} + \frac{m_1}{n_1+m_1}\cdot\frac{n_2}{n_2+m_2+1} =$$

$$= \frac{n_1(n_2+1)+m_1n_2}{(n_1+m_1)(n_2+m_2+1)}.$$

§ 3.4. ФОРМУЛА БАЙЕСА

Пусть $A_1, A_2, ..., A_n$ — полная группа событий. Если наступило некоторое случайное событие A, то вместо первоначальных вероятностей $P(A_1), P(A_2), ..., P(A_n)$, которые называются априорными, разумнее рассматривать вероятности $P(A_1|A), P(A_2|A), ..., P(A_n|A)$, которые называются апостериорными.

Пусть, например, $A_1, A_2, ..., A_n$ — возможные виды заболеваний в данном регионе в данное время, тогда вероятности $P(A_1), P(A_2), ..., P(A_n)$ дают представление о частоте этих заболеваний. Но если к врачу приходит пациент с симптомами A, то, очевидно, при диагностике заболевания важнее рассматривать вероятности $P(A_1|A), P(A_2|A), ..., P(A_n|A)$, дающие представление о частоте заболеваний при данных симптомах. Как вычислять апостериорные вероятности?

Теорема 3.5. *Если $A_1, A_2, ..., A_n$ — полная группа событий, то справедлива формула Байеса*

$$P(A_k|A) = \frac{P(A_k)P(A|A_k)}{\sum_{i=1}^n P(A_i)P(A|A_i)}, \quad k = 1, 2, ..., n.$$

Доказательство. По определению условной вероятности

$$P(A_k|A) = \frac{P(AA_k)}{P(A)}.$$

Заменив числитель на $P(A_k)P(A|A_k)$ по теореме умножения вероятностей, а знаменатель на $\sum_{i=1}^n P(A_i)P(A|A_i)$ по формуле полной вероятности, получим утверждение теоремы.

ПРИМЕР 3.4. Рассмотрим случайный эксперимент, описанный в примере 3.3. Пусть известно, что из второго ящика извлекается белый шар, найти вероятность того, что из первого ящика был извлечен белый шар.

Воспользуемся обозначениями, введенными в примере 3.3. Нас интересует вероятность $P(A_1|A)$, которая может быть вычислена по формуле Байеса:

$$P(A_1|A) = \frac{P(A_1)P(A|A_1)}{P(A_1)P(A|A_1) + P(A_2)P(A|A_2)}.$$

Подставляя найденные в соотношениях (3.11)–(3.13) выражения для всех вероятностей в правой части, получаем, что

$$P(A_1|A) = \frac{n_1(n_2+1)}{n_1(n_2+1)+m_1n_2}.$$

Контрольные вопросы

1. Выпишите полностью формулу для вероятности $P(A_1 + A_2 + A_3)$ (см. теорему 3.2).
2. Обоснуйте замечание 3.3.
3. Обоснуйте определение 3.1 условной вероятности на примере модели выбора наудачу точки в квадрате.
4. Среди случайных событий, связанных со случайным экспериментом, выделяют, с одной стороны, элементарные исходы. С другой стороны, среди случайных событий можно выделить полную группу событий. Что общего между этими двумя процедурами, а в чем их отличие?
5. Поясните, какие вероятности в случае двухступенчатого эксперимента рассчитываются с помощью формулы полной вероятности, а какие с помощью формулы Байеса.

Ответы

1. $P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3)$.

Лекция 4

НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ ИСПЫТАНИЯ

§ 4.1. НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ СОБЫТИЯ

Какие случайные события A и B следует считать независимыми в вероятностном смысле? Осуществилось или нет событие B — это никак не должно сказываться на вероятности события A, и наоборот, т. е.

$$P(A|B) = P(A), \quad P(B|A) = P(B).$$

По определению условной вероятности это означает, что

$$\frac{P(AB)}{P(B)} = P(A), \quad \frac{P(AB)}{P(A)} = P(B) \Rightarrow P(AB) = P(A)P(B).$$

Определение 4.1. Случайные события A и B называются *независимыми*, если $P(AB) = P(A)P(B)$.

Определение 4.2. Случайные события $A_1, A_2, \ldots, A_n$ называются *независимыми*, если $P(A_1A_2 \cdot \ldots \cdot A_n) = P(A_1) \times P(A_2) \cdot \ldots \cdot P(A_n)$ и аналогичное равенство выполняется для любой части этой совокупности событий. В частности, три события A, B, C называются независимыми, если

$$P(ABC) = P(A)P(B)P(C),$$

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C),$$

$$P(BC) = P(B)P(C).$$

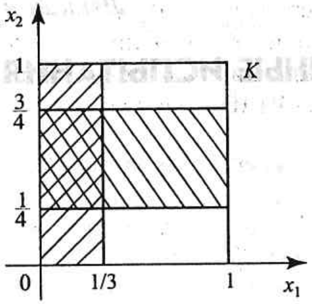


Рис. 4.1

ПРИМЕР 4.1. Пусть точка наудачу выбирается в квадрате K (рис. 4.1).

Тогда случайные события A_1 и A_2, связанные соответственно с первой и второй координатами этой точки, являются независимыми. Пусть, например,

$$A_1 = \{\text{первая координата меньше } 1/3\},$$

$$A_2 = \{\text{вторая координата принадлежит } [1/4, 3/4]\},$$

т. е.

$$A_1 = \{(x_1, x_2): \ 0 \leqslant x_1 < 1/3, \ 0 \leqslant x_2 \leqslant 1\},$$
$$A_2 = \{(x_1, x_2): \ 0 \leqslant x_1 \leqslant 1, 1/4 \leqslant x_2 \leqslant 3/4\}.$$

Тогда $S(A_1) = 1/3$, $S(A_2) = 1/2$. Очевидно,

$$A_1A_2 = \{(x_1, x_2): \ 0 \leqslant x_1 \leqslant 1/3, \ 1/4 \leqslant x_2 \leqslant 3/4\}$$

и $S(A_1A_2) = (1/3) \cdot (1/2)$. Вспоминая, как вычисляются геометрические вероятности, получаем, что

$$P(A_1) = 1/3, \ P(A_2) = 1/2, \ P(A_1A_2) = 1/6 \Rightarrow$$
$$\Rightarrow P(A_1A_2) = P(A_1)P(A_2),$$

таким образом, случайные события A_1 и A_2 являются независимыми.

Тот факт, что случайные события $A_1, A_2, ..., A_n$ являются независимыми, позволяет вычислять вероятность любого случайного события, полученного из $A_1, A_2, ..., A_n$ в результате применения операций объединения, пересечения, нахождения разности, зная лишь вероятности событий $A_1, A_2, ..., A_n$. В частности, справедлив следующий результат.

Теорема 4.1. *Если $A_1, A_2, ..., A_n$ — независимые случайные события, то*

$$P(A_1 + A_2 + ... + A_n) = 1 - \prod_{i=1}^{n}(1 - P(A_i)).$$

Доказательство. Событие, противоположное к событию $A_1 + A_2 + ... + A_n$, означает, что не произошло ни одно из событий $A_1, A_2, ..., A_n$, т. е. произошли сразу все события $\bar{A}_1, \bar{A}_2, ..., \bar{A}_n$. Поэтому по следствию 1.1 (см. лекцию № 1)

$$P(A_1 + A_2 + ... + A_n) = 1 - P(\bar{A}_1\bar{A}_2 \cdot ... \cdot \bar{A}_n). \tag{4.1}$$

Если события $A_1, A_2, ..., A_n$ независимые, то и противоположные события $\bar{A}_1, \bar{A}_2, ..., \bar{A}_n$ независимые, поэтому

$$P(\bar{A}_1\bar{A}_2 \cdot ... \cdot \bar{A}_n) = P(\bar{A}_1)P(\bar{A}_2) \cdot ... \cdot P(\bar{A}_n) = \prod_{i=1}^{n}(1 - P(A_i)). \tag{4.2}$$

Из соотношений (4.1) и (4.2) следует утверждение теоремы.

ПРИМЕР 4.2. Вероятность невыхода из строя в течение времени T i-го элемента цепи, указанной на рис. 4.2, равна p_i. Все элементы цепи функционируют независимо друг от друга. Найти вероятность того, что вся цепь не выйдет из строя в течение времени T.

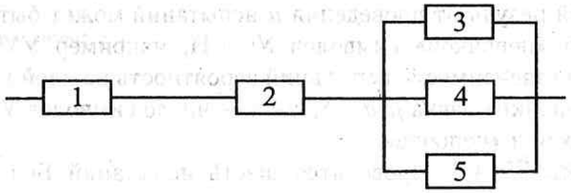


Рис. 4.2

Введем случайные события

$$A = \{\text{вся цепь не выйдет из строя в течение времени } T\},$$

$$A_i = \{i\text{-й элемент не выйдет из строя в течение времени } T\},$$
$$i = 1, ..., 5.$$

Тогда

$$A = A_1A_2(A_3 + A_4 + A_5).$$

Поскольку случайные события $A_1, A_2, ..., A_5$ независимые, то независимы события A_1, A_2 и $A_3 + A_4 + A_5$. По определению независимых событий

$$P(A) = P(A_1)P(A_2)P(A_3 + A_4 + A_5). \qquad (4.3)$$

Ясно, что $P(A_i) = p_i$, $i = 1, \ldots, 5$. По теореме 4.1

$$P(A_3+A_4+A_5) = 1-(1-P(A_3))(1-P(A_4))(1-P(A_5)). \quad (4.4)$$

Из формул (4.3) и (4.4) находим искомую вероятность:

$$P(A) = p_1p_2[1 - (1 - p_3)(1 - p_4)(1 - p_5)].$$

§ 4.2. ИСПЫТАНИЯ БЕРНУЛЛИ

Так называются независимые испытания с двумя исходами: успех (У) и неудача (Н), причем вероятность успеха в каждом испытании одна и та же p (тогда вероятность неудачи в каждом испытании $q = 1 - p$). Испытаниями Бернулли являются последовательные выстрелы по мишени (успех — попадание, неудача — промах), покупка лотерейных билетов (успех — выигрышный билет, неудача — проигрышный) и т. п. Каждый результат проведения n испытаний может быть представлен «цепочкой» символов У и Н, например УУНУ...Н. Ввиду независимости испытаний вероятность каждой конкретной «цепочки» равна p^kq^{n-k}, где k — число символов У, встречающихся в «цепочке».

ПРИМЕР 4.3. Проводятся шесть испытаний Бернулли с вероятностью успеха p. Найти вероятность того, что успех и неудача чередуются.

Подходят только две «цепочки» УНУНУН и НУНУНУ. Вероятность каждой из них равна p^3q^3. Поскольку эти «цепочки» не могут реализоваться одновременно, то искомая вероятность равна $2p^3q^3$.

Теорема 4.2. *Пусть μ_n означает число успехов в n испытаниях Бернулли с вероятностью успеха p. Справедлива формула Бернулли:*

$$P(\mu_n = k) = C_n^k\, p^k\, q^{n-k}, \quad k = 0, 1, \ldots, n.$$

Доказательство. Рассмотрим «цепочки», в которых символ У встречается k раз, а символ Н соответственно $(n - k)$ раз. Вероятность каждой такой «цепочки» равна p^kq^{n-k}. Число таких «цепочек» равно числу способов выбрать k мест из n возможных (в «цепочке» длины n) для символа У. Это число равно C_n^k. Разные «цепочки» представляют собой попарно несовместимые события, поэтому интересующая нас вероятность равна $C_n^k\, p^k\, q^{n-k}$, что и требовалось доказать.

Если n велико, то вычисления по формуле Бернулли слишком трудоемки, поэтому для нахождения $P(\mu_n = k)$ применяются различные приближенные формулы в зависимости от соотношения между n, p и k.

Теорема 4.3 (Пуассон). *Пусть $n \to \infty$, $p \to 0$ так, что $np \to a$, причем $a \in (0, +\infty)$. Тогда*

$$P(\mu_n = k) \to \frac{a^k}{k!}e^{-a}$$

при любом фиксированном k, $k = 0, 1, 2, \ldots$.

Доказательство. Положим $a_n = np$. Тогда по условию теоремы

$$\lim_{n\to\infty} a_n = a. \qquad (4.5)$$

По теореме 4.2

$$P(\mu_n = k) = C_n^k\, p^k\, q^{n-k}.$$

Заменим p на a_n/n, q на $(1 - a_n/n)$ и вспомним, что

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot \ldots \cdot (n-k+1)}{k!}.$$

Тогда

$$P(\mu_n = k) = \frac{n(n-1)\cdot \ldots \cdot (n-k+1)}{k!}\left(\frac{a_n}{n}\right)^k\left(1 - \frac{a_n}{n}\right)^{n-k} =$$

$$= \frac{1(1-1/n)\cdot \ldots \cdot (1-(k-1)/n)}{k!}a_n^k\left(1 - \frac{a_n}{n}\right)^{n-k}. \qquad (4.6)$$

При фиксированном k

$$\lim_{n\to\infty} \frac{1(1-1/n)\cdot \ldots \cdot (1-(k-1)/n)}{k!} = \frac{1}{k!}.$$

Ввиду (4.5)

$$\lim_{n\to\infty} a_n^k = a^k.$$

Вспоминая, как находятся пределы «типа e», получаем, что

$$\lim_{n\to\infty}\left(1-\frac{a_n}{n}\right)^{n-k} = e^{-\lim\limits_{n\to\infty}(n-k)a_n/n} = e^{-a}.$$

Следовательно, предел правой части соотношения (4.6) равен $a^k e^{-a}/k!$, что и требовалось доказать.

Замечание 4.1. Приближенную формулу

$$P(\mu_n = k) \approx \frac{a^k}{k!}e^{-a},$$

где $a = np$, целесообразно использовать, когда

$$n \gg k^2, \quad np^2 \ll 1, \quad kp \ll 1. \tag{4.7}$$

ПРИМЕР 4.4. Предполагая рождение ребенка в любой день года равновозможным, найти вероятность того, что в группе из 200 человек ровно трое родились 1 января.

Если ребенок родился 1 января, то будем считать это успехом, а в противоположном случае — неудачей. Вероятность успеха равна $p = 1/365$. Число испытаний равно $n = 200$. Нас интересует вероятность того, что число успехов равно $k = 3$. Условия (4.7) выполнены, поэтому

$$P(\mu_{200} = 3) \approx \frac{a^3}{3!}e^{-a}, \tag{4.8}$$

где $a = 200/365$. Вычисляя правую часть (4.8) с помощью калькулятора, получаем, что

$$P(\mu_{200} = 3) \approx 0,0159.$$

§ 4.3. ПРОСТЕЙШИЙ ПОТОК СОБЫТИЙ

Пусть мы наблюдаем за временем наступления некоторых однородных событий (телефонные звонки в какое-нибудь учреждение, дорожно-транспортные происшествия в каком-нибудь городе и т. п.). Поток событий называется *простейшим,* если он обладает тремя свойствами:

1) вероятностные характеристики потока событий на произвольном временном промежутке $[t_1, t_2]$ зависят от длины промежутка, но не зависят от его начала;

2) вероятность наступления более одного события за время Δt есть $o(\Delta t)$ при $\Delta t \to 0$;

3) поведение потока событий на произвольном временном промежутке $[t_1, t_2]$ не зависит (в вероятностном смысле) от его поведения до момента времени t_1.

Обозначим $P_k(t)$ вероятность наступления ровно k событий на временном промежутке длины t. Можно показать, что

$$P_1(\Delta t) = \lambda\Delta t + o(\Delta t), \tag{4.9}$$

где λ — положительная постоянная, называемая *интенсивностью* простейшего потока.

Теорема 4.4. *Если поток является простейшим с интенсивностью λ, то*

$$P_k(t) = \frac{a^k}{k!}e^{-a},$$

где $a = \lambda t$.

Доказательство. Разобьем промежуток длины t на n равных частей длины $\Delta t = t/n$. В силу свойства 2 простейшего потока можно считать, что на каждом из получившихся маленьких промежутков может произойти либо одно событие (будем считать это успехом), либо ни одного события (неудача). Таким образом, имеем n испытаний, причем в силу свойства 3 простейшего потока эти испытания независимые. Вероятность успеха ввиду (4.9) равна $p = (\lambda t)/n + o(1/n)$. Если устремить n в бесконечность, то $p \to 0$, $np \to \lambda t$ и, значит, в силу теоремы 4.3 вероятность того, что будет ровно k успехов в n испытаниях, стремится к $a^k e^{-a}/k!$, где $a = \lambda t$. Теорема доказана.

Замечание 4.2. Позже покажем, что интенсивность простейшего потока λ совпадает со средним числом событий за единицу времени.

ПРИМЕР 4.5. В некотором городе среднее число дорожно-транспортных происшествий за сутки равно 5. Найти вероятность того, что за двое суток будет 10 дорожно-транспортных происшествий.

Единица времени — сутки. Интенсивность простейшего потока равна $\lambda = 5$. Нас интересует $P_{10}(2)$. По теореме 4.4

$$P_{10}(2) = \frac{10^{10}}{10!}e^{-10} \approx 0,1251.$$

Контрольные вопросы

1. Вспомните определение несовместных событий и сравните его с определением независимых событий.
2. Покажите, что если A и B — независимые события, причем $P(A) \neq 0$, $P(B) \neq 0$, то A и B — совместные события.
3. Пусть A и B — независимые события. Найдите $P(A \backslash B)$.
4. Можно ли найти $P(A \backslash B)$, зная лишь вероятности случайных событий A и B, и без предположения о независимости A и B?
5. Покажите, что если A, B, C — независимые события, то $A + B$ и C — независимые события.
6. Приведите пример потока событий, не являющегося простейшим.

Ответы

2. $P(AB) = P(A)P(B) \neq 0 \Rightarrow AB \neq \phi$.
3. $P(A \backslash B) = P(A)(1 - P(B))$.
4. Нельзя.
5. $P((A+B)C) = P(AC+BC) = P(AC)+P(BC)-P(ABC) = P(A)P(C) + P(B)P(C) - P(A)P(B)P(C) = \big(P(A) + P(B) - P(AB)\big)P(C) = P(A+B)P(C)$.

Лекция 5

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 5.1. ОПРЕДЕЛЕНИЕ СЛУЧАЙНОЙ ВЕЛИЧИНЫ И ЕЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

Каждый случайный эксперимент завершается элементарным исходом ω. Лучше всего итог эксперимента подводить в виде числа, т. е. вместо ω рассматривать число $\xi(\omega)$.

Определение 5.1. Случайной величиной называется числовая функция от элементарного исхода ω, заданная на пространстве элементарных исходов Ω :

$$\xi = \xi(\omega), \quad \omega \in \Omega.$$

В математическом анализе в первую очередь исследуется зависимость значения функции от значения аргумента. В теории вероятностей эта зависимость играет не первую роль, важно знать вероятностные характеристики случайной величины, например вероятности следующих случайных событий: $\{\omega : \xi(\omega) \in [a, b]\}$, $\{\omega : \xi(\omega) > a\}$ и т. п. (здесь a, b — постоянные).

Определение 5.2. Функцией распределения случайной величины ξ называется числовая функция $F(x)$, определяемая равенством

$$F(x) = P(\xi \leqslant x), \quad x \in R.$$

Здесь для краткости написано $P(\xi \leqslant x)$ вместо $P(\{\omega : \xi(\omega) \leqslant x\})$. Множество $\{\omega : \xi(\omega) \leqslant x\}$ может и не принадлежать алгебре случайных событий $\mathcal{F}$, значит, вероятность этого множества может быть не определена. Поэтому будем предполагать, что если ξ — случайная величина, то для любого $x \in R$ случайное событие $\{\omega : \xi(\omega) \leqslant x\}$ принадлежит алгебре случайных событий $\mathcal{F}$.

ПРИМЕР 5.1. Пусть точка наудачу выбирается в треугольнике D (рис. 5.1) и ξ — первая координата этой точки. Найти функцию распределения случайной величины ξ.

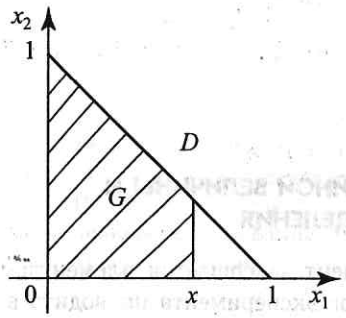


Рис. 5.1

Элементарные исходы — это точки (x_1, x_2), принадлежащие D. По определению $\xi(x_1, x_2) = x_1$. Рассмотрим $x \in [0,1]$. Случайному событию $\{\xi \leqslant x\}$ благоприятствуют элементарные исходы $\{(x_1, x_2) : x_1 \leqslant x\}$ (заштрихованная область G на рис. 5.1).

По определению геометрических вероятностей

$$P(\xi \leqslant x) = \frac{S(G)}{S(D)} = \frac{1/2-(1-x)^2/2}{1/2} = 1-(1-x)^2 = 2x - x^2.$$

Итак, при $x \in [0,1]$ $F(x) = 2x - x^2$. При $x \in (-\infty, 0)$, очевидно, событие $\{\xi \leqslant x\}$ является невозможным и его вероятность равна 0, следовательно, $F(x) = 0$. При $x \in (1, +\infty)$ событие $\{\xi \leqslant x\}$ является достоверным и его вероятность равна 1, следовательно, $F(x) = 1$. Таким образом,

$$F(x) = \begin{cases} 0, & x \in (-\infty, 0); \\ 2x - x^2, & x \in [0,1]; \\ 1, & x \in (1, +\infty). \end{cases}$$

§ 5.2. СВОЙСТВА ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

Для изучения свойств функций распределения потребуется еще одна аксиома вероятностного пространства.

Аксиома непрерывности. 1) Если случайные события $A_1, A_2, A_3, \ldots$ таковы, что $A_1 \supset A_2 \supset A_3 \supset \ldots$, то множество $\bigcap_{i=1}^{\infty} A_i$ (состоит из элементарных исходов, принадлежащих сразу всем A_i) является случайным событием и

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i).$$

2) Если случайные события $A_1, A_2, A_3, \ldots$ таковы, что $A_1 \subset A_2 \subset A_3 \subset \ldots$, то множество $\bigcup_{i=1}^{\infty} A_i$ (состоит из элементарных исходов, принадлежащих хотя бы одному A_i) является случайным событием и

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i).$$

Свойство 1. *Функция распределения $F(x)$ не убывает на всей числовой прямой.*

Доказательство. Пусть $x_2 > x_1$. Ясно, что

$$\{\xi \leqslant x_2\} = \{\xi \leqslant x_1\} + \{\xi \in (x_1, x_2]\},$$

причем последние два случайных события являются несовместными, поэтому по аксиоме $P3$

$$P(\xi \leqslant x_2) = P(\xi \leqslant x_1) + P(\xi \in (x_1, x_2]). \tag{5.1}$$

Поскольку $P(\xi \in (x_1, x_2]) \geqslant 0$, отсюда следует, что

$$P(\xi \leqslant x_2) \geqslant P(\xi \leqslant x_1) \;\Rightarrow\; F(x_2) \geqslant F(x_1).$$

Свойство 2. $\lim\limits_{x\to-\infty} F(x) = 0$.

Доказательство. Введем случайные события $A_i = \{\xi \leqslant -i\}$, $i = 1, 2, \ldots$ Ясно, что $A_1 \supset A_2 \supset A_3 \supset \ldots$ и $\bigcap_{i=1}^{\infty} A_i = \phi$, поэтому по аксиоме непрерывности

$$0 = P(\phi) = P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i) = \lim_{i\to\infty} P(\xi \leqslant -i) = \lim_{i\to\infty} F(-i).$$

Откуда, учитывая свойство 1, получаем требуемое утверждение.

Свойство 3. $\lim_{x\to\infty} F(x) = 1$.

Устанавливается аналогично свойству 2.

Свойство 4. Функция распределения $F(x)$ непрерывна справа на всей числовой оси. Если в точке x_0 функция $F(x)$ имеет разрыв, то он является разрывом первого рода и

$$P(\xi = x_0) = F(x_0) - F(x_0 - 0). \tag{5.2}$$

Доказательство. Сначала рассмотрим произвольное x и случайные события $A_i = \{\xi \leqslant x + 1/i\}$, $i = 1, 2, \ldots$ Ясно, что $A_1 \supset A_2 \supset A_3 \supset \ldots$ и $\bigcap_{i=1}^{\infty} A_i = \{\xi \leqslant x\}$. Поэтому по аксиоме непрерывности

$$F(x) = P(\xi \leqslant x) = P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i) =$$
$$= \lim_{i\to\infty} P(\xi \leqslant x + 1/i) = \lim_{i\to\infty} F(x + 1/i).$$

Откуда, опять учитывая свойство 1, получаем, что $F(x)$ непрерывна справа.

Для доказательства (5.2) рассмотрим случайные события $A_i = \{\xi \in (x_0 - 1/i, x_0]\}$, $i = 1, 2, \ldots$ Ясно, что $A_1 \supset A_2 \supset A_3 \supset \ldots$ и $\bigcap_{i=1}^{\infty} A_i = \{\xi = x_0\}$, и в силу аксиомы непрерывности и соотношения (5.1):

$$P(\xi = x_0) = P\left(\bigcap_{i=1}^{\infty} A_i\right) =$$
$$= \lim_{i\to\infty} P(A_i) = \lim_{i\to\infty} P(\xi \in (x_0 - 1/i, x_0]) =$$
$$= \lim_{i\to\infty} [P(\xi \leqslant x_0) - P(\xi \leqslant x_0 - 1/i)] =$$
$$= F(x_0) - \lim_{i\to\infty} F(x_0 - 1/i) = F(x_0) - F(x_0 - 0).$$

Замечание 5.1. Зная функцию распределения, можно найти вероятности $P(\xi \in B)$ для различных множеств B на числовой оси. В частности,

$$P(\xi > x) = 1 - F(x),$$
$$P(\xi \in (a, b]) = F(b) - F(a),$$
$$P(\xi \in [a, b]) = F(b) - F(a - 0),$$
$$P(\xi \in (a, b)) = F(b - 0) - F(a).$$

Из свойств функции распределения следует, что типичный график $F(x)$ имеет вид, изображенный на рис. 5.2.

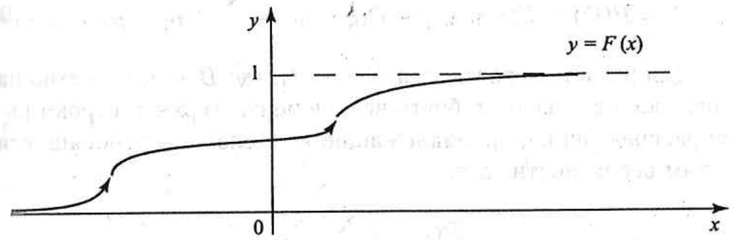


Рис. 5.2

§ 5.3. ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Определение 5.3. Случайная величина ξ называется *дискретной,* если она принимает конечное или счетное число значений.

Определение 5.4. Рядом распределения дискретной случайной величины ξ называется таблица, в верхней строке которой указываются возможные значения $a_1, a_2, \ldots, a_n, \ldots$ случайной величины ξ, а в нижней — вероятности $p_1, p_2, \ldots, p_n, \ldots$, с которыми эти значения принимает случайная величина ($p_n = P(\xi = a_n)$) :

ξ	a_1	a_2	...	a_n	...
P	p_1	p_2	...	p_n	...

Очевидно, $p_n \geqslant 0$ для любого n. Кроме того,

$$\sum_{n=1}^{\infty} p_n = 1.$$

Действительно (ограничимся случаем конечного числа возможных значений),

$$\Omega = \{\xi = a_1\} \cup \{\xi = a_2\} \cup \ldots,$$

причем все случайные события справа попарно несовместны, поэтому по следствию 1.2 (см. лекцию № 1)

$$1 = P(\Omega) = P(\xi = a_1) + P(\xi = a_2) + \ldots = p_1 + p_2 + \ldots$$

Для расчета вероятности $P(\xi \in B)$, где B — множество на числовой оси, надо отобрать все элементы верхней строки ряда распределения, принадлежащие B, и сложить относящиеся к ним вероятности, т. е.

$$P(B) = \sum_{\substack{n:\\ a_n \in B}} p_n.$$

ПРИМЕР 5.2. Пусть бросаются два игральных кубика и ξ — суммарное число выпавших очков. Найти ряд распределения ξ и по нему найти $P(\xi \leqslant 4)$.

Возможные значения ξ есть $2, 3, \ldots, 12$. Элементарные исходы — это пары чисел (n_1, n_2), где n_1 — число выпавших очков на первом кубике, а n_2 — на втором. Тогда

$$\xi(n_1, n_2) = n_1 + n_2.$$

Очевидно, $\{\xi = 2\} = \{(n_1, n_2) : n_1 + n_2 = 2\} = \{(1, 1)\}$. Поэтому

$$P(\xi = 2) = \frac{1}{36}.$$

Аналогично, $\{\xi = 3\} = \{(1, 2), (2, 1)\} \Rightarrow P(\xi = 3) = 2/36 = 1/18$ и т. д. В итоге получаем ряд распределения

ξ	2	3	4	5	6	7	8	9	10	11	12
P	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

Событие $\{\xi \leqslant 4\}$ осуществляется, если ξ принимает значения $2, 3, 4$. Складывая соответствующие вероятности, получаем, что

$$P(\xi \leqslant 4) = \frac{1}{36} + \frac{1}{18} + \frac{1}{12} = \frac{6}{36} = \frac{1}{6}.$$

В схеме испытаний Бернулли и в простейшем потоке используются следующие важные дискретные распределения.

Определение 5.5. Говорят, что случайная величина ξ имеет *биномиальное распределение* с параметрами p, n, где $p \in (0, 1)$, n — натуральное число, если ξ принимает значения $0, 1, \ldots, n$ и

$$P(\xi = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n \tag{5.3}$$

(здесь $q = 1 - p$).

Случайная величина μ_n, равная числу успехов в n испытаниях Бернулли, имеет биномиальное распределение.

Определение 5.6. Говорят, что случайная величина ξ имеет *распределение Пуассона* с параметром a, если ξ принимает значения $0, 1, 2, \ldots$ и

$$P(\xi = k) = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \ldots \tag{5.4}$$

Случайная величина, равная числу событий в простейшем потоке за время t, имеет распределение Пуассона с параметром λt, где λ — интенсивность простейшего потока.

§ 5.4. НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Определение 5.7. Случайная величина ξ называется *непрерывной*, если она принимает все значения из некоторого промежутка числовой оси и при этом существует такая неотрицательная числовая интегрируемая функция $p(x)$, что для любого отрезка $[a, b]$ (конечного или бесконечного):

$$P(\xi \in [a, b]) = \int_a^b p(x)dx. \tag{5.5}$$

Эта функция $p(x)$ называется *плотностью вероятностей.*

Итак, $p(x) \geqslant 0$ при любом x и ($a = -\infty, b = +\infty$) выполняется равенство:

$$\int_{-\infty}^{+\infty} p(x)dx = 1.$$

Имеется простая связь между плотностью вероятностей и функцией распределения:

а) $F(x) = \int\limits_{-\infty}^{x} p(u)du$;

б) $p(x) = F'(x)$ в точках непрерывности $p(x)$.

Свойство а) следует из соотношения (5.5) ($a = -\infty$, $b = x$). Свойство б) вытекает из теоремы о дифференцировании интеграла по переменному верхнему пределу.

ПРИМЕР 5.3. Найти плотность вероятностей случайной величины ξ из примера 5.1.

Поскольку $p(x) = F'(x)$ в точках непрерывности $p(x)$, то при $x \in (0, 1)$:

$$p(x) = F'(x) = (2x - x^2)' = 2 - 2x.$$

При $x \in (-\infty, 0) \cup (1, +\infty)$ $p(x) = F'(x) = 0$. Значения $p(x)$ в точках $x = 0$ и $x = 1$ не играют никакой роли, поскольку они не отразятся на вероятностях $P(\xi \in [a, b])$ из соотношения (5.5).

Очевидно, что если ξ — непрерывная случайная величина и $\triangle x$ мало, то

$$P(\xi \in [x, x + \triangle x]) \approx p(x)\triangle x. \qquad (5.6)$$

Отсюда понятен смысл названия $p(x)$ (вспомните, для того чтобы найти массу маленького отрезка $[x, x + \triangle x]$ прямой с плотностью массы $\rho(x)$, надо $\rho(x)$ умножить на $\triangle x$). Соотношение (5.6) удобно использовать для приближения непрерывной случайной величины ξ дискретной случайной величиной. Пусть ξ измеряется некоторым прибором с ценой деления $\triangle x$. В результате измерения вместо ξ получаем новую дискретную случайную величину $\hat{\xi}$, принимающую значения $x_1, x_2, \ldots$, причем $\triangle x_i = x_{i+1} - x_i = \triangle x$. Будем считать, что $\hat{\xi} = x_i$, если $\xi \in [x_i, x_{i+1})$. Тогда

$$P(\hat{\xi} = x_i) = P(\xi \in [x_i, x_{i+1})) \approx p(x_i)\triangle x.$$

Итак, непрерывная случайная величина ξ близка к дискретной случайной величине $\hat{\xi}$ с рядом распределения

$\hat{\xi}$	x_1	x_2	...
P	$p(x_1)\triangle x$	$p(x_2)\triangle x$	...

Дадим определение некоторых важных непрерывных распределений.

Определение 5.8. Говорят, что случайная величина ξ имеет *равномерное распределение* на отрезке $[a, b]$ (краткая запись: $\xi \sim U[a, b]$), если ξ непрерывна и

$$p(x) = \begin{cases} 1/(b-a), x \in [a, b]; \\ 0, \quad x \notin [a, b]. \end{cases}$$

Определение 5.9. Говорят, что случайная величина ξ имеет *показательное распределение* с параметром $\lambda \in (0, +\infty)$ (краткая запись: $\xi \sim \exp(\lambda)$), если ξ непрерывна и

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, x \geqslant 0; \\ 0, \quad x < 0. \end{cases}$$

Определение 5.10. Говорят, что случайная величина ξ имеет *нормальное распределение* с параметрами $a, \sigma^2 \neq 0$ (краткая запись: $\xi \sim N(a; \sigma^2)$), если ξ непрерывна и

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2/2\sigma^2}, \quad x \in (-\infty, +\infty).$$

Графики указанных плотностей распределения изображены на рис. 5.3.

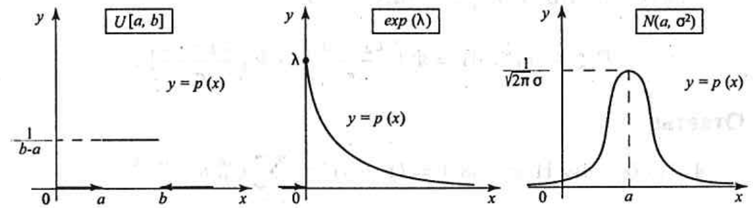


Рис. 5.3

Равномерное распределение используется в методе приближенных вычислений, известном под названием метод Монте-Карло. Показательному распределению подчинено время службы электротехнических приборов. Нормальному распределению подчинены такие случайные величины, как ошибка результата измерения какой-нибудь физической величины, отклонение снаряда от цели, рост и вес человека и т. п.

Контрольные вопросы

1. Докажите соотношения из замечания 5.1.

2. Почему на ваш взгляд понятие ряда распределения для дискретной случайной величины удобнее понятия функции распределения?

3. Как выглядит график функции распределения дискретной случайной величины?

4. Докажите непосредственно, что сумма всех вероятностей из формулы (5.3) действительно равна 1.

5. Докажите непосредственно, что сумма всех вероятностей из формулы (5.4) действительно равна 1.

6. Пусть $p(x)$ — плотность вероятностей некоторой случайной величины. О чем говорит неравенство $p(x_2) > p(x_1)$?

7. Пусть точка наудачу выбирается на отрезке $[a, b]$ и ξ — координата этой точки. Докажите, что $\xi \sim U[a, b]$.

8. Пусть τ — время ожидания первого события в простейшем потоке с интенсивностью λ. Докажите, что $\tau \sim \exp(\lambda)$.

9. Введем функцию Лапласа

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du.$$

Докажите, что если $\xi \sim N(a; \sigma^2)$, то

$$P(\xi \in [x_1, x_2]) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right).$$

Ответы

4. По биному Ньютона $1 = (p+q)^n = \sum\limits_{k=0}^{n} C_n^k p^k q^{n-k}$.

5. Поскольку $e^a = \sum\limits_{k=0}^{\infty} \frac{a^k}{k!}$, то $\sum\limits_{k=0}^{\infty} \frac{a^k}{k!} e^{-a} = 1$.

8. Случайное событие $\{\tau > t\}$ означает, что за время t не произошло ни одного события в простейшем потоке. Поэтому при $t \in [0, +\infty)$ $P(\tau > t) = P_0(t) = e^{-\lambda t} \Rightarrow P(\tau \leqslant t) = 1 - e^{-\lambda t}$. Чтобы найти плотность вероятностей $p(t)$, надо продифференцировать функцию распределения: $p(t) = (1 - e^{-\lambda t})' = \lambda e^{-\lambda t}$, $t \in [0, +\infty)$.

9. По определению плотности вероятностей

$$P(\xi \in [x_1, x_2]) = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2/2\sigma^2} dx.$$

Делая замену $\dfrac{x-a}{\sigma} = u$, получаем

$$P(\xi \in [x_1, x_2]) = \int_{(x_1-a)/\sigma}^{(x_2-a)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \Phi\left(\frac{x_2 - u}{\sigma}\right) - \Phi\left(\frac{x_1 - u}{\sigma}\right).$$

является квантиль $u_{1-\varepsilon}$ уровня $(1-\varepsilon)$ t-распределения с $(n+m-2)$ степенями свободы. Итак,

$$\sqrt{\frac{n+m-2}{(1/n)+(1/m)}}\cdot r = u_{1-\varepsilon} \Rightarrow r = \sqrt{\frac{(1/n)+(1/m)}{n+m-2}}\cdot u_{1-\varepsilon}.$$

Если $t \in A_0$, то принимается гипотеза H_0; если же $t \in A_1$, то принимается альтернатива H_1.

Замечание 15.1. Во всех разобранных в этой лекции примерах рассмотренные альтернативы можно заменить на альтернативы другого вида, как это было сделано в конце предыдущей лекции. Соответствующие изменения в построении критериев очевидны.

Контрольные вопросы

1. Опишите критерий проверки гипотезы о значении математического ожидания при известной дисперсии для произвольной выборки большого объема (не обязательно нормальной).

2. Опишите критерий проверки гипотезы о значении математического ожидания при неизвестной дисперсии в случае нормальной выборки. Можно ли этот критерий использовать в случае, когда дисперсия известна?

3. Пусть в случае нормальной выборки проверяется гипотеза H_0: $DX = \sigma_0^2$ против альтернативы H_1: $DX < \sigma_0^2$. Как будет выглядеть область принятия гипотезы H_0, если статистика критерия есть s^2?

4. Почему при проверке гипотезы о совпадении дисперсий двух нормальных выборок важно предположение о независимости этих выборок?

5. Объясните, почему статистики ξ и η, используемые при проверке гипотезы о совпадении математических ожиданий двух нормальных выборок, являются независимыми?

Ответы

3. Правая полуось, включающая точку σ_0^2.

5. Пары статистик $(\bar{x}, s_1^2)$ и $(\bar{y}, s_2^2)$ являются независимыми, причем элементы этих пар сами не зависят друг от друга, поэтому статистики $\bar{x}$, s_1^2, $\bar{y}$, s_2^2 являются независимыми. Следовательно, статистика ξ, образованная из $\bar{x}$ и $\bar{y}$, и статистика η, образованная из s_1^2 и s_2^2, являются независимыми.

ПРАКТИЧЕСКИЕ ЗАНЯТИЯ

Занятие 1

КОМБИНАТОРИКА

Пусть имеется несколько множеств элементов:

$$\{a_1, a_2, \ldots, a_t\}, \quad \{b_1, b_2, \ldots, b_s\}, \quad \ldots, \quad \{c_1, c_2, \ldots, c_k\}, \quad \ldots .$$

Сколькими способами можно составить новое множество $\{a, b, c, \ldots\}$, взяв из каждого исходного множества по одному элементу? Ответ на этот вопрос дает *основной комбинаторный принцип*:

Если некоторый первый выбор можно сделать t способами, для каждого первого выбора некоторый второй можно сделать s способами, для каждой пары первых двух — третий выбор можно сделать k способами и так далее, то число способов для последовательности таких выборов равно $t \cdot s \cdot k \cdot \ldots$.

Комбинаторные формулы в прикладных задачах теории вероятностей обычно связывают с выбором r элементов (*«выборкой объема r»*) из совокупности, состоящей из n элементов (элементов *«генеральной совокупности»*). Различают два способа выбора:

а) *повторный*, при котором выбранный элемент возвращается в генеральную совокупность и может быть выбран вновь;

б) *бесповторный*, при котором выбранный элемент в совокупность не возвращается и выборка не содержит повторяющихся элементов.

При повторном выборе выборку объема r можно сделать n^r способами. Например, повторную выборку объемом два из трех элементов $\{а,в,с\}$ можно сделать $3^2 = 9$ способами: *аа, ав, ва, вв, вс, св, ас, са, сс*.

При бесповторном выборе выборку объема r можно сделать $A_n^r = n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdot \ldots \cdot (n-r+1)$ способами. Число A_n^r называют *числом размещений* из n элементов по r. Размещения отличаются либо *составом* элементов, либо *порядком* их расположения. Например, размещений из трех элементов $\{a,в,c\}$ по два можно составить $A_3^2 = 3 \cdot 2 = 6$: *ав, ва, ас, са, вс, св*.

Выборки объема r, которые отличаются друг от друга только составом, можно сделать $C_n^r = \dfrac{n!}{r!(n-r)!}$ способами. Число C_n^r называют *числом сочетаний* из n элементов по r. Например, сочетаний из трех элементов $\{a,в,c\}$ по два существует $C_3^2 = 3$: *ав, ас, вс*.

При повторном выборе из n элементов число выборок объема r, которые отличаются только составом равно C_{n+r-1}^r.

Число перестановок из n элементов равно $n! = n \times \times (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$.

Совокупность из n элементов разделить на m групп по $k_1, k_2, \ldots, k_m$ элементов соответственно $(k_1 + k_2 + \ldots + k_m = n)$ можно $\dfrac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_m!}$ способами.

Порядок элементов внутри каждой из этих m групп не имеет значения.

Пусть $A_1, A_2, \ldots, A_k$ — множества, число элементов в каждом из которых равно соответственно $n_1, n_2, n_3, \ldots, n_k$. Составить множество B из m_1 элементов множества A_1, m_2 элементов множества A_2, ..., m_k элементов множества A_k, можно, согласно основному комбинаторному принципу,

$$C_{n_1}^{m_1} \cdot C_{n_2}^{m_2} \cdot \ldots \cdot C_{n_k}^{m_k}$$

способами.

Для безошибочного выбора комбинаторной формулы достаточно последовательно ответить на вопросы в следующей схеме:

	Что нас интересует при выборе?	Какой выбор?	Формула
Откуда выбор? $(n-?)$	Состав	Бесповторный	C_n^r
		Повторный	C_{n+r-1}^r
Сколько выбираем? $(r-?)$	Состав и порядок	Бесповторный	A_n^r
		Повторный	n^r
	Порядок	Бесповторный	$n!$
		Повторный	$\dfrac{n!}{k_1!k_2!\ldots k_m!}$

ПРИМЕР 1.1. Сколько сообщений можно послать посредством семи знаков точек или тире?

◁ Выбор знака производится из множества двух элементов: точка или тире $(n = 2)$. Повторным способом выбирается семь элементов $(r = 7)$. Поэтому число различных сообщений равно $2^7 = 128$. ▷

ПРИМЕР 1.2. Сколько комбинаций из четырех букв можно составить? Сколько из них содержат только разные буквы?

◁ Из совокупности 33 букв $(n = 33)$ необходимо выбрать четыре буквы $(r = 4)$. Если запрета на повторение букв нет, то выбор повторный и общее число комбинаций равно $(33)^4 = 1\,185\,921$. Если необходимо иметь только разные буквы, то выбор бесповторный и общее число комбинаций равно

$$A_{33}^4 = 33 \cdot 32 \cdot 31 \cdot 30 = 982\,080. \quad \triangleright$$

ПРИМЕР 1.3. Сколькими способами можно разложить восемь книг на две пачки по четыре книги в каждой? Сколькими способами можно разложить эти книги на четыре пачки по две книги в каждой? Сколькими способами можно разослать эти книги восьми различным адресатам?

◁ 1. Для разделения книг на две равные пачки достаточно из восьми книг ($n = 8$) выбрать бесповторным способом любые четыре ($r = 4$) для первой пачки, причем нас интересует только состав выбора, а остальные книги оставить для второй пачки. Поэтому общее число способов равно числу сочетаний из восьми элементов по четыре, т. е. $C_8^4 = \dfrac{8!}{4! \cdot 4!} = 70$.
2. Для комплектации четырех пачек по две книги в каждой необходимо сначала бесповторным способом выбрать состав первой пачки ($C_8^2 = 28$ способов), затем из оставшихся шести книг выбрать две книги для второй пачки ($C_6^2 = 15$ способов), после этого из оставшихся четырех книг выбрать две книги для третьей пачки ($C_4^2 = 6$ способов), а оставшиеся две книги составят четвертую пачку (формально, $C_2^2 = 1$ способ). Заметим, что при каждом выборе мы интересовались только составом. По комбинаторному принципу для описанной последовательности выборов существует $28 \cdot 15 \cdot 6 = 2520$ способов.
3. Рассылка восьми адресатам по одной книге каждому означает перестановку из восьми элементов, т. е. имеется $8! = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40\,320$ способов. ▷

Замечание. Для числа способов разделить n элементов на m групп по $k_1, k_2, k_3, ..., k_k$ соответственно в каждой $(k_1 + k_2 + k_3 + ... + k_m = n)$ известна формула

$$\frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_m!}.$$

По этой формуле решение примера 1.2(2) можно получить сразу: $\dfrac{8!}{2! \cdot 2! \cdot 2! \cdot 2!} = 2520$.

ПРИМЕР 1.4. Сколькими способами можно составить трехцветный флаг, если имеется материал пяти цветов?
◁ Для составления трехцветного флага нужно из пяти цветов ($n = 5$) выбрать три различных цвета ($r = 3$), иначе флаг не будет трехцветным. При выборе нас интересует состав выбора и порядок следования цветов. Поэтому число трехцветных флагов в нашем случае равно числу размещений из пяти по три, т. е. равно $A_5^3 = 5 \cdot 4 \cdot 3 = 60$. ▷

ПРИМЕР 1.5. Каких чисел от 1 до 10 000 000 больше — тех, в записи которых встречается единица, или тех, в записи которых нет ни одной единицы?
◁ Для того чтобы записать семизначное число, в записи которого нет ни одной единицы, необходимо повторным способом из девяти цифр $(0, 2, 3, 4, 5, 6, 7, 8, 9)$ выбрать семь цифр. Это можно сделать $9^7 = 4\,769\,847 < 5 \cdot 10^6$ способами. Следовательно, среди первых 10 миллионов чисел больше тех, в записи которых единица есть. ▷

ПРИМЕР 1.6. Каждый из 10 студентов может явиться на зачет в любой из двух назначенных дней. Сколькими способами могут студенты распределиться по дням явки на зачет? Сколькими способами могут распределиться студенты по дням явки на зачет, если каждый день должны сдавать зачет по пять студентов?
◁ При распределении по дням явки каждый из 10 студентов производит выбор между двумя возможностями. По комбинаторному принципу всего способов выбора имеется $2^{10} = 1\,024$. Если же в каждый из дней должно явиться на зачет по пять студентов, то достаточно выбрать студентов для первого дня зачета, а остальные будут сдавать зачет во второй день. Из 10 студентов ($n = 10$) следует выбрать бесповторным способом пять студентов ($r = 5$), причем нас интересует только состав выбора. Поэтому возможных комбинаций будет $C_{10}^5 = 252$. ▷

ПРИМЕР 1.7. В некотором государстве не было двух жителей с одинаковым набором зубов. Какой может быть наибольшая численность населения в этом государстве?
◁ В отношении каждого из 32 зубов может быть две возможности: есть этот зуб или его нет. Выбор из этих двух возможностей нужно произвести 32 раза. Поэтому число всех мыслимых комбинаций равно $2^{32} = 4\,294\,967\,296$. ▷

ПРИМЕР 1.8. Сколькими различимыми способами можно переставить между собой буквы:
а) A_1, A_2, B_1, B_2, B_3; б) A, A, B_1, B_2, B_3; в) A, A, B, B, B?
◁ а) Так как все буквы различны, то число перестановок из них равно $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.
б) В этом случае каждая пара перестановок отличающаяся только порядком расположения букв, например, A_1, B_1, A_2, B_2, B_3 и A_2, B_1, A_1, B_2, B_3, сливается в одну

перестановку A, B_1, A, B_2, B_3. Поэтому различимых перестановок будет $120/2 = 60$.

в) Итак, имеется 60 перестановок, которые отличаются друг от друга либо местами расположения букв B_1, B_2, B_3, либо порядком расположения этих букв на данных местах. Всего перестановок букв B_1, B_2, B_3 на заданных трех местах существует $3! = 6$. Поэтому при потере индексов у букв B_1, B_2, B_3 различимых комбинаций будет $60/6 = 10$.

Заметим, что число различимых перестановок равно числу способов выбора из пяти мест любых двух и постановки на них буквы A. На остальные места ставим буквы B. Это число равно $C_5^2 = 10$. ▷

ПРИМЕРЫ (для самостоятельного решения).

1. Перечислите все перестановки из элементов *А, В, С, Д*. Перечислите все размещения по два элемента и все сочетания по два элемента из элементов *А, В, С, Д*.

2. Сколькими способами можно из колоды карт (36 штук) выбрать пять карт так, чтобы среди них было два туза?

3. Сколькими способами можно из колоды карт выбрать четыре карты разных мастей?

4. Сколько различных перестановок можно сделать из букв слова:

а) наука;

б) математика?

5. Сколько автомобильных номеров можно составить из трех букв и трех цифр? Сколько таких номеров можно составить из различных букв и различных цифр?

6. Сколькими способами можно переставить между собой цифры 1, 2, 3, 4, 5, 6 так, чтобы на четных по порядку местах стояли четные цифры, а на нечетных — нечетные?

7. Сколькими способами можно расставить шесть книг по трем полкам? Сколько способов расставить книги так, чтобы ни одна полка не пустовала?

8. Сколько четырехзначных чисел состоят только из разных цифр?

9. Сколькими способами можно поставить на полку шесть книг так, чтобы три заданные книги оказались рядом (в произвольном порядке)?

10. Сколькими способами можно разделить 15 команд на три подгруппы в каждой по пять команд?

11. Сколько различных частных производных третьего порядка имеет функция трех переменных?

12. Сколькими способами можно расселить девять студентов в трех комнатах, рассчитанных на трех человек каждая? Сколькими способами это можно сделать, если какие-либо два из этих студентов отказываются поселиться в одной комнате?

13. Студенту нужно выбрать два факультативных курса из шести возможных. Сколькими способами он может это сделать?

14. Два города A и B соединены четырьмя различными дорогами. Сколькими способами можно проехать из A в B и обратно? Сколько существует таких способов, если на обратном пути непременно выбирать новую дорогу?

15. В соревнованиях принимают участие 16 равносильных команд. Сколькими способами могут быть распределены золотая и серебряная медали?

16. Сколькими способами можно выбрать путь из начала координат в точку с координатами (6,4), если каждый шаг равен 1, и его можно совершать только вправо или вверх?

Ответы:

2. 29 760;
3. 6 561;
4. а) 60, б) 151 200;
5. 35 937 000, 23 569 920;
6. 36;
7. 729, 540;
8. 4 536;
9. 144;
10. 756 756;
11. 10;
12. 1 680, 1 260;
13. 15;
14. 16, 12;
15. 240;
16. 210.

Занятие 2

НЕПОСРЕДСТВЕННОЕ ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ

Пусть опыт имеет n возможных исходов. Исходы опыта, при которых появляется событие A, называют исходами, благоприятствующими этому событию.

Классическое определение вероятности. Если исходы опыта *равновозможны*, то вероятностью события A называется отношение числа исходов, *благоприятствующих* данному событию, к числу всех возможных исходов опыта, т. е. $P(A) = \dfrac{m}{n}$, где m — число исходов опыта, благоприятствующих событию, а n — число всех возможных исходов.

Свойства вероятностей:

1. Вероятность любого события — есть число, заключенное между нулем и единицей, т. е. $0 \leqslant P(A) \leqslant 1$. Вероятность невозможного события равна 0, а вероятность достоверного события равна 1.

2. Если события A и B *несовместны,* то $P(A+B) = P(A) + P(B)$.

3. Вероятность любого события A в сумме с вероятностью противоположного события равна единице: $P(\bar{A}) + P(A) = 1$.

Если вероятность интересующего нас события A по каким-либо причинам вычислить трудно, то можно попытаться вычислить вероятность противоположного события, а затем с помощью свойства 3 вычислить искомую вероятность события A.

ПРИМЕР 2.1. Брошены две игральные кости. Найти вероятности следующих событий:

A — на обеих костях выпало одинаковое число очков;

B — число очков на первой кости больше, чем на второй;

C — сумма очков четная;

D — сумма числа очков больше двух.

$\lhd$ Число очков, благоприятствующих каждому из названных событий, легко подсчитать, если все возможные исходы опыта перечислить в виде таблицы. В каждой клетке таблицы первая цифра указывает число очков на первой кости, вторая — на второй кости.

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Если кости симметричны и однородны, то все перечисленные исходы опыта равновозможны. Непосредственный подсчет числа благоприятствующих исходов дает $P(A) = 6/36 = 1/6$, $P(B) = 15/36 = 5/12$, $P(C) = 18/36 = 1/2$, $P(D) = 35/36$. $\rhd$

ПРИМЕР 2.2. Из партии, содержащей 10 изделий, среди которых три бракованных, наугад извлекаются три изделия для контроля. Найти вероятности следующих событий:

A — среди выбранных изделий ровно два бракованных;

B — выбраны только бракованные изделия;

C — среди выбранных изделий содержится хотя бы одно бракованное.

$\lhd$ Выбрать любых три изделия из десяти можно C_{10}^3 способами. Поэтому имеем $n = C_{10}^3 = 120$ равновозможных исходов.

Событию A благоприятствуют те исходы, при которых из семи годных изделий выбирается одно (это можно сделать $C_7^1 = 7$ способами) и из трех бракованных — два (это можно сделать $C_3^2 = 3$ способами). По комбинаторному принципу число благоприятствующих событию A исходов равно $C_7^1 \cdot C_3^2 = 7 \cdot 3 = 21$. Поэтому $P(A) = 21/120 = 7/40 \approx 1/6$, т. е. примерно один шанс из шести.

Событию B благоприятствует всего один исход и его вероятность $P(B) = 1/120$.

Вероятность события C проще вычислить, определив сначала вероятность события $\bar{C}$, которое состоит в том, что выбраны все годные изделия. Выбрать три годных изделия из семи можно $C_7^3 = 35$ способами. Поэтому $P(\bar{C}) = 35/120$ и $P(C) = 1 - P(\bar{C}) = 1 - 35/120 = 17/24 \approx 2/3$. $\triangleright$

ПРИМЕР 2.3. Каждый из пяти студентов может сдавать зачет в один из пяти назначенных дней. Выбор каждым студентом любого дня равновозможен. Какова вероятность того, что каждый день на зачет будет приходить только один из этих студентов? Если студентов трое, а дней пять, то какова вероятность того, что эти студенты явятся на зачет в разные дни?

$\triangleleft$ Каждый из пяти студентов может выбрать любой из пяти дней, поэтому по дням явки на зачет студенты могут распределиться 5^5 способами.

Благоприятствующие способы можно перебрать, если распределить студентов по одному на каждый день и рассмотреть всевозможные их перестановки. Таких перестановок существует $A_5^5 = 5! = 120$. Поэтому вероятность явки каждый день по одному студенту равна $P = 5!/5^5 = 24/625$.

Если студентов трое, то возможных способов явки 5^3, а благоприятствующих из них $A_5^3 = 5 \cdot 4 \cdot 3 = 60$ (первый может явиться в любой из пяти дней, второй — в любой из четырех дней, третий — в любой из оставшихся трех дней). Вероятность интересующего нас события равна $p = 60/125 \approx 1/2$. $\triangleright$

ПРИМЕР 2.4. За семь дней недели независимо друг от друга происходит семь событий (скажем, семь аварий). Какова вероятность того, что каждый день будет происходить по одному событию?

$\triangleleft$ Для наглядности представим себе семь ящиков и семь шариков. Тогда распределение событий по дням недели равносильно раскладке шариков по ящикам. Первый шар можно положить в любой из семи ящиков, второй — также в любой из семи и т. д., поэтому, согласно комбинаторному принципу, всех возможных способов раскладки имеется 7^7. Для получения числа способов, благоприятствующих интересующему нас событию (событие A), разложим по одному шарику в каждый ящик, а затем станем менять местами шарики. Тогда число благоприятствующих способов равно числу перестановок из семи элементов, т. е. равно $7!$. В итоге имеем $P(A) = 7!/7^7 \approx 1/165$. $\triangleright$

ПРИМЕР 2.5. При раздаче тщательно перемешанных карт (в колоде 36 карт) игрок получает шесть карт. Какова вероятность того, что игрок получит два туза, два короля и две дамы любой масти?

$\triangleleft$ Шесть карт данному игроку можно сдать C_{36}^6 способами, так как выбор бесповторный и нас интересует только состав выбора. Выбрать два туза, два короля и две дамы можно $C_4^2 \cdot C_4^2 \cdot C_4^2 = 6^3 = 216$ способами. Поэтому искомая вероятность равна $P = 216/C_{36}^6 = 9/2618 \approx 0,003$. $\triangleright$

ПРИМЕР 2.6. Вы являетесь одним из восьми человек, среди которых по жребию распределяется три выигрыша. В розыгрыше каждого выигрыша участвуют все восемь человек. Найдите вероятности следующих событий: $A = \{$Вам достанутся все выигрыши$\}$; $B = \{$Вы не получите ни одного выигрыша$\}$; $C = \{$Вам достанется хотя бы один выигрыш$\}$.

$\triangleleft$ Три выигрыша среди 8 человек могут быть распределены $8^3 = 512$ способами. Событию A благоприятствует только один из этих способов распределения выигрышей. Поэтому $P(A) = 1/512$. Если каждый раз выбор будет производиться среди остальных семи человек (это можно сделать $7^3 = 343$ способами), то Вам не достанется ни одного выигрыша. Поэтому вероятность события B равна $P(B) = 343/512$. Событие C противоположно событию B. Поэтому

$$P(C) = 1 - P(B) = 1 - 343/512 = 169/512. \triangleright$$

Область применения классического определения вероятности — испытания с конечным числом равновозможных исходов. Существенным является условие равновозможности. От конечности числа исходов опыта можно отказаться и определять вероятности не с помощью числа исходов, а с помощью отношения длин, площадей и т. д.

Геометрическое определение вероятности. Пусть область g принадлежит области G. Если равновозможно попадание точки в любую точку области G, то вероятность попасть в область g равна отношению меры области g к мере области G:

$$P(\text{попасть в область } g) = \frac{\text{Мера области } g}{\text{Мера области } G},$$

где «мера» — означает:

1) длину, если область G часть прямой или кривой линии;

2) площадь, если G часть плоскости;

3) объем, если G часть пространства,

и так далее в зависимости от характера области G.

ПРИМЕР 2.7. Отрезок $[0,1]$ наугад делят на три части. Какова вероятность того, что из этих трех частей можно сложить треугольник?

$\lhd$ Обозначим первую из полученных частей отрезка через x, а вторую — через y. Тогда оставшаяся третья часть равна $1-x-y$. В треугольнике сумма двух любых сторон больше третьей стороны. Поэтому из частей отрезка получится треугольник, если выполнятся неравенства:

$$\begin{cases} x+y > 1-x-y; \\ x+(1-x-y) > y; \\ y+(1-x-y) > x. \end{cases} \Leftrightarrow \begin{cases} y > 1/2 - x; \\ y < 1/2; \\ x < 1/2. \end{cases}$$

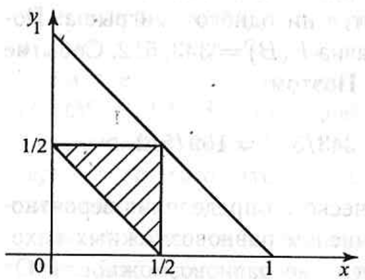


Рис. 2.1

Возможные значения для пары (x,y) составляют треугольник с вершинами $(0;0)$, $(0;1)$, $(1;0)$. Системе неравенств соответствует заштрихованная треугольная область на рис. 2.1, площадь которой равна $1/4$ площади исходного треугольника. Поэтому искомая вероятность, как отношение площадей, равна $1/4$. $\rhd$

ПРИМЕР 2.8. Две радиостанции в течение часа независимо друг от друга должны передать сообщения длительностью 10 мин и 20 мин соответственно. Какова вероятность того, что сообщения не перекроются по времени.

$\lhd$ Пусть x — момент начала сообщения первой радиостанции, а y — момент начала второго сообщения. Для того, чтобы сообщения уложились в отведенный час, должны выполняться условия: $0 \leqslant x \leqslant 50$ мин; $0 \leqslant y \leqslant 40$ мин. Сообщения не перекроются во времени, если выполнятся условия $y-x > 10$ и $x-y > 20$. Этим условиям удовлетворяют точки заштрихованных областей на рис. 2.2.

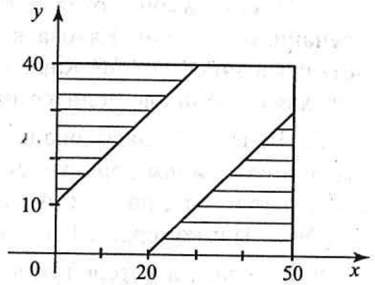


Рис. 2.2

Так как все положения точки (x,y) в прямоугольнике 50×40 равновозможны, то искомая вероятность равна отношению заштрихованной площади, которая равна 30×30, к площади прямоугольника. Поэтому $P = (30\times30)/(50\times40) = 9/20$. $\rhd$

ПРИМЕРЫ (для самостоятельного решения).

1. Десять команд случайным образом (по жребию) разбиваются на две равные подгруппы. Какова вероятность того, что две сильнейшие команды попадут в разные подгруппы? ... в одну подгруппу? ... в первую подгруппу?

2. Из полной колоды карт (52 штуки) наугад выбраны три карты. Какова вероятность того, что это «тройка», «семерка», «туз»? Какова вероятность того, что эти карты выбраны в указанной последовательности?

3. На отрезок OA длины L брошены «наугад» две точки B и C, причем точка C расположена правее точки B. Найти вероятность того, что длина отрезка BC меньше длины отрезка OB.

4. Десять билетов с номерами от 1 до 10 перемешаны на столе экзаменатора. Какова вероятность того, что эти билеты будут вытянуты студентами в порядке их номеров?

5. Четыре человека вошли в лифт на первом этаже девятиэтажного дома. Считая, что равновозможен выход каждого пассажира на любом из этажей со второго по девятый, найти вероятность того, что: а) все пассажиры выйдут на разных этажах; б) все пассажиры выйдут выше пятого этажа; в) на третьем этаже не выйдет ни одного пассажира.

6 Петрушко И. М.

6. Наугад выбираются четыре цифры и расставляются в случайном порядке. Какова вероятность того, что получится четырехзначное число? Какова вероятность того, что это четырехзначное число делится на 5?

7. Имеется 20 экзаменационных билетов, разложенных на столе в случайном порядке. Десять студентов один за другим выбирают наугад по одному билету. Какова вероятность того, что билеты с номерами 1 и 2 не будут выбраны?

8. Из урны, в которой лежат, шесть белых, четыре черных и два красных шара, наугад выбирают четыре шара. Какова вероятность того, что среди них только черные и красные шары?

9. Десять книг, из них три красные, в случайном порядке поставлены на полку. Какова вероятность того, что три красные книги в любом порядке стоят рядом?

10. Бросаются три игральные кости. Найдите вероятности следующих событий: $A = \{$кости выпадут разными гранями$\}$; $B = \{$на всех костях выпадет одинаковое число очков$\}$.

11. В шкафу лежат вперемежку пять пар ботинок. Наугад выбирается два ботинка. Какова вероятность того, что они образуют пару?

12. Каждый из шести призов в результате жеребьевки разыгрывается между десятью участниками. Какова вероятность того, что данные шесть участников получат по одному призу каждый?

13. Группа из четырех юношей и четырех девушек по жребию делится на две подгруппы по четыре человека. Какова вероятность того, что в каждую подгруппу попадет поровну юношей и девушек?

14. Внутрь круга радиусом R наугад брошена точка. Какова вероятность того, что она попадет внутрь вписанного в круг: а) квадрата, б) правильного треугольника?

15. В течение суток к причалу независимо друг от друга должны подойти и разгрузиться два сухогруза. Одному из них требуется для разгрузки шесть часов, другому — восемь часов. Какова вероятность того, что ни одному из сухогрузов не придется ждать очерёди для разгрузки?

Ответы:

1. $5/9$, $4/9$, $2/9$;
2. $16/5525 \approx 0,003$, $8/16\,575 \approx 0,0005$;
3. $1/2$;
4. $1/10!$;
5. а) $105/256$, б) $1/16$, в) $2401/4096 \approx 0,59$;
6. $0,9$, $0,18$;
7. $9/38$;
8. $1/33$;
9. $1/15$;
10. $P(A)=5/9$, $P(B)=1/36$;
11. $1/9$;
12. $0,00072$;
13. $18/35$;
14. а) $2/\pi \approx 2/3$, б) $(3\sqrt{3})/(4\pi) \approx 0,41$;
15. $25/72 \approx 1/3$.

Занятие 3

ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

В теории вероятностей события рассматривают на фоне комплекса условий, которые его порождают. Проще говоря, событие — это результат опыта, который проистекает в природе по воле человека, независимо от нее или ей вопреки. Рассмотрим множество событий, которые можно наблюдать в эксперименте при фиксированном комплексе условий. На множестве таких событий определим следующие понятия.

Суммой событий A и B называется событие, состоящее в появлении хотя бы одного из событий A или B. Сумму событий A и B обозначают через $A+B$.

Произведением событий A и B называют событие, состоящее в появлении событий A и B в одном и том же опыте. Обозначают произведение событий A и B через $A \cdot B$.

Событие, состоящее в не появлении события A, называется противоположным событием и обозначается через $\bar{A}$.

Вероятность события A, вычисленная при условии, что событие B произошло, называется условной вероятностью события A и обозначается $P(A/B)$.

Теорема 3.1 (умножения вероятностей). *Вероятность произведения событий равна вероятности одного события, умноженной на вероятность другого события, вычисленную при условии, что первое событие произошло, т. е.*

$$P(A \cdot B) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B). \tag{3.1}$$

События называются *независимыми*, если появление одного из них не изменяет вероятности появления другого.

Если события независимы, то $P(A/B) = P(A)$, $P(B/A) = P(B)$ и

$$P(A \cdot B) = P(A) \cdot P(B).$$

Для любого конечного числа событий вероятность произведения событий равна произведению вероятностей этих событий, причем вероятность каждого следующего события вычисляется при условии, что предыдущие события произошли, т. е.

$$P(A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_{n-1} \cdot A_n) =$$
$$= P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cdot A_2) \cdot \ldots \cdot P(A_n/A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_{n-1}).$$

Если события независимы, то

$$P(A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_{n-1} \cdot A_n) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot \ldots \cdot P(A_n).$$

Итак, прежде чем вычислять вероятность произведения событий, необходимо установить, зависимы ли события или нет.

Теорема 3.2 (сложения вероятностей). *Вероятность суммы событий A и B равна сумме вероятностей этих событий без вероятности их совместного появления:*

$$P(A+B) = P(A) + P(B) - P(A \cdot B).$$

События называются *несовместными*, если их появление в одном и том же опыте невозможно.

Если события A и B несовместны, то

$$P(A+B) = P(A) + P(B).$$

Для трех совместных событий теорема сложения вероятностей имеет вид:

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

Если события несовместны, то

$$P(A+B+C) = P(A) + P(B) + P(C).$$

Теорему сложения можно обобщить на любое конечное число слагаемых:

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j\leqslant n} P(A_i\cdot A_j) + \sum_{1\leqslant i<j<k\leqslant n} P(A_i\cdot A_j\cdot A_k) + \ldots + (-1)^{n+1} P(A_1\cdot A_2\cdot\ldots\cdot A_n).$$

Если события несовместны, то

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Итак, *прежде чем вычислять вероятность суммы событий следует выяснить, совместны они или нет.*

ПРИМЕР 3.1. Вероятности попадания в цель при одном выстреле для первого, второго и третьего стрелков равны соответственно 0,3; 0,6; 0,8. Все три стрелка выстрелили в цель. Какова вероятность того, что: а) цель поражена; б) произошло только одно попадание; в) произошло ровно два попадания; г) попадут все три стрелка; д) будет хотя бы один промах?

$\lhd$ Обозначим через A_i — событие, состоящее в попадании в цель i-го стрелка.

а) Поражение цели (событие A) равносильно появлению хотя бы одного из событий A_1 или A_2, или A_3. Поэтому $A = A_1 + A_2 + A_3$. Учитывая совместность событий, имеем $P(A) = P(A_1) + P(A_2) + P(A_3) - P(A_1\cdot A_2) - P(A_1\cdot A_3) - P(A_2\cdot A_3) + P(A_1\cdot A_2\cdot A_3)$, а так как события независимы, то $P(A) = 0,3 + 0,6 + 0,8 - 0,3\cdot 0,6 - 0,3\cdot 0,8 - 0,6\cdot 0,8 + 0,3\cdot 0,6\cdot 0,8 = 0,944$.

б) Рассмотрим три случая:

1) $B_1 = A_1\cdot\bar{A}_2\cdot\bar{A}_3$ — первый стрелок попал в цель и при этом второй не попал и третий не попал.

2) $B_2 = \bar{A}_1\cdot A_2\cdot\bar{A}_3$ — первый стрелок не попал и при этом второй попал и третий не попал.

3) $B_3 = \bar{A}_1\cdot\bar{A}_2\cdot A_3$ — первый и второй не попали и при этом третий попал.

Только одно попадание в цель (событие B) равносильно реализации хотя бы одного из несовместных событий B_1 или B_2, или B_3. Поэтому

$$B = A_1\cdot\bar{A}_2\cdot\bar{A}_3 + \bar{A}_1\cdot A_2\cdot\bar{A}_3 + \bar{A}_1\cdot\bar{A}_2\cdot A_3.$$

В силу независимости событий A_i имеем

$$P(B) = 0,3\cdot 0,4\cdot 0,2 + 0,7\cdot 0,6\cdot 0,2 + 0,7\cdot 0,4\cdot 0,8 = 0,332.$$

в) Два попадания в цель (событие C) равносильны реализации хотя бы одного из несовместных случаев: $A_1\cdot A_2\cdot\bar{A}_3$ или $A_1\cdot\bar{A}_2\cdot A_3$, или $\bar{A}_1\cdot A_2\cdot A_3$. В силу независимости событий A_i получаем

$$P(C) = 0,3\cdot 0,6\cdot 0,2 + 0,3\cdot 0,4\cdot 0,8 + 0,7\cdot 0,6\cdot 0,8 = 0,468.$$

г) Все три стрелка попадут в цель (событие D), если произойдут события A_1 и A_2, и A_3, т. е. $D = A_1\cdot A_2\cdot A_3$. В силу независимости событий A_i имеем

$$P(D) = P(A_1)\cdot P(A_2)\cdot P(A_3) = 0,3\cdot 0,6\cdot 0,8 = 0,144.$$

д) Хотя бы один промах (событие E) равносилен появлению хотя бы одного из событий $\bar{A}_1$ или $\bar{A}_2$, или $\bar{A}_3$, т. е. $E = \bar{A}_1 + \bar{A}_2 + \bar{A}_3$. Вместо вычисления вероятности суммы трех совместных событий, заметим, что событие E равносильно непоявлению события D. Поэтому

$$P(E) = P(\bar{D}) = 1 - P(D) = 1 - 0,144 = 0,856. \rhd$$

ПРИМЕР 3.2. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлены две системы (рис. 3.1).

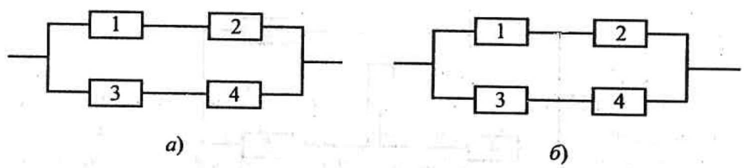


Рис. 3.1

Какая система надежнее? Иначе говоря, что выгоднее в системе дублировать, каждый элемент отдельно или всю систему в целом?

◁ Пусть событие A_i состоит в том, что i-й элемент работает безотказно. Безотказная работа первой системы (событие B_1) равносильна безотказной работе первого элемента и второго или третьего элемента и четвертого. Символически это можно записать в виде $B_1 = A_1 \cdot A_2 + A_3 \cdot A_4$. События $A_1 \cdot A_2$ и $A_3 \cdot A_4$ совместны, а события A_i независимы. Поэтому

$$P(B_1) = P(A_1 \cdot A_2) + P(A_3 \cdot A_4) - P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) =$$

$$= P(A_1) \cdot P(A_2) + P(A_3) \cdot P(A_4) - P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) =$$

$$= (0,8)^2 + (0,8)^2 - (0,8)^4 = 0,8704.$$

Безотказная работа второй системы (событие B_2) равносильна безотказной работе первого элемента или третьего и второго или четвертого, т. е. $B_2 = (A_1 + A_3) \cdot (A_2 + A_4)$. Тогда

$$P(B_2) = P(A_1 + A_3) \cdot P(A_2 + A_4) =$$
$$= [P(A_1) + P(A_3) - P(A_1) \cdot P(A_3)] \times$$
$$\times [P(A_2) + P(A_4) - P(A_2) \cdot P(A_4)] =$$
$$= (0,8 + 0,8 - 0,64)^2 = 0,9216.$$

Результаты вычислений свидетельствуют о том, что выгоднее дублировать каждый элемент отдельно. Система б) надежнее. ▷

ПРИМЕР 3.3. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлена система (рис. 3.2).

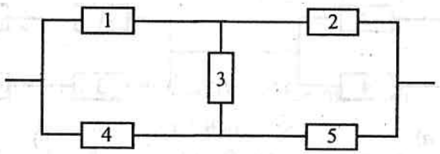


Рис. 3.2

Какова надежность системы?

◁ Сохраним обозначения примера 3.3. Если элемент № 3 не работает, то система совпадает с системой а) из этого примера.

Если же элемент № 3 работает, то система совпадает с системой б) из предыдущего примера. Поэтому безотказная работа системы (событие D) равносильна событию $D = \bar{A}_3 \cdot B_1 + A_3 \cdot B_2$. События независимы, а события $\bar{A}_3 \cdot B_1$ и $A_3 \cdot B_2$ несовместны. Поэтому с учетом результатов предыдущей задачи имеем

$$P(D) = P(\bar{A}_3) \cdot P(B_1) + P(A_3) \cdot P(B_2) =$$
$$= 0,2 \cdot 0,8704 + 0,8 \cdot 0,9216 = 0,91136. \ ▷$$

ПРИМЕР 3.4. В одной урне пять белых, семь черных и три красных шара, а во второй соответственно четыре белых, два черных и четыре красных шара. Из каждой урны вынимают наугад по одному шару. Какова вероятность того, что будут выбраны шары одного цвета?

◁ Обозначим извлечение из i-й урны белого, черного и красного шара соответственно через B_i, C_i, K_i. Тогда извлечение шаров одного цвета (событие A) можно записать следующим образом:

$$A = B_1 \cdot B_2 + C_1 \cdot C_2 + K_1 \cdot K_2.$$

Так как события, образующие сумму, несовместны, то

$$P(A) = P(B_1 \cdot B_2) + P(C_1 \cdot C_2) + P(K_1 \cdot K_2).$$

В силу независимости событий

$$P(A) = P(B_1) \cdot P(B_2) + P(C_1) \cdot P(C_2) + P(K_1) \cdot P(K_2) =$$

$$= \frac{5}{15} \cdot \frac{4}{10} + \frac{7}{15} \cdot \frac{2}{10} + \frac{3}{15} \cdot \frac{4}{10} = \frac{23}{75} \cong \frac{1}{3}. \ ▷$$

ПРИМЕР 3.5. В партии из 25 деталей четыре бракованных. Детали выбирают для проверки наугад по одной пока не попадется бракованная. Какова вероятность того, что будет проверено ровно три детали?

◁ Обозначим через A интересующее нас событие, а через A_i — событие, состоящее в выборе годной детали при i-м выборе. Событие A произойдет, если первая и вторая детали окажутся годными и лишь третья по счету окажется бракованной. Это означает, что $A = A_1 \cdot A_2 \cdot \bar{A}_3$, причем события зависимы. Поэтому

$$P(A) = P(A_1) \cdot P(A_2/A_1) \cdot P(\bar{A}_3/A_1, A_2) = \\ = \frac{21}{25} \cdot \frac{20}{24} \cdot \frac{4}{23} = \frac{42}{345} \cong 0,12. \rhd$$

ПРИМЕР 3.6. Урна содержит шесть занумерованных шаров с номерами от 1 до 6. Шары извлекаются по одному без возвращения. Пусть событие A состоит в том, что шары будут извлечены в порядке их номеров, а событие B — в том, что хотя бы один раз номер шара совпадет с порядковым номером его извлечения. Найти вероятности событий A и B и определить предельные вероятности этих событий при неограниченном увеличении числа шаров в урне.

$\lhd$ Обозначим через A_i — событие, состоящее в том, что порядок извлечения i-го шара совпадает с его номером. Тогда событие $A = A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_6$. Вместо рассмотрения произведения зависимых событий заметим, что шары в указанном порядке можно извлечь только одним способом, а всего равновозможных способов извлечения существует 6!. Поэтому $P(A) = \frac{1}{6!} = \frac{1}{720}$. При увеличении числа шаров $P(A) \to 0$. Событие B произойдет, если появится хотя бы одно из событий A_1 или A_2, или ... или A_6. Поэтому $B = A_1 + A_2 + A_3 + A_4 + A_5 + A_6$, причем события совместны. При переходе к противоположному событию придется рассматривать произведение шести зависимых событий $\bar{A}_i$, что в данном случае сделать сложно. Поэтому вычислим вероятность суммы непосредственно:

$$P(B) = \sum_{i=1}^{6} P(A_i) - \sum_{1 \leqslant i<j \leqslant 6} P(A_i) \cdot P(A_j) + \\ + \sum_{1 \leqslant i<j<k \leqslant 6} P(A_i) \cdot P(A_j) \cdot P(A_k) - \ldots - P(A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6) = \\ = 6 \cdot \frac{1}{6} - C_6^2 \cdot \frac{1}{6} \cdot \frac{1}{5} + C_6^3 \cdot \frac{1}{6} \cdot \frac{1}{5} \cdot \frac{1}{4} - C_6^4 \cdot \frac{1}{6} \cdot \frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} + C_6^5 \cdot \frac{1}{6} \cdot \frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} - \\ - \frac{1}{6} \cdot \frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot 1 = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} + \frac{1}{120} - \frac{1}{720} = \frac{454}{720} \approx 0,63.$$

Заметим, что искомая вероятность является частичной суммой ряда Тейлора функции $1 - e^{-x}$ при $x = -1$. Поэтому при больших n имеем

$$P(A_1 + A_2 + A_3 + \ldots + A_n) \approx 1 - e^{-1} \approx 0,63. \rhd$$

ПРИМЕР 3.7. В колоде 36 карт. Каждому из четырех игроков раздается по шесть карт. Какова вероятность того, что каждый игрок получит по одному тузу?

$\lhd$ Обозначим интересующее нас событие через A. Пусть A_i означает, что i-й игрок получил при раздаче одного туза. Тогда $A = A_1 \cdot A_2 \cdot A_3 \cdot A_4$. События A_i зависимы. Поэтому $P(A) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1A_2) \cdot P(A_4/A_1A_2A_3)$. Карты для первого игрока могут быть выбраны C_{36}^6 возможными способами, событию A_1 благоприятствует $C_4^1 \cdot C_{32}^5 = 4 \cdot C_{32}^5$ способов. Для второго игрока карты, из числа оставшихся, могут быть выбраны C_{30}^6 способами, событию A_2, с учетом появления события A_1, благоприятствует $C_3^1 \cdot C_{27}^5 = 3 \cdot C_{27}^5$ способов. Для третьего игрока, с учетом появления событий A_1 и A_2, карты могут быть выбраны C_{24}^6 способами, из них событию A_3 благоприятствует $C_2^1 \cdot C_{22}^5$ способов. Наконец событию A_4, с учетом появления событий A_1, A_2 и A_3, благоприятствует $C_1^1 \cdot C_{17}^5$ способов из числа C_{18}^6 возможных. Поэтому искомую вероятность запишем в виде

$$P(A) = \frac{C_4^1 \cdot C_{32}^5}{C_{36}^6} \cdot \frac{C_3^1 \cdot C_{27}^5}{C_{30}^6} \cdot \frac{C_2^1 \cdot C_{22}^5}{C_{24}^6} \cdot \frac{C_1^1 \cdot C_{17}^5}{C_{18}^6} = \frac{144}{6545} \approx 0,02. \rhd$$

ПРИМЕРЫ (для самостоятельного решения).

1. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлены три системы (см. рис. 3.3). Какова надежность каждой из систем?

2. Из урны, содержащей семь белых и три черных шара, наугад последовательно извлекают по одному шару до появления черного шара. Найти вероятность того, что придется извлечь четыре шара в предположении, что выбор производится:

а) с возвращением;

б) без возвращения.

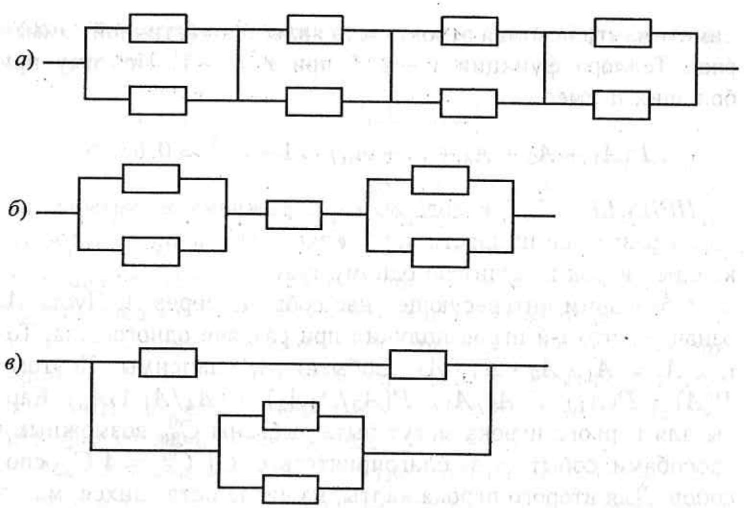


Рис. 3.3

3. Из полного набора костей домино (28 штук) наугад выбирают семь костей. Какова вероятность того, что среди них окажется по меньшей мере одна кость с шестью очками?

4. Подбрасываются четыре игральные кости. Какова вероятность того, что на них выпадут разные грани?

5. В урне пять белых, семь красных и восемь синих шаров. Наугад выбрано два шара. Какова вероятность того, что шары одного цвета?

6. Два дуэлянта одновременно стреляют друг в друга. Для каждого вероятность убить противника равна 0,2. Какова вероятность того, что дуэль закончится гибелью одного из дуэлянтов?

7. Двадцать футбольных команд, среди которых четыре призера предыдущего первенства, по жеребьевке разбиваются на четыре занумерованных подгруппы по пять команд в каждой. Какова вероятность того, что в каждую подгруппу попадет по одному призеру? Какова вероятность того, что в первую подгруппу не попадет ни одного призера?

8. В первой урне два белых и три черных шара, во второй один белый и два синих шара, в третьей три белых и один

красный шар. Из каждой урны наугад вынули по одному шару. Найдите вероятности следующих событий: $A = \{$вынут только один белый шар$\}$; $B = \{$вынут хотя бы один белый шар$\}$; $C = \{$вынуты шары разных цветов$\}$.

9. Из колоды в 36 карт выбрали наугад две. Какова вероятность того, что обе карты красной масти?

10. В ящике находятся три неисправные лампочки и семь исправных. Лампочки извлекают наугад по одной и проверяют, пока не будут выбраны две исправные. Какова вероятность того, что придется проверить половину лампочек из ящика?

Ответы:

1. а) $\approx 0,85$, б) $\approx 0,74$, в) $\approx 0,98$;
2. а) $\frac{1029}{10\,000} \approx 0,1$, б) $\frac{1}{8}$;
3. $\frac{2966}{3289} \approx 0,9$;
4. $\frac{5}{18}$;
5. $\frac{59}{190} \approx \frac{1}{3}$;
6. 0,32;
7. $\frac{125}{969}$, $\frac{91}{323}$;
8. $P(A) = \frac{5}{12}$, $P(B) = 0,9$, $P(C) = \frac{31}{60}$;
9. $\frac{17}{70}$;
10. $\frac{1}{30}$.

Лекция 6

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

§ 6.1. ОПРЕДЕЛЕНИЕ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Математическое ожидание случайной величины дает представление о ее среднем значении после проведения большого числа экспериментов. Рассмотрим дискретную случайную величину ξ с рядом распределения

ξ	a_1	a_2	...	a_k
P	p_1	p_2	...	p_k

Пусть осуществлено n реализаций случайного эксперимента, в котором наблюдается случайная величина ξ, и при этом значение a_1 появилось n_1 раз, значение a_2 появилось n_2 раз и т. д. Тогда среднее значение ξ после n экспериментов принимает вид

$$\frac{a_1n_1 + a_2n_2 + \ldots + a_kn_k}{n} = a_1\frac{n_1}{n} + a_2\frac{n_2}{n} + \ldots + a_k\frac{n_k}{n}. \quad (6.1)$$

В силу устойчивости частот:

$$\lim_{n\to\infty}\frac{n_1}{n} = P(\xi = a_1) = p_1, \qquad \lim_{n\to\infty}\frac{n_2}{n} = P(\xi = a_2) = p_2, \ldots .$$

Поэтому правая часть (6.1) имеет при $n \to \infty$ предел, равный $a_1p_1 + a_2p_2 + \ldots + a_kp_k$. На первый взгляд кажется, что мы вычислили среднее значение ξ, но в действительности, свойство устойчивости частот не выведено нами из аксиом теории

вероятностей. Тем не менее, приведенные соображения можно использовать для определения понятия, характеризующего среднее значение случайной величины.

Определение 6.1. ***Математическим ожиданием*** дискретной случайной величины ξ с рядом распределения

ξ	a_1	a_2	...	a_n	...
P	p_1	p_2	...	p_n	...

называется число

$$M\xi = \sum_{n=1}^{\infty} a_np_n.$$

Замечание 6.1. Если возможные значения ξ пронумеровать по-другому, то в последнем ряду изменится порядок слагаемых, а это может, как известно, изменить сумму ряда. Чтобы этого не произошло, будем предполагать, что ряд $\sum_{n=1}^{\infty} a_np_n$ сходится абсолютно.

ПРИМЕР 6.1. Пусть ξ имеет распределение Пуассона с параметром a, т. е. ξ принимает значения $0, 1, 2, \ldots$, причем

$$P(\xi = n) = \frac{a^n}{n!}e^{-a}, \quad n = 0, 1, 2, \ldots .$$

Найти $M\xi$.

По определению математического ожидания

$$M\xi = \sum_{n=0}^{\infty} nP(\xi = n) = \sum_{n=1}^{\infty} n\frac{a^n}{n!}e^{-a} = \sum_{n=1}^{\infty}\frac{a^n}{(n-1)!}e^{-a} =$$

$$= ae^{-a}\sum_{n=1}^{\infty}\frac{a^{n-1}}{(n-1)!} = ae^{-a}\left(1 + \frac{a}{1!} + \frac{a^2}{2!} + \ldots\right) = ae^{-a}e^{a} = a.$$

Замечание 6.2. Рассмотрим простейший поток с интенсивностью λ. Обозначим μ_t число событий в простейшем потоке за время t. Известно, что μ_t имеет распределение Пуассона с параметром $a = \lambda t$. Поэтому $M\mu_t = \lambda t$. При $t = 1$ получаем, что интенсивность простейшего потока λ совпадает со средним числом событий в простейшем потоке за единицу времени.

Рассмотрим теперь непрерывную случайную величину ξ с плотностью вероятностей $p(x)$. Как известно из лекции № 5, к ней близка дискретная случайная величина $\hat{\xi}$, принимающая значения $x_1, x_2, \dots$ с вероятностями $p(x_1)\triangle x$, $p(x_2)\triangle x$, … . Математическое ожидание $\hat{\xi}$ равно $\sum_{n=1}^{\infty} x_n p(x_n)\triangle x$. Но последняя сумма близка к $\int_{-\infty}^{+\infty} xp(x)dx$. Это приводит нас к следующему определению.

Определение 6.2. Математическим ожиданием непрерывной случайной величины ξ с плотностью вероятностей $p(x)$ называется число

$$M\xi = \int\limits_{-\infty}^{+\infty} xp(x)dx$$

(предполагается, что интеграл сходится абсолютно).

ПРИМЕР 6.2. Пусть ξ имеет показательное распределение, $\xi \sim \exp(\lambda)$. Это означает, что

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0; \\ 0, & x < 0. \end{cases}$$

Найти $M\xi$.

По определению

$$M\xi = \int\limits_{-\infty}^{+\infty} xp(x)dx = \int\limits_{0}^{+\infty} x\lambda e^{-\lambda x}dx = -\int\limits_{0}^{+\infty} xde^{-\lambda x} =$$

$$= -xe^{-\lambda x}\Big|_0^{+\infty} + \int\limits_{0}^{+\infty} e^{-\lambda x}dx = \int\limits_{0}^{+\infty} e^{-\lambda x}dx = \frac{1}{\lambda}.$$

§ 6.2. СВОЙСТВА МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Свойство 1. Если случайная величина ξ является постоянной C, то $M\xi = C$.

Доказательство. Случайная величина ξ является дискретной, принимающей одно значение C с вероятностью 1, поэтому $M\xi = C \cdot 1 = C$.

Свойство 2. Постоянный множитель можно выносить за знак математического ожидания, т. е.

$$M(C\xi) = CM\xi.$$

Доказательство (в дискретном случае). Если случайная величина ξ принимает значения $a_1, a_2, \dots$ с вероятностями $p_1, p_2, \dots$, то случайная величина $C\xi$ принимает значения $Ca_1, Ca_2, \dots$ с теми же вероятностями $p_1, p_2, \dots$, поэтому

$$M(C\xi) = \sum_{n=1}^{\infty} Ca_n p_n = C\sum_{n=1}^{\infty} a_n p_n = CM\xi.$$

Свойство 3. Математическое ожидание суммы равно сумме математических ожиданий, т. е.

$$M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2.$$

Доказательство (в дискретном случае). Пусть ξ_1 принимает значения x_n, а ξ_2 — значения y_m. Положим

$$p_{nm} = P(\xi_1 = x_n,\ \xi_2 = y_m).$$

Будем для простоты изложения считать, что всевозможные суммы $x_n + y_m$ различны. Тогда случайная величина $\xi_1 + \xi_2$ принимает значение $x_n + y_m$ с вероятностью p_{nm} и, следовательно,

$$M(\xi_1 + \xi_2) = \sum_{n,m} (x_n + y_m)p_{nm}. \tag{6.2}$$

Заметим, что

$$\{\xi_1 = x_n\} = \bigcup_m \{\xi_1 = x_n, \xi_2 = y_m\},$$

причем случайные события справа являются попарно несовместными, поэтому

$$P(\xi_1 = x_n) = \sum_m P(\xi_1 = x_n, \xi_2 = y_m) = \sum_m p_{nm}. \tag{6.3}$$

Аналогично показывается, что

$$P(\xi_2 = y_m) = \sum_n p_{nm}. \tag{6.4}$$

Преобразуем правую часть (6.2):

$$\sum_{n,m}(x_n + y_m)p_{nm} = \sum_{n,m} x_n p_{nm} + \sum_{n,m} y_m p_{nm} =$$

$$= \sum_n x_n \sum_m p_{nm} + \sum_m y_m \sum_n p_{nm}.$$

Откуда, учитывая соотношения (6.3), (6.4), получаем, что

$$\sum_{n,m}(x_n + y_m)p_{nm} = \sum_n x_n P(\xi_1 = x_n) + \sum_m y_m P(\xi_2 = y_m) =$$

$$= M\xi_1 + M\xi_2.$$

Ввиду (6.2), требуемое утверждение доказано.

Свойство 4. Если ξ — неотрицательная случайная величина, то $M\xi \geqslant 0$, причем $M\xi = 0$ тогда и только тогда, когда $\xi = 0$ с вероятностью 1.

Доказательство (в дискретном случае). Пусть ξ принимает значения $a_1, a_2, \ldots$ с вероятностями $p_1, p_2, \ldots$. Так как ξ неотрицательна, то $a_1, a_2, \ldots$ неотрицательны. Следовательно, $M\xi = \sum_{n=1}^{\infty} a_n p_n \geqslant 0$. Пусть теперь $\sum_{n=1}^{\infty} a_n p_n = 0$, но все слагаемые неотрицательны, значит, $a_n p_n = 0$ при любом n. А это означает, что если $a_n > 0$ при некотором n, то $p_n = 0$, что и требовалось доказать.

Свойство 5. Пусть $f(x)$ — числовая функция, а ξ — случайная величина, тогда $\eta = f(\xi)$ является случайной величиной и

$$M\eta = \sum_{n=1}^{\infty} f(a_n)p_n,$$

если ξ — дискретная случайная величина с рядом распределения

ξ	a_1	a_2	…	a_n	…
P	p_1	p_2	…	p_n	…

.

Если же ξ — непрерывная случайная величина с плотностью вероятностей $p(x)$, то

$$M\eta = \int_{-\infty}^{+\infty} f(x)p(x)dx.$$

Доказательство (в дискретном случае). Случайная величина $\eta = f(\xi)$ принимает значения $f(a_1)$, $f(a_2)$, … с вероятностями p_1, p_2, …, поэтому

$$M\eta = \sum_{n=1}^{\infty} f(a_n)p_n.$$

ПРИМЕР 6.3. Пусть ξ имеет равномерное распределение, $\xi \sim U[0, \pi]$, т. е.

$$p(x) = \begin{cases} 1/\pi, & x \in [0, \pi]; \\ 0, & x \notin [0, \pi]. \end{cases}$$

Найти математическое ожидание случайной величины $\sin \xi$.

В данном случае $f(x) = \sin x$, поэтому

$$M \sin \xi = \int_{-\infty}^{+\infty} \sin x\, p(x)dx = \frac{1}{\pi}\int_0^{\pi} \sin x dx = \frac{2}{\pi}.$$

Замечание 6.3. В курсе математического анализа средним значением интегрируемой функции $f(x)$ на отрезке $[a, b]$ называется число

$$\frac{1}{b-a}\int_a^b f(x)dx.$$

Нетрудно показать, что оно совпадает с $Mf(\xi)$, где $\xi \sim U[a, b]$.

§ 6.3. ДИСПЕРСИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ И ЕЕ СВОЙСТВА

Пусть ξ — случайная величина. Наряду с математическим ожиданием важными числовыми характеристиками ξ являются ее *начальные моменты*

$$m_k = M(\xi^k), \quad k = 1, 2, \ldots$$

и ее *центральные моменты*

$$a_k = M(\xi - m)^k,$$

где $m = m_1$ — математическое ожидание ξ.

Можно показать, что если случайная величина ξ сосредоточена на конечном промежутке, то по ее моментам однозначно восстанавливается функция распределения.

Среди центральных моментов особо выделяют момент второго порядка.

Определение 6.3. Дисперсией случайной величины ξ называется число

$$D\xi = M(\xi - m)^2,$$

где m — математическое ожидание ξ.

Из определения дисперсии видно, что она характеризует меру отклонения случайной величины от своего среднего значения. Другими словами, дисперсия является характеристикой разброса случайной величины. Если ξ — некоторая физическая величина, то дисперсия имеет размерность, равную квадрату размерности ξ. Поэтому наряду с $D\xi$ часто рассматривают так называемое *стандартное отклонение* случайной величины ξ, равное $\sqrt{D\xi}$.

Установим свойства дисперсии.

Свойство 1. Если случайная величина ξ является постоянной, то $D\xi = 0$.

Доказательство. Пусть $\xi \equiv C$, тогда по свойству 1 математических ожиданий $M\xi = C$ и, следовательно,

$$D\xi = M(\xi - C)^2 = M(C - C)^2 = 0.$$

Свойство 2. Добавление постоянной к случайной величине не меняет ее дисперсии, т. е.

$$D(\xi + C) = D\xi.$$

Доказательство. По свойствам 1 и 3 математического ожидания $M(\xi + C) = M\xi + C$, поэтому

$$D(\xi+C) = M((\xi+C) - M(\xi+C))^2 = M(\xi+C-M\xi-C)^2 =$$

$$= M(\xi - M\xi)^2 = D\xi.$$

Свойство 3. Постоянный множитель выносится за знак дисперсии в квадрате, т. е.

$$D(C\xi) = C^2 D\xi.$$

Доказательство. По свойству 2 математического ожидания $M(C\xi) = CM\xi$, поэтому

$$D(C\xi) = M((C\xi) - M(C\xi))^2 = M(C\xi - CM\xi)^2 =$$

$$= M\left[C^2(\xi - M\xi)^2\right] = C^2 M(\xi - M\xi)^2 = C^2 D\xi.$$

Свойство 4. Дисперсия произвольной случайной величины неотрицательна, причем она равна нулю тогда и только тогда, когда случайная величина ξ является постоянной с вероятностью 1.

Доказательство. Дисперсия случайной величины ξ есть математическое ожидание неотрицательной случайной величины $(\xi - m)^2$, где $m = M\xi$. Поэтому по свойству 4 математического ожидания $D\xi \geqslant 0$, причем равенство будет, если $(\xi - m)^2 = 0$ с вероятностью 1, а это означает, что $\xi = m$ с вероятностью 1.

Свойство 5. Для произвольной случайной величины ξ

$$D\xi = M(\xi^2) - (M\xi)^2.$$

Доказательство. Пусть $M\xi = m$, тогда по свойствам 1–3 математического ожидания

$$D\xi = M(\xi - m)^2 = M(\xi^2 - 2m\xi + m^2) = M(\xi^2) - 2mM\xi + m^2 =$$

$$= M(\xi^2) - 2m^2 + m^2 = M(\xi^2) - m^2 = M(\xi^2) - (M\xi)^2.$$

Следствие 6.1. Для произвольной случайной величины

$$M(\xi^2) \geqslant (M\xi)^2.$$

Доказательство. Требуемое неравенство следует из свойств 4 и 5.

Объясним, как вычислять дисперсию. Пусть ξ — дискретная случайная величина с рядом распределения

ξ	a_1	a_2	...	a_n	...
P	p_1	p_2	...	p_n	...

По свойству 5 математического ожидания

$$D\xi = M(\xi - m)^2 = \sum_{n=1}^{\infty}(a_n - m)^2 p_n,$$

где $m = M\xi$ (здесь к ξ применена функция $f(x) = (x-m)^2$). Если воспользоваться свойством 5 дисперсии, то получаем следующую формулу:

$$D\xi = M(\xi^2) - m^2 = \sum_{n=1}^{\infty} a_n^2 p_n - m^2.$$

Если же ξ — непрерывная случайная величина с плотностью вероятностей $p(x)$, то

$$D\xi = \int_{-\infty}^{+\infty}(x-m)^2 p(x)dx = \int_{-\infty}^{+\infty} x^2 p(x)dx - m^2,$$

где $m = M\xi$.

ПРИМЕР 6.4. Пусть ξ — дискретная случайная величина с рядом распределения

ξ	-1	2	4
P	$1/2$	$1/4$	$1/4$

Найти $M\xi$ и $D\xi$.

1) $M\xi = (-1)\cdot\frac{1}{2} + 2\cdot\frac{1}{4} + 4\cdot\frac{1}{4} = 1.$

2) $D\xi = M(\xi-1)^2 = (-1-1)^2\frac{1}{2} + (2-1)^2\frac{1}{4} + (4-1)^2\frac{1}{4} =$

$$= 2 + \frac{1}{4} + \frac{9}{4} = 4{,}5.$$

ПРИМЕР 6.5. Пусть $\xi \sim U[a,b]$. Найти $M\xi$ и $D\xi$.

1) $M\xi = \int_{-\infty}^{+\infty} xp(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\cdot\frac{x^2}{2}\Big|_a^b = \frac{a+b}{2}.$

2) $$D\xi = \int_{-\infty}^{+\infty}\left(x - \frac{a+b}{2}\right)^2 p(x)dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}dx =$$

$$= \left(\text{замена } x - \frac{a+b}{2} = t\right) = \frac{1}{b-a}\int_{(a-b)/2}^{(b-a)/2} t^2 dt =$$

$$= \frac{1}{3(b-a)}t^3\Big|_{(a-b)/2}^{(b-a)/2} = \frac{2}{3(b-a)}\left(\frac{b-a}{2}\right)^3 = \frac{(b-a)^2}{12}.$$

Контрольные вопросы

1. Каков вероятностный смысл математического ожидания?
2. Почему в формуле для математического ожидания в дискретном случае предполагается, что ряд сходится абсолютно?
3. Докажите, что принятое в математическом анализе определение среднего значения интегрируемой функции $f(x)$ на отрезке $[a,b]$ совпадает с $Mf(\xi)$, где $\xi \sim U[a,b]$.
4. Каков вероятностный смысл дисперсии случайной величины?
5. Как вы думаете, верно ли равенство $D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2$?
6. Докажите неравенство

$$\sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} \geqslant \left(\sum_{k=0}^{n} k C_n^k p^k q^{n-k}\right)^2.$$

Ответы

5. Это равенство, вообще говоря, неверно. Например, если $\xi_1 = \xi_2$, то $D(\xi_1 + \xi_2) = D(2\xi_1) = 4D\xi_1 \neq 2D\xi_1$.

6. Это неравенство вытекает из следствия 6.1, если рассмотреть случайную величину ξ, имеющую биноминальное распределение.

Лекция 7

СЛУЧАЙНЫЕ ВЕКТОРЫ

§ 7.1. ОПРЕДЕЛЕНИЕ СЛУЧАЙНОГО ВЕКТОРА. ЕГО ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

В некоторых случаях более целесообразно подводить итог случайного эксперимента в виде нескольких чисел. Например, при стрельбе по плоской мишени требуется знать обе координаты точки попадания. Таким образом, каждому элементарному исходу ω сопоставляется не одно, а сразу несколько чисел $\xi_1(\omega)$, $\xi_2(\omega)$, ... $\xi_k(\omega)$.

Определение 7.1. Пусть на одном и том же вероятностном пространстве задано несколько случайных величин ξ_1, ξ_2, ... ξ_k. Тогда говорят, что задан *случайный вектор* $\bar{\xi} = \{\xi_1, \xi_2, ... \xi_k\}$.

Как и в случае одной случайной величины в теории вероятностей важно знать вероятности $P(\bar{\xi} \in B)$, где B — произвольное подмножество k-мерного арифметического пространства R^k. С этой целью введем понятие функции распределения случайного вектора.

Определение 7.2. Функцией распределения случайного вектора $\bar{\xi} = \{\xi_1, \xi_2, ... \xi_k\}$ называется числовая функция нескольких числовых переменных $F(x_1, ..., x_k)$, задаваемая равенством

$$F(x_1, ..., x_k) = P(\xi_1 \leqslant x_1, \xi_2 \leqslant x_2, ..., \xi_k \leqslant x_k).$$

При этом предполагается, что множество

$$\{\omega : \xi_1(\omega) \leqslant x_1, \xi_2(\omega) \leqslant x_2, ..., \xi_k(\omega) \leqslant x_k\}$$

принадлежит алгебре случайных событий.

Функция распределения случайного вектора обладает свойствами, схожими со свойствами функции распределения случайной величины. Более того, доказательства этих свойств для случайного вектора мало чем отличаются от доказательств для случайной величины, поэтому некоторые из них опустим.

Свойство 1. Функция распределения $F(x_1, x_2, ..., x_k)$ не убывает по каждой переменной $x_1, x_2, ..., x_k$.

Свойство 2. Для каждой переменной $x_i, i = 1, 2, ..., k$,

$$\lim_{x_i \to -\infty} F(x_1, x_2, ..., x_k) = 0.$$

Свойство 3.

$$\lim_{\substack{x_1 \to +\infty, \\ \dots \\ x_k \to +\infty}} F(x_1, x_2, ..., x_k) = 1.$$

Свойство 4. Функция распределения непрерывна по каждой переменной справа.

Свойство 5. Если устремить к $+\infty$ все переменные, кроме x_i, то предел $F(x_1, ..., x_k)$ совпадет с функцией распределения случайной величины $\xi_i (i = 1, ..., k)$.

Доказательство. Покажем, например, что при $k = 2$

$$P(\xi_1 \leqslant x_1) = \lim_{x_2 \to +\infty} F(x_1, x_2). \tag{7.1}$$

Рассмотрим случайные события $A_i = \{\xi_1 \leqslant x_1, \xi_2 \leqslant i\}, i = 1, 2, ...$ Ясно, что $A_1 \subset A_2 \subset A_3 \subset ...$ и $\bigcup\limits_{i=1}^{\infty} A_i = \{\xi_1 \leqslant x_1\}$.

Поэтому по аксиоме непрерывности

$$\lim_{i \to \infty} F(x_1, i) = \lim_{i \to \infty} P(A_i) = P(\bigcup_{i=1}^{\infty} A_i) = P(\xi_1 \leqslant x_1),$$

откуда, учитывая свойство 1, получаем (7.1).

Замечание 7.1. Свойство 5 показывает, что зная распределение случайного вектора $\bar{\xi}$, можно найти распределения отдельных компонент этого вектора (их называют маргинальными распределениями). Однако по маргинальным распределениям нельзя восстановить однозначно распределение самого случайного вектора $\bar{\xi}$.

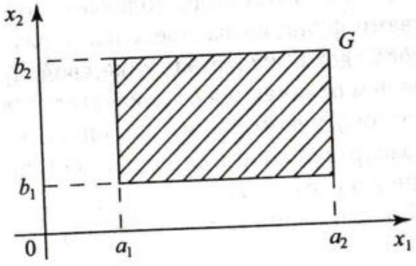


Рис. 7.1

Знание функции распределения позволяет находить различные вероятности, связанные со случайным вектором $\bar{\xi}$. Покажем, например, как в случае $k = 2$ по $F(x_1, x_2)$ найти вероятность того, что $\bar{\xi}$ попадет в прямоугольник G, изображенный на рис. 7.1.

Утверждение 7.1. Справедлива следующая формула

$$P(a_1 < \xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2) = \\ = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1).$$

Доказательство. Нетрудно понять, что

$$\{\xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2\} = \{\xi_1 < a_1, b_1 < \xi_2 \leqslant b_2\} + \\ + \{a_1 < \xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2\},$$

причем случайные события в правой части несовместны, поэтому по аксиоме $P3$

$$P(\xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2) = P(\xi_1 \leqslant a_1, b_1 < \xi_2 \leqslant b_2) + \\ + P(a_1 < \xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2). \quad (7.2)$$

Аналогично можно показать, что

$$P(\xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2) = P(\xi_1 \leqslant a_2, \xi_2 \leqslant b_2) - \\ - P(\xi_1 \leqslant a_2, \xi_2 \leqslant b_1) = F(a_2, b_2) - F(a_2, b_1), \quad (7.3)$$

$$P(\xi_1 \leqslant a_1, b_1 < \xi_2 \leqslant b_2) = F(a_1, b_2) - F(a_1, b_1). \quad (7.4)$$

Подставляя формулы (7.3) и (7.4) в соотношение (7.2), получаем, что

$$F(a_2, b_2) - F(a_2, b_1) = F(a_1, b_2) - F(a_1, b_1) + \\ + P(a_1 < \xi_1 \leqslant a_2, b_1 < \xi_2 \leqslant b_2),$$

откуда следует требуемое утверждение.

Разбивая произвольную область $B \subset R^2$, для которой определено понятие площади, на достаточно маленькие прямоугольники, мы можем в виду утверждения 1, зная функцию распределения, вычислить вероятность того, что случайный вектор $\bar{\xi}$ окажется в B.

§ 7.2. ДИСКРЕТНЫЕ И НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕКТОРЫ

Определение 7.3. Случайный вектор $\bar{\xi} = \{\xi_1, ..., \xi_k\}$ называется *дискретным*, если он принимает конечное или счетное число значений $\bar{a}_1, \bar{a}_2, ...$ ($\bar{a}_1, \bar{a}_2, ...$ принадлежат R^k).

Определение 7.4. Рядом распределения случайного вектора $\bar{\xi}$ называется таблица, в верхней строке которой записаны возможные значения $\bar{\xi}$, а в нижней — вероятности, с которыми эти значения принимает случайный вектор ($p_l = P(\bar{\xi} = \bar{a}_l)$, $l = 1, 2, ...$):

$\bar{\xi}$	$\bar{a}_1$	$\bar{a}_2$	...	$\bar{a}_l$	...
P	p_1	p_2	...	p_l	...

Как и в одномерном случае показывается, что $p_l \geqslant 0$ при всех $l = 1, 2, ...$ и $\sum_{l=1}^{\infty} p_l = 1$. Кроме того, для произвольного множества $B \subset R^k$

$$P(\bar{\xi} \in B) = \sum_{\substack{l: \\ \bar{a}_l \in B}} p_l.$$

Определение 7.5. Случайный вектор $\bar{\xi} = \{\xi_1, ..., \xi_k\}$ называется *непрерывным*, если он принимает все значения из некоторой области D (ненулевого объема) и существует такая неотрицательная функция нескольких переменных $p(x_1, ..., x_k)$, что для любого параллелепипеда $G = \{(x_1, ..., x_k) : x_i \in [a_i, b_i], i = 1, ..., k\}$

$$P(\bar{\xi} \in G) = \int_G \cdots \int p(x_1, ..., x_k) dx_1 ... dx_k. \quad (7.5)$$

Замечание 7.2. Чтобы избежать трудностей, связанных с многомерным интегралом, читатель может считать, что $k = 2$ или $k = 3$.

Плотность вероятностей $p(x_1, ..., x_k)$ неотрицательна и

$$\int\limits_{R^k} ... \int p(x_1, ..., x_k) dx_1 ... dx_k = 1.$$

Разбивая произвольную область $B \subset R^k$, для которой определено понятие объема, на достаточно маленькие параллелепипеды и применяя к каждому из них формулу (7.5), получаем, что

$$P(\bar{\xi} \in B) = \int\limits_B ... \int p(x_1, ..., x_k) dx_1 ... dx_k.$$

Как и в одномерном случае, плотность вероятностей случайного вектора и его функция распределения $F(x_1, ..., x_k)$ связаны между собой:

$$F(x_1, ..., x_k) = \int\limits_{-\infty}^{x_1} ... \int\limits_{-\infty}^{x_k} p(u_1, ..., u_k) du_1 ... du_k,$$

$$p(x_1, ..., x_k) = \frac{\partial^k F(x_1, ..., x_k)}{\partial x_1 ... \partial x_k}$$

(последнее равенство выполняется в точках непрерывности $p(x_1, ..., x_k)$).

Приведем примеры непрерывных распределений.

а) *Равномерное распределение.* Случайный вектор $\bar{\xi} = \{\xi_1, ..., \xi_k\}$ имеет равномерное распределение в области $D \subset R^k$, если он является непрерывным и его плотность вероятностей равна

$$p(x_1, ..., x_k) = \begin{cases} C, (x_1, ..., x_k) \in D, \\ 0, (x_1, ..., x_k) \notin D, \end{cases}$$

где C — постоянная. Нетрудно вычислить, что постоянная C равна величине, обратной объему области D.

Случайный вектор, равный радиусу-вектору точки, наудачу выбранной в области D, имеет указанное распределение.

б) *Нормальное распределение* ($k = 2$). Случайный вектор $\bar{\xi} = \{\xi_1, \xi_2\}$ имеет нормальное распределение, если он является непрерывным и его плотность вероятностей равна

$$p(x_1, x_2) = Ce^{-k(x_1 - a_1, x_2 - a_2)/2},$$

где $k(x_1 - a_1, x_2 - a_2)$ — квадратичная форма относительно $x_1 - a_1$ и $x_2 - a_2$, т. е.

$$k(x_1 - a_1, x_2 - a_2) = a_{11}(x_1 - a_1)^2 + \\ + 2a_{12}(x_1 - a_1)(x_2 - a_2) + a_{22}(x_2 - a_2)^2$$

(здесь a_1, a_2, a_{11}, a_{12}, a_{22} — постоянные).

При этом предполагается, что $k(x_1 - a_1, x_2 - a_2) > 0$, если x_1 и x_2 не равны одновременно a_1 и a_2 соответственно. Можно показать, что

$$C = \frac{1}{2\pi}\sqrt{|A|},$$

где

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}.$$

В теории артиллерийских стрельб считается, что радиус-вектор точки попадания снаряда имеет нормальное распределение.

§ 7.3. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНОГО ВЕКТОРА

*Определение 7.6. **Математическим ожиданием*** (или ***центром рассеяния***) случайного вектора $\bar{\xi} = \{\xi_1, \xi_2, ..., \xi_k\}$ называется вектор

$$M\bar{\xi} = \{M\xi_1, M\xi_2, ..., M\xi_k\}.$$

*Определение 7.7. **Ковариационной матрицей*** случайного вектора $\bar{\xi} = \{\xi_1, \xi_2, ..., \xi_k\}$ называется матрица

$$M(\bar{\xi} - \bar{m})^T(\bar{\xi} - \bar{m}),$$

где $\bar{m} = M\bar{\xi}$ (чтобы найти математическое ожидание матрицы $(\bar{\xi} - \bar{m})^T(\bar{\xi} - \bar{m})$ надо найти математическое ожидание каждой ее компоненты).

Очевидно, ковариационная матрица имеет размеры $k \times k$ и ее элемент, стоящий в пересечении i-й строки и j-го столбца, равен

$$a_{ij} = M(\xi_i - m_i)(\xi_j - m_j),$$

где m_i — i-я координата вектора $\bar{m}$, т. е. $m_i = M\xi_i$.

Определение 7.8. Ковариацией двух случайных величин ξ и η называется число

$$\mathrm{cov}(\xi, \eta) = M(\xi - a)(\eta - b),$$

где $a = M\xi, b = M\eta$.

Таким образом, a_{ij} является ковариацией ξ_i и ξ_j. Следовательно, ковариационная матрица случайного вектора $\bar{\xi}$ — это матрица, составленная из ковариаций его компонент (отсюда и название).

Очевидно, ковариационная матрица является симметричной и на ее диагонали находятся дисперсии случайных величин $\xi_1, ..., \xi_k$.

Разберемся с вероятностным смыслом ковариации $\mathrm{cov}(\xi, \eta)$, но для этого сначала заметим, что размерность ковариации равна произведению размерностей случайных величин ξ и η. Поэтому разумно вместо ковариации рассматривать безразмерную величину

$$r(\xi, \eta) = \frac{\mathrm{cov}(\xi, \eta)}{\sqrt{D\xi D\eta}},$$

которая называется *коэффициентом корреляции* случайных величин ξ и η (предполагается, что $D\xi \neq 0$, $D\eta \neq 0$). Коэффициент корреляции является мерой зависимости случайных величин ξ и η (об этом см. следующую лекцию). Укажем свойства коэффициента корреляции.

Свойство 1. Справедливо равенство

$$r(\xi, \eta) = r(\eta, \xi).$$

Свойство 2. Пусть $\tilde{\xi} = a_1\xi + b_1, \tilde{\eta} = a_2\xi + b_2$, где a_1, a_2, b_1, b_2 — постоянные, причем $a_1 > 0, a_2 > 0$. Тогда

$$r(\tilde{\xi}, \tilde{\eta}) = r(\xi, \eta).$$

Свойство 3. Справедливо неравенство

$$|r(\xi, \eta)| \leqslant 1, \tag{7.6}$$

причем равенство возможно тогда и только тогда, когда с вероятностью единица $\eta = c\xi + d$, где c, d — постоянные, причем $c \neq 0$.

Доказательство. В силу свойства 2 достаточно доказать это утверждение для случайных величин $\tilde{\xi} = \xi - M\xi$ и $\tilde{\eta} = \eta - M\eta$. Заметим, что $M\tilde{\xi} = M\tilde{\eta} = 0$, поэтому

$$\mathrm{cov}(\tilde{\xi}, \tilde{\eta}) = M\tilde{\xi}\tilde{\eta}, D\tilde{\xi} = M(\tilde{\xi}^2), D\tilde{\eta} = M(\tilde{\eta}^2). \tag{7.7}$$

Рассмотрим случайную величину $\zeta = \lambda\tilde{\xi} + \tilde{\eta}$. Очевидно,

$$0 \leqslant M(\zeta^2) = M(\lambda\tilde{\xi} + \tilde{\eta})^2 = \lambda^2 M(\tilde{\xi}^2) + 2\lambda M\tilde{\xi}\tilde{\eta} + M(\tilde{\eta}^2).$$

Относительно λ последнее выражение является квадратным трехчленом. Из неотрицательности этого квадратного трехчлена следует, что его дискриминант меньше или равен нулю, т. е.

$$(M\tilde{\xi}\tilde{\eta})^2 - M(\tilde{\xi})^2 M(\tilde{\eta})^2 \leqslant 0.$$

Вспоминая (7.7), получаем, что

$$|\mathrm{cov}(\tilde{\xi}, \tilde{\eta})| \leqslant \sqrt{D\tilde{\xi} D\tilde{\eta}}.$$

Но это как раз означает справедливость (7.6).

Если $|r(\xi, \eta)| = 1$, то дискриминант упомянутого квадратного трехчлена равен 0 и, следовательно, существует корень $\lambda_0 \neq 0$ этого трехчлена. Но это означает, что $M(\lambda_0\tilde{\xi} + \tilde{\eta})^2 = 0$. По свойству 4 математического ожидания $\lambda_0\tilde{\xi} + \tilde{\eta} = 0$ с вероятностью единица, ч. т. д.

Покажем, как в непрерывном случае вычисляются числовые характеристики случайного вектора. Для этого воспользуемся следующим утверждением, аналогичным свойству 5 математического ожидания случайной величины (см. лекцию № 6).

Утверждение 7.2. Если $\bar{\xi} = \{\xi_1, ..., \xi_k\}$ — непрерывный случайный вектор с плотностью вероятностей $p(x_1, ..., x_k)$ и $f(x_1, ..., x_k)$ — числовая функция нескольких числовых переменных, то $\eta = f(\xi_1, ..., \xi_k)$ является случайной величиной и

$$M\eta = \int\limits_{R^k} ... \int f(x_1, ..., x_k) p(x_1, ..., x_k) dx_1 ... dx_k.$$

Из этого утверждения следует, что

$$m_i = M\xi_i = \int\limits_{R^k} ... \int x_i p(x_1, ..., x_k) dx_1 ... dx_k, \quad i = 1, ..., k; \tag{7.8}$$

$$a_{ij} = M(\xi_i - m_i)(\xi_j - m_j) =$$

$$= \int\limits_{R^k} ... \int (x_i - m_i)(x_j - m_j) p(x_1, ..., x_k) dx_1 ... dx_k, \quad i, j = 1, ..., k. \tag{7.9}$$

Легко проверить, что

$$\operatorname{cov}(\xi, \eta) = M\xi\eta - M\xi M\eta,$$

поэтому a_{ij} можно вычислять иначе:

$$a_{ij} = \int\limits_{R^k} ... \int x_i x_j p(x_1, ..., x_k) dx_1 ... dx_k - m_i m_j, i, j = 1, ..., k. \tag{7.10}$$

Математическое ожидание случайного вектора $\bar{\xi}$ есть вектор $\bar{m} = \{m_1, ..., m_k\}$, а ковариационная матрица $\bar{\xi}$ — матрица $A = \{a_{ij}, \; i, j = 1, ..., k\}$.

ПРИМЕР 7.1. Пусть точка наудачу выбирается в треугольнике D, изображенном на рис. 7.2, и $\bar{\xi} = \{\xi_1, \xi_2\}$ — радиус-вектор этой точки. Найти $M\bar{\xi}$ и ковариационную матрицу $\bar{\xi}$.

Вектор $\bar{\xi}$ имеет равномерное распределение в треугольнике D, т. е. является непрерывным и его плотность вероятностей равна

$$p(x_1, x_2) = \begin{cases} 2, (x_1, x_2) \in D; \\ 0, (x_1, x_2) \notin D, \end{cases}$$

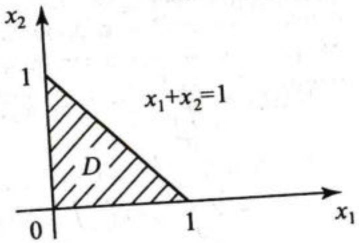


Рис. 7.2

поскольку площадь треугольника D равна 1/2. Сначала найдем $M\bar{\xi}$:

$$M\xi_1 = \iint\limits_{R^2} x_1 p(x_1, x_2) dx_1 dx_2 = \iint\limits_{D} 2x_1 dx_1 dx_2 =$$

$$= 2\int\limits_0^1 dx_1 \int\limits_0^{1-x_1} x_1 dx_2 = 2\int\limits_0^1 x_1(1 - x_1) dx_1 =$$

$$= 2\left(\frac{x_1^2}{2} - \frac{x_1^3}{3}\right)\bigg|_0^1 = 2\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{3},$$

аналогично $M\xi_2 = \dfrac{1}{3}$. Итак,

$$M\bar{\xi} = \left\{\frac{1}{3}; \frac{1}{3}\right\}.$$

Для нахождения ковариационной матрицы найдем $M\xi_1^2, M\xi_2^2$ и $M\xi_1\xi_2$:

$$M\xi_1^2 = \iint\limits_{R^2} x_1^2 p(x_1, x_2) dx_1 dx_2 = \iint\limits_{D} 2x_1^2 dx_1 dx_2 =$$

$$= 2\int\limits_0^1 dx_1 \int\limits_0^{1-x_1} x_1^2 dx_2 = 2\int\limits_0^1 x_1^2(1 - x_1) dx_1 =$$

$$= 2\left(\frac{x_1^3}{3} - \frac{x_1^4}{4}\right)\bigg|_0^1 = 2\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{1}{6};$$

аналогично $M\xi_2^2 = \dfrac{1}{6}$;

$$M\xi_1\xi_2 = \iint\limits_{R^2} x_1x_2p(x_1,x_2)dx_1dx_2 = \iint\limits_{D} 2x_1x_2dx_1dx_2 =$$

$$= 2\int_0^1 x_1dx_1 \int_0^{1-x_1} x_2dx_2 = \int_0^1 x_1(1-x_1)^2dx_1 =$$

$$= \int_0^1 (x_1 - 2x_1^2 + x_1^3)dx_1 = \frac{1}{2} - \frac{2}{3} + \frac{1}{4} = \frac{1}{12}.$$

Следовательно,

$$D\xi_1 = M(\xi_1^2) - (M\xi_1)^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}, \quad D\xi_2 = \frac{1}{18},$$

$$\mathrm{cov}(\xi_1,\xi_2) = M\xi_1\xi_2 - M\xi_1M\xi_2 = \frac{1}{12} - \left(\frac{1}{3}\right)^2 = -\frac{1}{36}.$$

Итак, ковариационная матрица случайного вектора $\bar{\xi}$ имеет вид

$$\begin{pmatrix} D\xi_1 & \mathrm{cov}(\xi_1,\xi_2) \\ \mathrm{cov}(\xi_1,\xi_2) & D\xi_2 \end{pmatrix} = \begin{pmatrix} 1/18 & -1/36 \\ -1/36 & 1/18 \end{pmatrix}.$$

Контрольные вопросы

1. Объясните, почему по маргинальным распределениям неоднозначно восстанавливается распределение случайного вектора. Приведите пример.

2. Пусть $F(x_1,x_2,x_3)$ — функция распределения случайного вектора $\{\xi_1,\xi_2,\xi_3\}$. Как по ней найти функцию распределения случайного вектора $\{\xi_1,\xi_2\}$?

3. Докажите, что

$$\mathrm{cov}(\xi,\eta) = M\xi\eta - M\xi M\eta.$$

4. Докажите свойство 1 коэффициента корреляции.

5. Докажите свойство 2 коэффициента корреляции. Что будет, если рассмотреть другие варианты знаков a_1 и a_2?

6. Докажите, что если $r(\xi,\eta) = 1$, то $\eta = c\xi + d$, где c, d — постоянные, причем $c > 0$. Если же $r(\xi,\eta) = -1$, то $c < 0$.

Ответы

2. Функция распределения $\{\xi_1,\xi_2\}$ равна $\lim\limits_{x_3\to+\infty} F(x_1,x_2,x_3)$.

3. $\mathrm{cov}(\xi,\eta) = M(\xi - M\xi)(\eta - M\eta) = M(\xi\eta - \eta M\xi - \xi M\eta + M\xi M\eta) = M\xi\eta - M\eta M\xi - M\xi M\eta + M\xi M\eta = M\xi\eta - M\xi M\eta$.

5. Если, например, $a_1 > 0$, $a_2 < 0$, то $r(\bar{\xi},\bar{\eta}) = -r(\xi,\eta)$.

Лекция 8

ФУНКЦИИ ОТ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 8.1. НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Определение 8.1. Случайные величины $\xi_1, \xi_2, ..., \xi_k$ (определенные на одном и том же вероятностном пространстве) называются ***независимыми***, если для любых отрезков $[a_i, b_i]$, $i = 1, ..., k$, выполняется равенство

$$P(\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2], ..., \xi_k \in [a_k, b_k]) =$$

$$= P(\xi_1 \in [a_1, b_1]) P(\xi_2 \in [a_2, b_2]) \cdot ... \cdot P(\xi_k \in [a_k, b_k]).$$

ПРИМЕР 8.1. Пусть точка наудачу выбирается в квадрате K, изображенном на рис. 8.1, и $\bar{\xi} = \{\xi_1, \xi_2\}$ — радиус-вектор этой точки. Тогда ξ_1 и ξ_2 — независимые случайные величины.

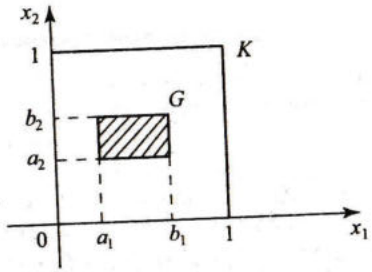


Рис. 8.1

Действительно, выберем произвольные отрезки $[a_1, b_1]$ и $[a_2, b_2]$ $(0 \leqslant a_1 < b_1 \leqslant 1$, $0 \leqslant a_2 < b_2 \leqslant 1)$. Тогда по определению геометрических вероятностей

$$P(\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2]) =$$
$$= P(\bar{\xi} \in G) = S(G) =$$
$$= (b_1 - a_1)(b_2 - a_2),$$

где G — прямоугольник $\{(x_1, x_2) : a_1 \leqslant x_1 \leqslant b_1, a_2 \leqslant x_2 \leqslant b_2\}$, а $S(G)$ — его площадь. Аналогично,

$$P(\xi_1 \in [a_1, b_1]) = P(\bar{\xi} \in \{(x_1, x_2) : a_1 \leqslant x_1 \leqslant b_1, 0 \leqslant x_2 \leqslant 1\}) =$$
$$= (b_1 - a_1)(1 - 0) = b_1 - a_1,$$

$$P(\xi_2 \in [a_2, b_2]) = b_2 - a_2.$$

Итак,

$$P(\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2]) = P(\xi_1 \in [a_1, b_1]) P(\xi_2 \in [a_2, b_2]),$$

что и требовалось доказать.

ПРИМЕР 8.2. Пусть точка наудачу выбирается в треугольнике D, изображенном на рис. 8.2, и $\bar{\xi} = \{\xi_1, \xi_2\}$ — радиус-вектор этой точки. Тогда случайные величины ξ_1 и ξ_2 являются зависимыми.

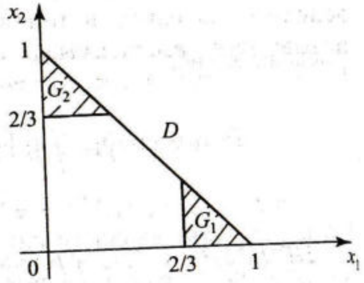


Рис. 8.2

Действительно, если взять $a_1 = 2/3$, $b_1 = 1$, $a_2 = 2/3$, $b_2 = 1$, то

$$P(\xi_1 \in [a_1, b_1]) = P(\bar{\xi} \in G_1) =$$
$$= S(G_1)/S(D) = (1/3)^2 = 1/9,$$

аналогично $P(\xi_2 \in [a_2, b_2]) = = P(\xi \in G_2) = 1/9$, но

$$P(\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2]) =$$
$$= P(\bar{\xi} \in \phi)) = 0.$$

Следовательно,

$$P(\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2]) \neq P(\xi_1 \in [a_1, b_1]) P(\xi_2 \in [a_2, b_2]).$$

Утверждение 8.1. Если $\xi_1, \xi_2, ..., \xi_k$ — независимые непрерывные случайные величины с плотностями вероятностей $p_1(\cdot), p_2(\cdot), ..., p_k(\cdot)$, то случайный вектор $\bar{\xi} = \{\xi_1, \xi_2, ..., \xi_k\}$ является непрерывным и его плотность вероятностей записывается в виде

$$p(x_1, x_2, ..., x_k) = p_1(x_1) p_2(x_2) \cdot ... \cdot p_k(x_k).$$

Доказательство (для $k=2$). Рассмотрим произвольный прямоугольник $G=\{(x_1,x_2):\ x_1\in[a_1,b_1],\ x_2\in[a_2,b_2]\}$. Тогда ввиду независимости ξ_1 и ξ_2

$$P(\bar\xi\in G)=P(\xi_1\in[a_1,b_1],\xi_2\in[a_2,b_2])=$$
$$=P(\xi_1\in[a_1,b_1])P(\xi_2\in[a_2,b_2]).$$

По определению плотности вероятностей

$$P(\xi_1\in[a_1,b_1])=\int\limits_{a_1}^{b_1}p_1(x_1)dx_1,$$

$$P(\xi_2\in[a_2,b_2])=\int\limits_{a_2}^{b_2}p_2(x_2)dx_2.$$

Следовательно,

$$P(\bar\xi\in G)=\int\limits_{a_1}^{b_1}p_1(x_1)dx_1\cdot\int\limits_{a_2}^{b_2}p_2(x_2)dx_2.$$

Но по свойству двойного интеграла произведение интегралов в правой части равно $\iint\limits_G p_1(x_1)p_2(x_2)dx_1dx_2$. Итак,

$$P(\bar\xi\in G)=\iint\limits_G p_1(x_1)p_2(x_2)dx_1dx_2,$$

но это означает, что случайный вектор $\bar\xi$ является непрерывным с плотностью вероятностей $p(x_1,x_2)=p_1(x_1)p_2(x_2)$, ч. т. д.

Теорема 8.1. *Если $\xi_1,\xi_2,...,\xi_k$ — независимые случайные величины с конечными математическими ожиданиями, то*

$$M\xi_1\xi_2\cdot...\cdot\xi_k=M\xi_1\cdot M\xi_2\cdot...\cdot M\xi_k.$$

Доказательство (в непрерывном случае, $k=2$). Ввиду утверждения 8.1 случайный вектор $\{\xi_1,\xi_2\}$ имеет плотность

вероятностей $p(x_1,x_2)=p_1(x_1)p_2(x_2)$, где $p_1(\cdot),p_2(\cdot)$ — плотности вероятностей случайных величин ξ_1 и ξ_2. В силу утверждения 7.2

$$M\xi_1\xi_2=\iint\limits_{R^2}x_1x_2p(x_1,x_2)dx_1dx_2=$$

$$=\iint\limits_{R^2}x_1x_2p_1(x_1)p_2(x_2)dx_1dx_2=\int\limits_{R^1}x_1p_1(x_1)dx_1\int\limits_{R^1}x_2p_2(x_2)dx_2,$$

но последние два интеграла есть $M\xi_1$ и $M\xi_2$. Теорема доказана.

Теорема 8.2. *Если ξ и η — независимые случайные величины, то $cov(\xi,\eta)=0$ и $r(\xi,\eta)=0$.*

Доказательство. Известно, что

$$\mathrm{cov}(\xi,\eta)=M\xi\eta-M\xi\cdot M\eta.$$

По теореме 8.1, ввиду независимости ξ и η,

$$M\xi\eta=M\xi\cdot M\eta,$$

откуда следует, что

$$\mathrm{cov}(\xi,\eta)=0,$$

а следовательно, $r(\xi,\eta)=0$.

Замечание 8.1. Ранее установлено, что равенство $|r(\xi,\eta)|=1$ означает, что одна случайная величина линейно выражается через другую (см. свойство 3 коэффициента корреляции из лекции № 7). Независимость ξ и η влечет равенство $r(\xi,\eta)=0$. Именно эти два свойства коэффициента корреляции наводят на мысль использовать его как меру зависимости случайных величин: близость модуля коэффициента корреляции к единице означает сильную зависимость случайных величин, а близость к нулю — слабую зависимость случайных величин.

Следствие 8.1. Если компоненты $\xi_1,\xi_2,...,\xi_n$ случайного вектора $\bar\xi$ независимы, то его ковариационная матрица является диагональной.

Доказательство. Как мы знаем, ковариационная матрица состоит из элементов $a_{ij} = \operatorname{cov}(\xi_i, \xi_j)$. Ввиду независимости ξ_i и ξ_j при $i \neq j$, $a_{ij} = 0$, ч. т. д.

Ранее отмечалось, что дисперсия суммы случайных величин, вообще говоря, не равна сумме дисперсий (см. лекцию № 6). На самом деле, справедлив следующий факт.

Утверждение 8.2. Для произвольных случайных величин $\xi_1, \xi_2, ..., \xi_k$ (заданных на одном и том же вероятностном пространстве):

$$D(\xi_1 + \xi_2 + ... + \xi_k) = \sum_{i,j=1}^{k} \operatorname{cov}(\xi_i, \xi_j). \tag{8.1}$$

Доказательство. Положим $M\xi_i = m_i$, $i = 1, ..., k$. Тогда $M(\xi_1 + \xi_2 + ... + \xi_k) = m_1 + m_2 + ... + m_k$. Следовательно,

$$D(\xi_1+\xi_2+...+\xi_k) = M((\xi_1+\xi_2+...+\xi_k)-(m_1+m_2+...+m_k))^2 =$$

$$= M((\xi_1 - m_1) + (\xi_2 - m_2) + ... + (\xi_k - m_k))^2 =$$

$$= \sum_{i,j=1}^{k} M((\xi_i - m_i)(\xi_j - m_j)) = \sum_{i,j=1}^{k} \operatorname{cov}(\xi_i, \xi_j).$$

Теорема 8.3. *Если $\xi_1, \xi_2, ..., \xi_k$ — независимые случайные величины, то*

$$D(\xi_1 + \xi_2 + ... + \xi_k) = D\xi_1 + D\xi_2 + ... + D\xi_k.$$

Доказательство. Ввиду независимости ξ_i и ξ_j при $i \neq j$, $\operatorname{cov}(\xi_i, \xi_j) = 0$, поэтому из (8.1) получаем, что

$$D(\xi_1 + \xi_2 + ... + \xi_k) = \sum_{i=1}^{k} \operatorname{cov}(\xi_i, \xi_i).$$

Осталось вспомнить, что $\operatorname{cov}(\xi_i, \xi_i) = D\xi_i$, $i = 1, ..., k$. Теорема доказана.

ПРИМЕР 8.3. Пусть проводятся испытания Бернулли с вероятностью успеха в одном испытании p и μ_n — число успехов в n испытаниях. Найдем $M\mu_n$ и $D\mu_n$.

Обозначим ξ_i — число успехов в i-м испытании, тогда нетрудно понять, что

$$\mu_n = \xi_1 + \xi_2 + ... + \xi_n. \tag{8.2}$$

Так как испытания Бернулли не зависят друг от друга, то случайные величины $\xi_1, \xi_2, ..., \xi_n$ являются независимыми. Распределение случайной величины ξ_i задается рядом распределения

ξ_i	1	0
P	p	q

.

Поэтому $M\xi_i = 1 \cdot p + 0 \cdot q = p$, $D\xi_i = M(\xi_i^2) - (M\xi_i)^2 = (1^2 \cdot p + 0^2 \cdot q) - p^2 = p - p^2 = p(1-p) = pq$. Из равенства (8.2) следует, что

$$M\mu_n = \sum_{i=1}^{n} M\xi_i = np.$$

Из равенства (8.2) и независимости $\xi_1, \xi_2, ..., \xi_n$ вытекает (см. теорему 8.3), что

$$D\mu_n = \sum_{i=1}^{n} D\xi_i = npq.$$

§ 8.2. ФУНКЦИИ ОТ СЛУЧАЙНЫХ ВЕЛИЧИН И ВЕКТОРОВ

Пусть ξ — случайная величина, а $f(x)$ — числовая функция. Тогда $\eta = f(\xi)$ является случайной величиной. Мы уже знаем, как находить математическое ожидание η (см. свойство 5 математического ожидания из лекции № 6). Ответим на вопрос, как находить распределение случайной величины η. Пусть $F_\eta(x)$ — функция распределения η, тогда

$$F_\eta(x) = P(\eta \leqslant x) = P(f(\xi) \leqslant x).$$

Рассмотрим при фиксированном x множество на числовой оси $B_x = \{u : f(u) \leqslant x\}$. Тогда случайные события $\{\omega : f(\xi) \leqslant x\}$ и $\{\omega : \xi \in B_x\}$ совпадают и поэтому

$$P(f(\xi) \leqslant x) = P(\xi \in B_x).$$

Если, например, ξ — непрерывная случайная величина с плотностью вероятностей $p(\cdot)$, то

$$P(\xi \in B_x) = \int\limits_{B_x} p(u)du$$

и, следовательно,

$$F_\eta(x) = \int\limits_{B_x} p(u)du.$$

В свою очередь, по функции распределения $F_\eta(x)$ можно найти плотность вероятностей $p_\eta(x)$ случайной величины η, используя равенство

$$p_\eta(x) = F'_\eta(x).$$

ПРИМЕР 8.4. Пусть $\eta \sim U[-1,1]$. Найти функцию распределения и плотность вероятностей случайной величины $\eta = \xi^3$.

Очевидно, ξ^3 принимает значения из отрезка $[-1,1]$. Возьмем $x \in [-1,1]$. Тогда

$$F_\eta(x) = P(\xi^3 \leqslant x) = P(\xi \leqslant \sqrt[3]{x}).$$

Плотность вероятностей случайной величины η имеет вид

$$p_\xi(x) = \begin{cases} 1/2, & x \in [-1,1]; \\ 0, & x \notin [-1,1]. \end{cases}$$

Поэтому

$$P(\xi \leqslant \sqrt[3]{x}) = \int\limits_{-\infty}^{\sqrt[3]{x}} p_\xi(u)du = \int\limits_{-1}^{\sqrt[3]{x}} \frac{1}{2}du = \frac{1}{2}\left(\sqrt[3]{x}+1\right).$$

Итак, при $x \in [-1,1]$

$$F_\eta = \frac{1}{2}\left(\sqrt[3]{x}+1\right).$$

Ясно, что при $x < -1$ $F_\eta(x) = 0$ и при $x > 1$ $F_\eta(x) = 1$. Найдем плотность вероятностей случайной величины η:

$$p_\eta(x) = F'_\eta(x) = \begin{cases} \dfrac{1}{6\sqrt[3]{x^2}}, & x \in (-1,1); \\ 0, & x \notin (-1,1). \end{cases}$$

ПРИМЕР 8.5. Пусть $\eta \sim U[-1,1]$. Найти функцию распределения и плотность вероятностей случайной величины $\eta = \xi^2$.

Очевидно, ξ^2 принимает значения из отрезка $[0,1]$. Возьмем $x \in [0,1]$. Тогда

$$F_\eta(x) = P(\xi^2 \leqslant x) = P(-\sqrt{x} \leqslant \xi \leqslant \sqrt{x}) =$$

$$= \int\limits_{-\sqrt{x}}^{\sqrt{x}} p_\xi(u)du = \int\limits_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{2}du = \sqrt{x}.$$

При $x < 0$ $F_\eta(x) = 0$, а при $x > 1$ $F_\eta(x) = 1$. Плотность вероятностей случайной величины η имеет вид:

$$p_\eta(x) = F'_\eta(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & x \in (0,1); \\ 0 & x \notin (0,1). \end{cases}$$

Пусть теперь $\bar{\xi} = \{\xi_1, \xi_2, ..., \xi_k\}$ — случайный вектор, а $f(x_1, x_2, ..., x_k)$ — числовая функция нескольких числовых переменных. Тогда $\eta = f(\xi_1, \xi_2, ..., \xi_k)$ — случайная величина. О том, как находить математическое ожидание η говорится в лекции № 7. Пусть $F_\eta(x)$ — функция распределения случайной величины η. Предположим, что $\bar{\xi}$ — непрерывный случайный вектор с плотностью вероятностей $p(x_1, x_2, ..., x_k)$. Рассмотрим при фиксированном x множество из R^k:

$$B_x = \{(x_1, x_2, ..., x_k) : \ f(x_1, x_2, ..., x_k) \leqslant x\}.$$

Тогда так же, как и в одномерном случае, получаем, что

$$F_\eta(x) = P(\eta \leqslant x) = P(f(\xi_1, \xi_2, ..., \xi_k) \leqslant x) =$$

$$= P((\xi_1, \xi_2, ..., \xi_k) \in B_x) = \int\limits_{B_x}\!\!...\!\int p(x_1, x_2, ..., x_k)dx_1dx_2...dx_k.$$

В качестве иллюстрации этого приема найдем распределение суммы двух независимых непрерывных случайных величин ξ_1 и ξ_2 с плотностями вероятностей $p_1(\cdot)$ и $p_2(\cdot)$. Оказывается, что случайная величина $\eta = \xi_1 + \xi_2$ является непрерывной и ее плотность вероятностей можно определить по формуле

$$p_\eta(x) = \int\limits_{-\infty}^{+\infty} p_1(x_1)p_2(x - x_1)dx_1$$

(как известно, последний интеграл называется сверткой функций $p_1(\cdot)$ и $p_2(\cdot)$). Действительно, в силу утверждения 8.1 плотность вероятностей случайного вектора $\{\xi_1, \xi_2\}$ равна $p(x_1, x_2) = p_1(x_1)p_2(x_2)$, поэтому для произвольного x

$$P(\eta \leqslant x) = P(\xi_1 + \xi_2 \leqslant x) = \iint\limits_{B_x} p_1(x_1)p_2(x_2)dx_1dx_2,$$

где $B_x = \{(x_1, x_2): \ x_1 + x_2 \leqslant x\}$ (см. рис. 8.3).

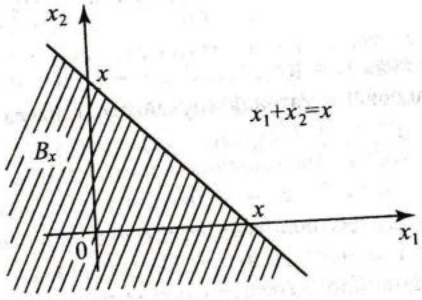


Рис. 8.3

Перейдем от двойного интеграла к повторному:

$$\iint\limits_{B_x} p_1(x_1)p_2(x_2)dx_1dx_2 = \int\limits_{-\infty}^{+\infty} p_1(x_1)dx_1 \int\limits_{-\infty}^{x-x_1} p_2(x_2)dx_2. \quad (8.3)$$

Делая замену $u = x_2 + x_1$ во внутреннем интеграле, получаем, что правая часть (8.3) имеет вид

$$\int\limits_{-\infty}^{+\infty} p_1(x_1)dx_1 \int\limits_{-\infty}^{x} p_2(u - x_1)du = \int\limits_{-\infty}^{+\infty} dx_1 \int\limits_{-\infty}^{x} p_1(x_1)p_2(u - x_1)du.$$

И, наконец, меняя порядок переменных в последнем интеграле, в итоге находим, что

$$P(\eta \leqslant x) = \int\limits_{-\infty}^{x} du \int\limits_{-\infty}^{+\infty} p_1(x_1)p_2(u - x_1)dx_1 = \int\limits_{-\infty}^{x} p_\eta(u)du,$$

ч. т. д.

Контрольные вопросы

1. Какой геометрический факт лежит в основе доказательства независимости ξ_1 и ξ_2 из примера 8.1?

2. Объясните, почему в случае, когда компоненты случайного вектора независимы, его распределение однозначно восстанавливается по маргинальным распределениям.

3. Следует ли из равенства $r(\xi, \eta) = 0$ независимость случайных величин ξ и η?

4. Докажите непосредственно соотношение $M\mu_n = np$ из примера 8.3.

5. Напишите формулу для дисперсии суммы двух произвольных случайных величин (см. (8.1)).

6. Ковариационная матрица случайного вектора $\{\xi_1, \xi_2\}$ имеет вид $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Найти $D(3\xi_1 - 5\xi_2)$.

Ответы

1. Формула для площади прямоугольника.

3. Не следует.

5. $D(\xi_1 + \xi_2) = D\xi_1 + 2\text{cov}(\xi_1, \xi_2) + D\xi_2$.

6. $D(3\xi_1 - 5\xi_2) = D(3\xi_1) + 2\text{cov}(3\xi_1, -5\xi_2) + D(-5\xi_2) = 9D\xi_1 - 30\text{cov}(\xi_1, \xi_2) + 25D\xi_2 = 9 \cdot 1 - 30 \cdot 2 + 25 \cdot 4 = 49$.

Лекция 9

ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

§ 9.1. ОПРЕДЕЛЕНИЕ ХАРАКТЕРИСТИЧЕСКОЙ ФУНКЦИИ И ЕЕ ВЫЧИСЛЕНИЕ

Рассмотрим произвольную случайную величину ξ и применим к ней функцию $f(x) = e^{itx}$, где t — действительный параметр. Известно, что

$$e^{itx} = \cos tx + i \sin tx,$$

таким образом, $f(x)$ является комплекснозначной функцией. Рассмотрение именно такой функции вызвано, с одной стороны, тем, что $|f(x)| = 1$, а с другой — тем, что $f(x)$ является показательной функцией. Применяя $f(x)$ к случайной величине ξ, получаем случайную величину $\eta = f(\xi)$. Математическое ожидание η будет зависеть от t. Оказывается, совокупность всех таких математических ожиданий несет исчерпывающую информацию о случайной величине ξ, причем представленную в очень удобном виде.

Определение 9.1. Характеристической функцией случайной величины ξ называется следующая комплекснозначная функция действительного переменного t:

$$\varphi(t) = Me^{it\xi} = M\cos t\xi + iM \sin t\xi.$$

По свойству 5 математического ожидания (см. лекцию № 6)

$$\varphi(t) = \sum_{k=1}^{\infty} e^{ita_k} p_k,$$

если ξ является дискретной случайной величиной с рядом распределения

ξ	a_1	a_2	...	a_k	...
P	p_1	p_2	...	p_k	...

и

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} p(x)dx,$$

если ξ является непрерывной случайной величиной с плотностью вероятностей $p(x)$.

Замечание 9.1. Из курса математического анализа известно, что в непрерывном случае характеристическая функция является преобразованием Фурье плотности вероятностей.

ПРИМЕР 9.1. Пусть случайная величина ξ имеет распределение Пуассона с параметром a. Найти характеристическую функцию ξ.

Случайная величина ξ принимает значения $0, 1, 2, ...$, причем

$$P(\xi = k) = \frac{a^k}{k!} e^{-a}, k = 0, 1, 2, ...$$

По определению характеристической функции

$$\varphi(t) = \sum_{k=0}^{\infty} e^{itk} P(\xi = k) = \sum_{k=0}^{\infty} e^{itk} \frac{a^k}{k!} e^{-a} = e^{-a} \sum_{k=0}^{\infty} \frac{(e^{it}a)^k}{k!}.$$

Известно, что при всех комплексных z

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!},$$

поэтому

$$\sum_{k=0}^{\infty} \frac{(e^{it}a)^k}{k!} = e^{ae^{it}}.$$

Итак,

$$\varphi(t) = e^{-a} e^{ae^{it}} = e^{-a(1-e^{it})}.$$

ПРИМЕР 9.2. Пусть $\xi \sim N(a; \sigma^2)$. Найти характеристическую функцию ξ.

а) Рассмотрим сначала случай, когда ξ имеет *стандартное* нормальное распределение, т. е. $\xi \sim N(0;1)$. Тогда плотность вероятностей ξ имеет вид

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, x \in R.$$

По определению характеристической функции

$$\varphi(t) = \int\limits_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{itx - x^2/2} dx.$$

Поскольку

$$itx - \frac{x^2}{2} = -\frac{x^2 - 2itx}{2} = -\frac{(x-it)^2 + t^2}{2},$$

то делая замену $x - it = u$, получаем, что

$$\varphi(t) = e^{-t^2/2} \int\limits_{-\infty - it}^{+\infty + it} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du.$$

Последний интеграл равен 1 при $t = 0$, поскольку он представляет собою интеграл от плотности вероятностей по всей числовой оси. И при всех действительных t он равен нулю, что можно показать методами теории функций комплексного переменного. Итак,

$$\varphi(t) = e^{-t^2/2}.$$

б) Рассмотрим теперь общий случай: $\xi \sim N(a;\sigma^2)$. Тогда плотность вероятностей ξ имеет вид

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2/(2\sigma^2)}, x \in R,$$

и характеристическая функция определяется соотношением

$$\varphi(t) = \int\limits_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2/(2\sigma^2)} dx.$$

Осуществляя замену переменных $(x-a)/\sigma = u$, получаем, что

$$\varphi(t) = \int\limits_{-\infty}^{+\infty} e^{it(\sigma u + a)} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = e^{ita} \int\limits_{-\infty}^{+\infty} e^{i(t\sigma)u} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du,$$

но последний интеграл есть характеристическая функция стандартного нормального распределения, взятая в точке $t\sigma$, и поэтому этот интеграл равен $e^{-(t\sigma)^2/2}$. Следовательно,

$$\varphi(t) = e^{ita - (t^2\sigma^2)/2}.$$

§ 9.2. СВОЙСТВА ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ

Свойство 1. Характеристическая функция определена при всех $t \in R$, непрерывна на всей числовой оси и

$$\lim_{t \to +\infty} \varphi(t) = \lim_{t \to -\infty} \varphi(t) = 0. \tag{9.1}$$

Доказательство (в непрерывном случае). Характеристическая функция есть несобственный интеграл $\int_{-\infty}^{+\infty} e^{itx} p(x) dx$. Он сходится, если сходится интеграл от модуля подынтегральной функции, но

$$\left| e^{itx} p(x) \right| = \left| e^{itx} \right| p(x) = 1 p(x),$$

а интеграл от $p(x)$ по всей числовой оси равен 1.

Непрерывность $\varphi(t)$ и соотношение (9.1) следуют из свойств преобразования Фурье.

Свойство 2. Характеристическая функция при $t = 0$ равна 1.

Доказательство. Очевидно, что

$$\varphi(0) = Me^{i0\xi} = Me^0 = M1 = 1.$$

Свойство 3. Если существует $M\left(|\xi|^k\right)$ при некотором натуральном k, то существует непрерывная производная k-го порядка $\varphi^{(k)}(t)$ характеристической функции $\varphi(t)$ и

$$M\left(\xi^k\right) = i^{-k} \varphi^{(k)}(0).$$

Доказательство (в непрерывном случае). Сходимость интеграла $\int_{-\infty}^{+\infty}|x|^k p(x)dx$ позволяет при дифференцировании функции

$$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}p(x)dx$$

осуществлять это дифференцирование под знаком интеграла:

$$\varphi^{(k)}(t)=\int_{-\infty}^{+\infty}(ix)^k e^{itx}p(x)dx.$$

Подставляя $t=0$, получаем, что

$$\varphi^{(k)}(0)=i^k\int_{-\infty}^{+\infty}x^k p(x)dx=i^k M(\xi^k),$$

откуда вытекает требуемое равенство.

Свойство 4. Пусть $\varphi(t)$ — характеристическая функция случайной величины ξ. Тогда характеристическая функция случайной величины $a\xi+b$, где a,b — постоянные, равна $e^{itb}\varphi(at)$.

Доказательство. Действительно,

$$Me^{it(a\xi+b)}=M\left(e^{ita\xi}e^{itb}\right)=e^{itb}Me^{i(ta)\xi}=e^{itb}\varphi(ta).$$

Свойство 5. Если $\xi_1,\xi_2,...,\xi_k$ — независимые случайные величины с характеристическими функциями $\varphi_1(t),\varphi_2(t),...,$ $\varphi_k(t)$, то характеристическая функция суммы $\xi_1+\xi_2+...+\xi_k$ равна $\varphi_1(t)\varphi_2(t)...\varphi_k(t)$.

Доказательство. Нетрудно доказать, что если $\xi_1,\xi_2,...,$ ξ_k — независимые случайные величины и $f(x)$ — произвольная числовая функция (в том числе, комплекснозначная), то $f(\xi_1),f(\xi_2),...,f(\xi_k)$ также являются независимыми случайными величинами. Откуда, учитывая теорему 8.1, получаем при $f(x)=e^{itx}$, что

$$Me^{it(\xi_1+\xi_2+...+\xi_k)}=Me^{it\xi_1}e^{it\xi_2}...e^{it\xi_k}=$$
$$=Me^{it\xi_1}Me^{it\xi_2}...Me^{it\xi_k}=\varphi_1(t)\varphi_2(t)...\varphi_k(t).$$

Свойство 6. По характеристической функции $\varphi(t)$ функция распределения случайной величины восстанавливается однозначно.

ПРИМЕР 9.3. Пусть случайная величина ξ имеет распределение Пуассона с параметром a. По характеристической функции найти $M\xi$ и $D\xi$.

В примере 9.1 установлено, что характеристическая функция ξ имеет вид

$$\varphi(t)=e^{-a(1-e^{it})}.$$

По свойству 3

$$M\xi=i^{-1}\varphi'(0),M(\xi^2)=-\varphi''(0). \tag{9.2}$$

Найдем первые две производные $\varphi(t)$:

$$\varphi'(t)=iae^{it}e^{-a(1-e^{it})},\varphi''(t)=-ae^{it}e^{-a(1-e^{it})}-a^2e^{2it}e^{-a(1-e^{it})}.$$

Подставляя $t=0$, получаем, что

$$\varphi'(0)=ia,\varphi''(0)=-a-a^2.$$

Следовательно,

$$M\xi=a,M(\xi^2)=a+a^2\Rightarrow D\xi=M(\xi^2)-(M\xi)^2=a.$$

ПРИМЕР 9.4. Пусть $\xi\sim N(a;\sigma^2)$. По характеристической функции найти $M\xi$ и $D\xi$.

В примере 9.2 найдена характеристическая функция ξ:

$$\varphi(t)=e^{ita-(t^2\sigma^2)/2}.$$

Найдем $\varphi'(t)$ и $\varphi''(t)$:

$$\varphi'(t)=(ia-t\sigma^2)e^{ita-(t^2\sigma^2)/2},\varphi''(t)=$$
$$=-\sigma^2e^{ita-(t^2\sigma^2)/2}+(ia-t\sigma^2)^2e^{ita-(t^2\sigma^2)/2}.$$

Подставляя $t=0$, получаем, что

$$\varphi'(0)=ia,\varphi''(0)=-\sigma^2-a^2.$$

Из равенств (9.2) находим, что

$$M\xi = a, M(\xi^2) = \sigma^2 + a^2 \Rightarrow D\xi = M(\xi^2) - (M\xi)^2 = \sigma^2.$$

Итак, параметры a и σ^2 нормального распределения имеют простой вероятностный смысл: a является математическим ожиданием, а σ^2 — дисперсией случайной величины, имеющей такое распределение.

ПРИМЕР 9.5. Доказать, что если $\xi \sim N(a;\sigma^2)$, то случайная величина $\eta = c\xi + d$, где c, d — постоянные, имеет нормальное распределение $N(ca + d; c^2\sigma^2)$.

Характеристическая функция ξ имеет вид

$$\varphi_\xi(t) = e^{ita-(t^2\sigma^2)/2}.$$

Тогда по свойству 4 характеристическая функция η имеет вид

$$\varphi_\eta(t) = e^{itd}\varphi_\xi(ct) = e^{itd}e^{itca-(t^2(c\sigma)^2)/2} = e^{i(ca+d)t-(t^2(c\sigma)^2)/2}.$$

Последнее выражение является характеристической функцией нормального распределения $N(ca + d; c^2\sigma^2)$. Поскольку по свойству 6 функция распределения восстанавливается однозначно по характеристической функции, то $\eta \sim N(ca + d; c^2\sigma^2)$.

ПРИМЕР 9.6. Пусть $\xi_1, \xi_2, ..., \xi_k$ — независимые случайные величины, причем $\xi_m \sim N(a_m; \sigma_m^2), m = 1, 2, ..., k$. Тогда

$$\xi_1 + \xi_2 + ... + \xi_k \sim N\left(\sum_{m=1}^{k} a_m; \sum_{m=1}^{k} \sigma_m^2\right).$$

Действительно, характеристическая функция случайной величины ξ_m имеет вид

$$\varphi_m(t) = e^{ita_m-(t^2\sigma_m^2)/2}.$$

По свойству 5 характеристическая функция случайной величины $\xi_1 + \xi_2 + ... + \xi_k$ выражается в виде

$$\prod_{m=1}^{k} \varphi_m(t) = \prod_{m=1}^{k} e^{ita_m-(t^2\sigma_m^2)/2} = e^{it\sum_{m=1}^{k} a_m - t^2/2\sum_{m=1}^{k}\sigma_m^2}.$$

Последнее выражение является характеристической функцией нормального распределения $N(\sum_{m=1}^{k} a_m; \sum_{m=1}^{k} \sigma_m^2)$. Ввиду того, что функция распределения восстанавливается однозначно по характеристической функции, получаем требуемое утверждение.

§ 9.3. СХОДИМОСТЬ СЛУЧАЙНЫХ ВЕЛИЧИН

Важные результаты теории вероятностей — закон больших чисел, центральная предельная теорема и т. п. связаны с понятием сходимости последовательности случайных величин.

Определение 9.2. Пусть $\xi_1, \xi_2, ...$ — последовательность случайных величин (заданных, вообще говоря, на разных вероятностных пространствах). Говорят, что эта последовательность *сходится по распределению* к случайной величине ξ, если

$$\lim_{n\to\infty} P(\xi_n \leqslant x) = P(\xi \leqslant x),$$

т. е.

$$\lim_{n\to\infty} F_n(x) = F(x),$$

где $F_n(x)$ — функция распределения ξ_n, а $F(x)$ — функция распределения ξ. Предполагается, что соотношение (9.3) выполняется для всех тех x, при которых функция $F(x)$ непрерывна. Кратко сходимость по распределению записывают в следующем виде:

$$\xi_n \xrightarrow{D} \xi.$$

Справедлив следующий замечательный результат, подчеркивающий важность понятия характеристической функции.

Теорема 9.1. *Последовательность случайных величин $\xi_1, \xi_2, ...$ сходится по распределению к случайной величине ξ тогда и только тогда, когда последовательность характеристических функций $\varphi_1(t), \varphi_2(t), ...$ этих случайных величин сходится к характеристической функции $\varphi(t)$ случайной величины ξ при любом t, т. е. $\lim_{n\to\infty}\varphi_n(t) = \varphi(t), t \in R$.*

Рассмотрим теперь более сильное понятие сходимости случайных величин.

Определение 9.3. Пусть $\xi_1, \xi_2, ...$ — последовательность случайных величин, заданных на одном и том же вероятностном пространстве. Говорят, что эта последовательность *сходится почти наверное* или с *вероятностью* 1 к случайной величине ξ, если

$$\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)$$

для всех $\omega \in \Omega$, за исключением тех ω, которые составляют множество нулевой вероятности.

Примеры теорем, касающихся сходимости последовательности случайных величин по распределению или почти наверное будут приведены в следующей лекции. Теоремы такого типа получили название предельных и составляют основное содержание современной теории вероятностей. Теоремы о сходимости почти наверное требуют более сложного математического аппарата и здесь доказываться не будут. Существенную роль в разработке методов их доказательств сыграл А. Н. Колмогоров.

Контрольные вопросы

1. При определении характеристической функции использовалась показательная функция $f(x) = e^{itx}$, где t — действительный параметр. Как вы думаете, для каких случайных величин удобно использовать показательную функцию $g(x) = e^{-\lambda x}$, где λ — положительный параметр? Вспомните, с каким известным преобразованием в математическом анализе связана эта показательная функция?

2. Если $f(x)$ — показательная функция, то $f(x_1 + x_2) = f(x_1)f(x_2)$. При доказательстве какого свойства характеристических функций было использовано это соотношение?

3. Приведите пример случайной величины, характеристическая функция которой есть e^{ita}, где a — действительная постояннная.

4. Докажите, что для произвольной случайной величины ξ с характеристической функций $\varphi(t)$

$$D\xi = -\varphi''(0) + (\varphi'(0))^2,$$

если $M(\xi^2)$ существует.

5. Пусть $\xi \sim N(a, \sigma^2)$. Докажите, что

$$\frac{\xi - a}{\sigma} \sim N(0; 1).$$

6. Пусть независимые случайные величины $\xi_1, \xi_2, ..., \xi_n$ имеют одинаковое распределение $N(a; \sigma^2)$. Докажите, что

$$\frac{\xi_1, \xi_2, ..., \xi_n - na}{\sigma\sqrt{n}} \sim N(0; 1).$$

Ответы

1. Для положительных случайных величин.
2. Свойство 5.
3. $\xi \equiv a$.
4. $D\xi = M(\xi^2) - (M\xi)^2 = -\varphi''(0) - \left(\frac{\varphi'(0)}{i}\right)^2 = -\varphi''(0) + (\varphi'(0))^2.$

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

§ 10.1. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Рассмотрим последовательность независимых одинаково распределенных случайных величин $\xi_1, \xi_2, \ldots$. Это означает, что при любом натуральном $n \geqslant 2$ случайные величины $\xi_1, \xi_2, \ldots \xi_n$ независимы и все случайные величины $\xi_1, \xi_2, \ldots$ имеют одинаковую функцию распределения. Отсюда, кстати, следует, что математические ожидания этих случайных величин совпадают (если они существуют). То же самое верно относительно дисперсий. Во многих задачах теории вероятностей и математической статистики важно знать распределение суммы

$$S_n = \xi_1 + \xi_2 + \ldots + \xi_n$$

при довольно больших значениях n. Именно на этот вопрос отвечают рассматриваемые предельные теоремы.

Теорема 10.1 (Хинчин). *Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины с математическим ожиданием a. Тогда при $n \to \infty$*

$$\frac{S_n}{n} \xrightarrow{D} a. \tag{10.1}$$

Доказательство. Пусть $\varphi(t)$ — характеристическая функция ξ_1 (а следовательно, и $\xi_2, \xi_3, \ldots$). По свойствам 2 и 3 характеристических функций (см. лекцию № 9)

$$\varphi(0) = 1, \quad \varphi'(0) = iM\xi_1 = ia,$$

причем $\varphi'(t)$ непрерывна на всей числовой оси. Следовательно, по формуле Тейлора при всех $t \in R$

$$\varphi(t) = \varphi(0) + \varphi'(0)t + h(t)t = 1 + iat + h(t)t,$$

где $\lim\limits_{t\to 0} h(t) = 0$. По свойству 5 характеристическая функция S_n равна $\varphi^n(t)$, а по свойству 4 характеристическая функция S_n/n равна

$$\varphi^n\left(\frac{t}{n}\right) = \left(1 + \frac{iat}{n} + h\left(\frac{t}{n}\right)\frac{t}{n}\right)^n,$$

причем $\lim\limits_{n\to\infty} h(t/n) = 0$. Вспоминая, как находятся пределы «типа e», получаем, что

$$\lim_{n\to\infty} \varphi^n\left(\frac{t}{n}\right) = \lim_{n\to\infty}\left(1 + \frac{iat}{n} + h\left(\frac{t}{n}\right)\frac{t}{n}\right)^n =$$

$$= e^{\lim\limits_{n\to\infty}((iat)/n + h(t/n)(t/n))n} = e^{iat}$$

Но e^{iat} является характеристической функцией случайной величины $\xi \equiv a$. Итак, характеристическая функция случайной величины S_n/n сходится при $n \to \infty$ к характеристической функции постоянной a. По теореме 9.1 отсюда следует утверждение теоремы 10.1.

Следствие 10.1. Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины с математическим ожиданием a, тогда при $n \to \infty$ для любого положительного числа ε

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - a\right| > \varepsilon\right) = 0. \qquad (10.2)$$

Доказательство. Соотношение (10.2) равносильно следующему утверждению

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - a\right| \leqslant \varepsilon\right) = 1. \qquad (10.3)$$

Заметим, что

$$P\left(\left|\frac{S_n}{n} - a\right| \leqslant \varepsilon\right) = P\left(a - \varepsilon \leqslant \frac{S_n}{n} \leqslant a + \varepsilon\right) \geqslant$$

$$\geqslant P\left(a - \varepsilon < \frac{S_n}{n} \leqslant a + \varepsilon\right) =$$

$$= P\left(\frac{S_n}{n} \leqslant a + \varepsilon\right) - P\left(\frac{S_n}{n} \leqslant a - \varepsilon\right).$$

Итак,

$$1 \geqslant P\left(\left|\frac{S_n}{n} - a\right| \leqslant \varepsilon\right) \geqslant P\left(\frac{S_n}{n} \leqslant a + \varepsilon\right) - P\left(\frac{S_n}{n} \leqslant a - \varepsilon\right). \qquad (10.4)$$

Функция распределения $F_\xi(x)$ случайной величины $\xi \equiv a$ имеет вид

$$F_\xi(x) = \begin{cases} 1, & x \geqslant a; \\ 0, & x < a. \end{cases}$$

Точки $(a + \varepsilon)$ и $(a - \varepsilon)$ являются точками непрерывности $F_\xi(x)$, поэтому по теореме 10.1

$$\lim_{n\to\infty} P\left(\frac{S_n}{n} \leqslant a + \varepsilon\right) = 1, \quad \lim_{n\to\infty} P\left(\frac{S_n}{n} \leqslant a - \varepsilon\right) = 0.$$

Значит, предел крайних членов неравенства (10.4) равен 1. Таким образом, и промежуточный член (10.4) будет иметь предел 1, что доказывает (10.3) и следствие.

Утверждения типа (10.1) и (10.2) получили название *закона больших чисел* (ЗБЧ). Фактически они свидетельствуют о том, что при больших n среднее значение $(\xi_1 + \xi_2 + \ldots + \xi_n)/n$ независимых и одинаково распределенных случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ мало отличается от постоянной, равной их математическому ожиданию.

Следствие 10.2. Пусть проводятся испытания Бернулли с вероятностью успеха в одном испытании p, и μ_n — число успехов в n испытаниях, тогда для произвольного положительного числа ε

$$\lim_{n\to\infty} P\left(\left|\frac{\mu_n}{n} - p\right| > \varepsilon\right) = 0. \qquad (10.5)$$

Доказательство. Вспомним (см. пример 8.3 лекции № 8), что

$$\mu_n = \xi_1 + \xi_2 + \dots + \xi_n,$$

где ξ_i означает число успехов в i-м испытании, $i = 1, 2, \dots$, причем $\xi_1, \xi_2, \dots$ — независимые и одинаково распределенные случайные величины с математическим ожиданием p. Отсюда ввиду следствия 10.1 получаем требуемое утверждение.

Замечание 10.1. Соотношение (10.5) является исторически первой формой закона больших чисел и принадлежит Я. Бернулли. Соотношение (10.2) с более обременительными условиями впервые было доказано П. Л. Чебышевым.

Наиболее сильный результат для последовательности независимых одинаково распределенных случайных величин принадлежит А. Н. Колмогорову и носит название *усиленного закона больших чисел*. Приведем его формулировку.

Теорема 10.2 (Колмогоров). *Пусть $\xi_1, \xi_2, \dots$ — независимые одинаково распределенные случайные величины с математическим ожиданием a, тогда почти наверное*

$$\lim_{n\to\infty} \frac{S_n}{n} = a.$$

Следствие 10.3 (Борель). Пусть проводятся испытания Бернулли с вероятностью успеха в одном испытании p, и μ_n — число успехов в n испытаниях. Тогда почти наверное

$$\lim_{n\to\infty} \frac{\mu_n}{n} = p.$$

Замечание 10.2. Результат Бореля легко следует из теоремы Колмогорова (см. доказательство следствия 10.2), но исторически он предшествовал ей.

Усиленный закон больших чисел имеет глубокий смысл.

Во-первых, он объясняет, почему математическое ожидание случайной величины ξ характеризует ее среднее значение после большого числа экспериментов. Действительно, когда мы говорим, что случайный эксперимент, в котором наблюдается случайная величина ξ, воспроизведен n раз, то имеем в виду, что в результате получено n независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$, распределенных как ξ и заданных

при одном и том же элементарном исходе ω. Поэтому среднее значение ξ после n экспериментов есть $(\xi_1(\omega) + \xi_2(\omega) + \dots + \xi_n(\omega))/n$. Но по усиленному закону больших чисел предел этой величины при $n \to \infty$ с вероятностью 1 равен $M\xi$.

Во-вторых, усиленный закон больших чисел позволяет установить свойство устойчивости частот. Пусть осуществляется наблюдение за некоторым случайным событием A, относящимся к некоторому случайному эксперименту. Наступление события A будем считать успехом, и вероятность успеха равна $p = P(A)$. Если случайный эксперимент воспроизведен n раз, то это означает, что проведено n независимых испытаний (экспериментов), т. е. получаем схему Бернулли. Частота $\nu_n(A)$ случайного события A равна отношению числа испытаний, в которых произошло событие A, к n, т. е. отношению числа успехов μ_n к n. Поэтому по следствию 10.3 с вероятностью 1

$$\lim_{n\to\infty} \nu_n(A) = \lim_{n\to\infty} \frac{\mu_n}{n} = p = P(A),$$

что и требовалось доказать.

§ 10.2. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

По закону больших чисел $S_n/n \approx a$. Каков порядок разности $S_n - na$? Центральная предельная теорема (ЦПТ), представленная ниже, показывает, что он равен $\sqrt{n}$.

Теорема 10.3 (Линдеберг–Леви). *Пусть $\xi_1, \xi_2, \dots$ — независимые одинаково распределенные случайные величины с математическим ожиданием a и дисперсией σ^2. Тогда при $n \to \infty$*

$$\frac{S_n - na}{\sigma\sqrt{n}} \xrightarrow{D} N(0;1).$$

Замечание 10.3. Здесь значок $N(0;1)$ использован для обозначения случайной величины со стандартным нормальным распределением.

Доказательство. Пусть $\varphi(t)$ — характеристическая функция случайной величины $\xi_1 - a$ (а следовательно, $\xi_2 - a$, $\xi_3 - a, \ldots$). По свойствам характеристической функции

$$\varphi(0) = 1, \quad \varphi'(0) = i^{-1}M(\xi_1 - a) = i^{-1}(a - a) = 0,$$

$$\varphi''(0) = -M(\xi_1 - a)^2 = -\sigma^2,$$

причем $\varphi''(t)$ непрерывна на всей числовой оси. Следовательно, по формуле Тейлора

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + h(t)t^2 = 1 - \frac{\sigma^2}{2}t^2 + h(t)t^2,$$

где $\lim\limits_{t\to 0} h(t) = 0$. По свойству 5 характеристических функций (см. лекцию № 9) случайная величина $S_n - na = (\xi_1 - a) + (\xi_2 - a) + \ldots + (\xi_n - a)$, где слагаемые $\xi_1 - a, \xi_2 - a, \ldots, \xi_n - a$ — независимые случайные величины, имеет характеристическую функцию $\varphi^n(t)$. Следовательно, по свойству 4 характеристическая функция случайной величины $\dfrac{S_n - na}{\sigma\sqrt{n}}$ задается выражением

$$\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(1 - \frac{\sigma^2}{2}\cdot\frac{t^2}{\sigma^2 n} + h\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{\sigma^2 n}\right)^n,$$

причем $\lim\limits_{n\to\infty} h\left(\dfrac{t}{\sigma\sqrt{n}}\right) = 0$. Воспользовавшись приемом нахождения пределов «типа e», получаем, что

$$\lim_{n\to\infty}\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right) = \lim_{n\to\infty}\left(1 - \frac{t^2}{2n} + h\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{\sigma^2 n}\right)^n =$$

$$= e^{\lim\limits_{n\to\infty}\left(-t^2/(2n) + h\left(t/(\sigma\sqrt{n})\right)t^2/(\sigma^2 n)\right)n} = e^{-t^2/2}.$$

Вспомним (см. пример 9.2 лекции № 9), что $e^{-t^2/2}$ — характеристическая функция стандартного нормального распределения. Итак, характеристическая функция случайной величины $(S_n - na)/(\sigma\sqrt{n})$ сходится при $n\to\infty$ к характеристической функции стандартного нормального распределения. По теореме 9.1 отсюда следует утверждение теоремы 10.3.

Обозначим $\Phi(x)$ функцию распределения случайной величины, имеющей стандартное нормальное распределение, т. е.

$$\Phi(x) = \int\limits_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}du.$$

Эта функция называется *функцией Лапласа*. Из математического анализа известно, что функция Лапласа не выражается через элементарные функции, поэтому для нее составлены таблицы. Нетрудно понять, что $\Phi(x) + \Phi(-x) = 1$, поэтому таблицы составлены лишь для неотрицательных значений x. Если вспомнить определение сходимости по распределению, то ЦПТ означает, что для произвольного x

$$\lim_{n\to\infty} P\left(\frac{S_n - na}{\sigma\sqrt{n}} \leqslant x\right) = \Phi(x). \tag{10.6}$$

Отсюда нетрудно видеть, что

$$P(S_n \in [x_1, x_2]) = P\left(\frac{S_n - na}{\sigma\sqrt{n}} \in \left[\frac{x_1 - na}{\sigma\sqrt{n}}, \frac{x_2 - na}{\sigma\sqrt{n}}\right]\right) \approx$$

$$\approx \Phi\left(\frac{x_2 - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{x_1 - na}{\sigma\sqrt{n}}\right), \tag{10.7}$$

если n велико и $(x_2 - na)$ и $(x_1 - na)$ имеют порядок $\sqrt{n}$.

Часто под функцией Лапласа понимают функцию

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-u^2/2}du.$$

Эта функция нечетна, т. е. $\Phi_0(-x) = -\Phi_0(x)$. Поэтому таблицы для нее составлены только для неотрицательных значений x. График этой функции при $x \geqslant 0$ приведен на рис. 10.1.

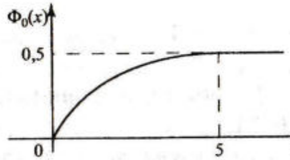


Рис. 10.1

При значениях $x \geqslant 5$ значение $\Phi_0(x)$ практически равно 0,5. Поэтому таблицы для $\Phi_0(x)$ составлены только для $x \in [0,5]$ (см. приложение, табл. 2). Поскольку $\Phi(x) = 0,5 + \Phi_0(x)$, из соотношения (10.7) следует, что

$$P(S_n \in [x_1, x_2]) \approx \Phi_0\left(\frac{x_2 - na}{\sigma\sqrt{n}}\right) - \Phi_0\left(\frac{x_1 - na}{\sigma\sqrt{n}}\right). \qquad (10.8)$$

ПРИМЕР 10.1. Стрелок попадает в «десятку» с вероятностью 0,4, в «девятку» с вероятностью 0,3, в «восьмерку» с вероятностью 0,2 и в «семерку» с вероятностью 0,1. Он произвел 25 выстрелов. Найти вероятность того, что суммарное число выбитых очков находится в пределах от 220 до 230.

Пусть ξ_i — число очков, выбитых при i-м выстреле. Ясно, что $\xi_1, \xi_2, \ldots$ — независимые случайные величины, причем с одинаковым распределением

ξ_i	7	8	9	10
P	0,1	0,2	0,3	0,4

Найдем числовые характеристики случайной величины ξ_i:

$$M\xi_i = 7 \cdot 0,1 + 8 \cdot 0,2 + 9 \cdot 0,3 + 10 \cdot 0,4 = 9,$$

$$D\xi_i = (7-9)^2 \cdot 0,1 + (8-9)^2 \cdot 0,2 + (9-9)^2 \cdot 0,3 + (10-9)^2 \cdot 0,4 = 1.$$

Суммарное число очков после 25 выстрелов выражается в виде

$$S_{25} = \xi_1 + \xi_2 + \ldots + \xi_{25}.$$

Ввиду соотношения (10.8)

$$P(S_{25} \in [220, 230]) \approx \Phi_0\left(\frac{230 - 25 \cdot 9}{1\sqrt{25}}\right) - \Phi_0\left(\frac{220 - 25 \cdot 9}{1\sqrt{25}}\right) =$$

$$= \Phi_0(1) - \Phi_0(-1) = 2\Phi_0(1).$$

Из таблиц (см. приложение, табл. 2) находим, что $\Phi_0(1) = 0,3413$.

Следовательно, искомая вероятность приблизительно равна 0,6826.

В качестве следствия центральной предельной теоремы получим результат, называемый теоремой Муавра-Лапласа.

Теорема 10.4. *Пусть проводятся испытания Бернулли с вероятностью успеха в одном испытании p, и μ_n — число успехов в n испытаниях. Тогда для произвольного x*

$$\lim_{n \to \infty} P\left(\frac{\mu_n - np}{\sqrt{npq}} \leqslant x\right) = \Phi(x).$$

Доказательство. Пусть ξ_i — число успехов в i-м испытании Бернулли. Случайные величины $\xi_1, \xi_2, \ldots$ являются независимыми и одинаково распределенными с математическим ожиданием $a = p$ и дисперсией $\sigma^2 = pq$ (см. пример 8.3). Далее,

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n.$$

Поэтому по центральной предельной теореме при $n \to \infty$

$$\frac{\mu_n - np}{\sqrt{npq}} \xrightarrow{D} N(0;1),$$

откуда (сравните с (10.6)) вытекает требуемое утверждение.

Контрольные вопросы

1. Объясните, почему соотношение (10.2) эквивалентно соотношению (10.3).
2. Объясните, почему именно усиленный ЗБЧ устанавливает согласованность аксиоматического и статистического определений понятия вероятности.
3. Выведите усиленный закон больших чисел из свойства устойчивости частот в случае, когда ξ_i $(i = 1, 2, \ldots)$ принимает конечное число значений.
4. Выведите закон больших чисел (следствие 10.1) из центральной предельной теоремы (теорема 10.3).
5. В условиях теоремы 10.3 докажите, что $\dfrac{S_{2n} - S_n - na}{\sigma\sqrt{n}} \xrightarrow{D} N(0;1)$ при $n \to \infty$.

Ответы

3. Вывод приведен в начале лекции № 6.
5. Требуемое утверждение следует из того, что случайная величина $S_{2n} - S_n$ имеет такое же распределение, как и S_n.

Занятие 4

ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ (ПРОДОЛЖЕНИЕ)

ПРИМЕР 4.1. У телефонного номера забыта последняя цифра и набирается наугад. Какова вероятность того, что придется звонить более четырех раз?

◁ Обозначим через A_i набор неправильной последней цифры при i-й попытке. Звонить придется более четырех раз (событие A), если при первой и при второй, и при третьей, и при четвертой попытках будет набраны неправильные цифры, т. е. $A = A_1 \cdot A_2 \cdot A_3 \cdot A_4$. События зависимы. Поэтому

$$P(A) = P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) =$$

$$= P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cdot A_2) \cdot P(A_4/A_1 \cdot A_2 \cdot A_3) =$$

$$= (9/10) \cdot (8/9) \cdot (7/8) \cdot (6/7) = 3/5. \; \triangleright$$

ПРИМЕР 4.2. В бригаде из шести человек было решено каждый день по жребию выбирать двоих для выполнения некоторой работы. Какова вероятность того, что в течение трех дней жребий падет по одному разу на каждого члена бригады? Какова вероятность для заданного члена бригады за три дня ни разу не вытянуть жребий? Какова вероятность того, что на данного члена бригады трижды падет жребий? Какова вероятность для данного члена бригады хотя бы один раз за три дня вытянуть жребий?

◁ Обозначим через A_i событие, состоящее в выборе в i-й день членов бригады, которые до этого еще ни разу не были выбраны.

Жребий падет по одному разу на каждого члена бригады (событие A), если в первый и во второй, и в третий день будут выбраны новые члены бригады, т. е. $A_1 \cdot A_2 \cdot A_3$. События зависимы. Поэтому $P(A) = P(A_1) \cdot P(A_2/A_1) \cdot (P(A_3/A_1 \cdot A_2)$. Каждый день производится выборка из шести по два бесповторным способом. Число возможных исходов жребия равно $C_6^2 = 15$. В первый день все эти исходы являются благоприятствующими, во второй день благоприятствующих исходов будет $C_4^2 = 6$, в третий день — $C_2^2 = 1$. Поэтому $P(A_1) = 1$, $P(A_2/A_1) = 6/15 = 2/5$, $P(A_3/A_1 \cdot A_2) = 1/15$ и $P(A) = 2/75$.

Обозначим через B_i событие, состоящее в том, что данный член бригады не был выбран по жребию в i-й день. Тогда данный член бригады не будет выбран ни разу (событие B), если произойдут событие B_1 и B_2, и B_3. Следовательно, $B = B_1 \cdot B_2 \cdot B_3$, причем события независимы, так как выбор каждый день производится в неизменных для данного члена бригады условиях. Возможных способов выбора каждый день $C_6^2 = 15$, а благоприятствуют те, при которых выбор производится из числа пяти остальных членов бригады. Число таких способов равно $C_5^2 = 10$. Поэтому

$$P(B) = P(B_1) \cdot P(B_2) \cdot P(B_3) = (10/15)^3 = (2/3)^3 = 8/27.$$

Обозначим через C_i событие, состоящее в том, что в i-й день на данного члена бригады пал жребий. На данного члена бригады жребий падет трижды (событие C), если произойдут события C_1 и C_2, и C_3. События C_i независимы, так как жребий каждый раз производится в неизменных по отношению к данному члену бригады условиях. Исходы жребия благоприятствуют событию C_i, если в пару к данному члену бригады выбрать одного из остальных пяти рабочих. Поэтому

$$P(C) = P(C_1) \cdot P(C_2) \cdot P(C_3) = (5/15)^3 = (1/3)^3 = 1/27.$$

Обозначим через D событие, состоящее в том, что на данного члена бригады жребий падет хотя бы один раз. Заметим, что это событие противоположно событию B. Поэтому

$$P(D) = 1 - P(B) = 1 - 8/27 = 19/27. \; \triangleright$$

ПРИМЕР 4.3. Два стрелка по очереди стреляют в цель до первого попадания. Вероятность попадания в цель при одном выстреле равна для них соответственно 1/3 и 1/2. Каждый стрелок имеет право только на два выстрела. Какова вероятность того, что цель будет поражена? Какова вероятность того, что цель поразит первый стрелок?

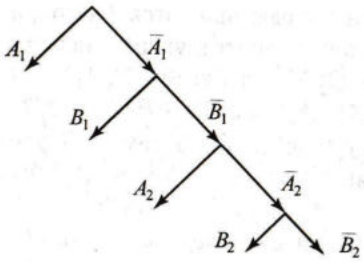


Рис. 4.1

◁ Обозначим через A_i попадание первого стрелка при i-м выстреле, а через B_i — попадание второго стрелка при i-м выстреле.

«Дерево» всех возможных способов протекания стрельбы изображено на рис. 4.1. Цель не будет поражена (событие $\bar{C}$), если произойдут события $\bar{A}_1$ и $\bar{B}_1$, и $\bar{A}_2$, и $\bar{B}_2$. Так как события независимы, то

$$P(\bar{C}) = P(\bar{A}_1) \cdot P(\bar{B}_1) \cdot P(\bar{A}_2) \cdot P(\bar{B}_2) = \\ = (2/3) \cdot (1/2) \cdot (2/3) \cdot (1/2) = 1/9.$$

Поэтому вероятность поражения цели

$$P(C) = 1 - P(\bar{C}) = 1 - 1/9 = 8/9.$$

Цель поразит первый стрелок (событие A), если он попадет при первом выстреле или при первом выстреле он не попадет в цель и второй стрелок при своем первом выстреле не попадет в цель и после этого первый стрелок попадет в цель.

Поэтому $A = A_1 + \bar{A}_1 \cdot \bar{B}_1 \cdot A_2$. События A_1 и $\bar{A}_1 \cdot \bar{B}_1 \cdot A_2$ несовместны. В силу независимости событий получаем

$$P(A) = P(A_1) + P(\bar{A}_1) \cdot P(\bar{B}_1) \cdot P(A_2) = \\ = 1/3 + 2/3 \cdot 1/2 \cdot 1/3 = 4/9. \triangleright$$

ПРИМЕР 4.4. Контракт между производителем и заказчиком предусматривает, что из каждой партии, состоящей из 100

изделий, наугад для проверки выбирают два изделия. Если выбранные изделия оба доброкачественные, то партия принимается. Если одно изделие доброкачественное, а другое дефектное, то наугад выбирают еще одно изделие и в зависимости от его качества решают вопрос о приеме партии. Какова вероятность приема партии, если на самом деле в ней содержится три бракованных изделия?

◁ Партия будет принята (событие A), если первые два проверенных изделия будут доброкачественными (событие B) или только одно из двух проверенных изделий будет доброкачественным (событие C) и третье проверенное изделие будет доброкачественным (событие D). Поэтому $A = B + C \cdot D$. Так как события B и $C \cdot D$ несовместны, а возможность появления события D зависит от события C, то

$$P(A) = P(B) + P(C) \cdot P(D/C) = \\ = \frac{C_{97}^2}{C_{100}^2} + \frac{C_{97}^1 \cdot C_3^1}{C_{100}^2} \cdot \frac{96}{98} \approx 0,998. \triangleright$$

ПРИМЕРЫ (для самостоятельного решения).

1. Из колоды карт одну за другой выбирают четыре карты. Какова вероятность того, что они все разных мастей? Какова вероятность того, что они все разного достоинства?

2. В ящике 10 теннисных мячей, в том числе семь новых и три побывавших в игре. Для игры наугад выбирают два мяча и после игры возвращают их обратно. Затем для второй игры также наугад извлекаются еще два мяча. Какова вероятность того, что вторая игра будет производиться новыми мячами?

3. На каждом крыле самолета по два мотора. Вероятность отказа в течение полета для каждого мотора равна p. Полет завершается благополучно, если на каждом крыле сохраняет работоспособность хотя бы один мотор. Какова вероятность благополучного полета?

4. Студент знает 20 вопросов из 30. Для сдачи экзамена достаточно правильно ответить на два предложенных вопроса или на один из предложенных вопросов и на один дополнительный вопрос. Какова вероятность того, что студент сдаст экзамен?

5. Два белых и два черных шара распределяют по двум урнам так, чтобы в каждой урне был хотя бы один шар. Затем наугад выбирают урну и из нее — один шар. Как следует распределить шары по урнам, чтобы вероятность вынуть белый шар (событие A) была наибольшей и какова эта вероятность?

6. Монету подбрасывают до тех пор, пока она два раза подряд не выпадет одной стороной. Найдите вероятность событий: $A =$ {опыт закончится до шестого броска}; $B =$ = {понадобится более четырех бросков}.

7. Из колоды карт (36 штук) одну за другой наугад выбирают карты. Найдите вероятности следующих событий: $A =$ {первый туз появится при третьем извлечении карты}; $B =$ {первый туз появится не ранее третьего извлечения карты}.

8. Из колоды карт (36 штук) выбирают карты по одной, пока не будет выбрана карта красной масти. Какова вероятность того, что придется выбрать более трех карт?

9. В урне восемь белых, шесть черных и два синих шара. Один за другим извлекаются наугад три шара. Какова вероятность того, что первый шар будет белым, второй — черным, третий — синим? Ответить на вопрос задачи в предположении: а) повторного выбора шаров; б) бесповторного выбора.

10. В первой урне четыре белых и два черных шара, а во второй урне три белых и пять черных шаров. Сначала подбрасывают монету и в зависимости от результата (герб или цифра) выбирают одну из урн, из которой затем наугад извлекают шар. Какова вероятность того, что шар окажется белым?

Ответы:

1. $\frac{729}{6545} \approx 0,11$, $\frac{512}{935} \approx 0,55$;
2. $\approx 0,29$;
3. $(1-p^2)^2$;
4. $\frac{152}{203} \approx \frac{3}{4}$;
5. Следует в одну урну положить белый шар, а остальные — во вторую урну; тогда $P(A) = \frac{2}{3}$;
6. $P(A) = \frac{15}{16}$, $P(B) = \frac{1}{8}$;
7. $P(A) = \frac{496}{5355} \approx 0,09$, $P(B) = \frac{248}{315} \approx 0,79$;
8. $\frac{4}{35}$;
9. а) $\frac{3}{128}$, б) $\frac{1}{35}$;
10. $\frac{25}{48}$.

ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ И ФОРМУЛЫ БАЙЕСА

Пусть событие A может произойти *с одним и только с одним из несовместимых событий* $B_1, B_2, ..., B_n$, *образующих полную группу*. Иными словами, событие A появится, если произойдет событие B_1 и при этом появится событие A, или произойдет событие B_2 и при этом появится событие A и т. д. Символическая запись этой фразы имеет вид

$$A = B_1 \cdot A + B_2 \cdot A + ... + B_n \cdot A. \tag{5.1}$$

В силу несовместимости событий можно записать

$$P(A) = P(B_1 \cdot A) + P(B_2 \cdot A) + ... + P(B_n \cdot A).$$

Используя теорему умножения вероятностей, получаем формулу

$$P(A) = P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) + ... \\ ... + P(B_n) \cdot P(A/B_n),$$

которая и называется *формулой полной вероятности*.

Мы привели целиком вывод формулы полной вероятности потому, что целесообразно в решении задач следовать именно этой схеме действий: сначала ввести обозначения, затем после соответствующих рассуждений получить запись вида (5.1), и только после этого вычислять вероятности. Тогда не нужно будет даже задаваться вопросом: на формулу полной вероятности данная задача или нет? Логика рассуждений неминуемо к ней приведет.

ПРИМЕР 5.1. Представим себе странника, который выходит из пункта O и на разветвлении дорог выбирает наугад один из возможных путей. Схема дорог изображена на рис. 5.1. Какова вероятность того, что путник, двигаясь описанным образом, попадет в пункт A?

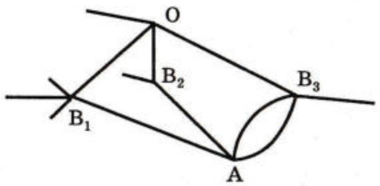


Рис. 5.1

◁ Обозначим прибытие путника в пункт той же буквой, что и сам пункт. Легко видеть, что $P(B_1) = P(B_2) = P(B_3) = 1/4$. Так как равновозможен выбор любого из путей, ведущих из пункта O. Путник попадет в пункт A, если он выберет дорогу в пункт B_1 и оттуда дорогу в пункт A, или он выберет дорогу в пункт B_2 и оттуда дорогу в пункт A, или он попадет сначала в пункт B_3 и оттуда — в пункт A. Символически это рассуждение можно записать в виде

$$A = B_1 \cdot A + B_2 \cdot A + B_3 \cdot A.$$

Из чего следует, что

$$P(A) = P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) + P(B_3) \cdot P(A/B_3) =$$
$$= 1/4 \cdot 1/4 + 1/4 \cdot 1/2 + 1/4 \cdot 2/3 = 17/48 \approx 1/3,$$

где вероятности $P(A/B_1)$, $P(A/B_2)$, $P(A/B_3)$ определены с учетом числа равновозможных путей, ведущих из соответствующего города. ▷

ПРИМЕР 5.2. Из полного набора костей домино (28 штук) наугад выбирают одну кость, затем вторую. Какова вероятность того, что вторую кость можно приставить (в игровом смысле) к первой?

◁ Различают кости двух типов:

а) «дупель» — кость с одной и той же цифрой на концах. Дупелей семь от «пусто–пусто» до «шесть–шесть». Вероятность выбрать сначала дупель равна $7/28 = 1/4$;

б) обычная кость с разными цифрами на концах. Заметим, что каждая цифра содержится на семи костях, включая дупель. Вероятность в качестве первой кости выбрать обычную кость равна $21/28 = 3/4$.

Вторую кость можно будет приставить к первой (событие A), если будет выбран дупель (событие B_1) и к нему можно будет приставить вторую кость, или сначала будет выбрана обычная кость (событие B_2) и к ней можно будет приставить вторую кость. Символически это можно записать в виде

$$A = B_1 \cdot A + B_2 \cdot A.$$

Так как события B_1 и B_2 несовместны, то

$$P(A) = P(B_1 \cdot A) + P(B_2 \cdot A) =$$
$$= P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) =$$
$$= (1/4) \cdot (6/27) + (3/4) \cdot (12/27) = 7/18 \approx 1/3. \ ▷$$

ПРИМЕР 5.3. Имеется две коробки деталей, в каждой из которых по 10 деталей. В первой коробке среди деталей две низкого сорта, а во второй — четыре низкосортных детали. Из первой коробки для нужд производства взяли наугад половину деталей, а оставшиеся высыпали во вторую коробку. Через некоторое время из второй коробки взяли наугад деталь. Какова вероятность того, что это деталь низкого сорта?

◁ Обозначим через A событие, состоящее в выборе из второй коробки детали низкого сорта. Возможность этого выбора зависит от того, какие именно детали были добавлены во вторую коробку. На этот счет можно выдвинуть следующие предположения: B_1 — во вторую коробку добавили пять годных деталей; B_2 — добавили одну деталь низкого сорта и четыре доброкачественные; B_3 — добавили две детали низкого сорта и три доброкачественные. Пять деталей во вторую коробку можно переложить $n = C_{10}^5 = \dfrac{10!}{5! \cdot 5!} = 252$ способами. Из них событию B_1 благоприятствует $C_8^5 = 56$, событию B_2 — $C_2^1 \cdot C_8^4 = 2 \cdot 70 = 140$, а событию B_3 — $C_2^2 \cdot C_8^3 = 56$ способов. Событие A произойдет, если появится событие B_1 и после этого произойдет событие A, или появится событие B_2 и после этого произойдет событие A, или появится B_3 и после этого произойдет A. Символически: $A = B_1 \cdot A + B_2 \cdot A + B_3 \cdot A$.

Учитывая несовместность событий B_i, имеем

$$P(A) = P(B_1 \cdot A) + P(B_2 \cdot A) + P(B_3 \cdot A) =$$

$$= P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) + P(B_3) \cdot P(A/B_3) =$$

$$= (56/252) \cdot (4/15) + (140/252) \cdot (5/15) + (56/252) \cdot (6/15) = 1/3. \triangleright$$

Пусть событие A может наступить только при появлении одного из несовместных событий $B_1, B_2, ..., B_n$. В этих условиях вероятность события A можно вычислить по формуле полной вероятности. События B_i иногда называют «гипотезами», поскольку можно лишь предполагать какое именно из них произойдет. Предположим, что известны вероятности $P(B_1), P(B_2), ..., P(B_n)$.

Проделан опыт, в результате которого событие A произошло. Вычислим вероятности событий B_i, $i = 1, 2, ..., n$, при условии появления события A. По теореме умножения вероятностей

$$P(A \cdot B_i) = P(A) \cdot P(B_i/A) = P(B_i) \cdot P(A/B_i).$$

Откуда

$$P(B_i/A) = \frac{P(B_i) \cdot P(A/B_i)}{P(A)}.$$

Используя формулу полной вероятности для события A, получаем *формулы Байеса*:

$$P(B_i/A) = \frac{P(B_i) \cdot P(A/B_i)}{\sum\limits_{j=1}^{n} P(B_j) \cdot P(A/B_j)}, \quad i = 1, 2, ..., n. \qquad (5.2)$$

Формулы Байеса позволяют переоценивать вероятности гипотез (событий B_i) с учетом информации, которую содержит в себе факт появления события A.

ПРИМЕР 5.4. По каналу связи передают сигналы «0» и «1» с вероятностями 0,3 и 0,7 соответственно. Под действием помех возникают ошибки приема сигнала. Вероятность принять «1», если передавался «0», равна 0,05. Вероятность принять «0», если передавалась «1», равна 0,1. Какова вероятность правильного приема сигнала? Если принята «1», то какова вероятность того, что действительно передавалась «1»?

$\triangleleft$ Обозначим передачу сигнала «1» через B_1, а сигнала «0» через B_2. Сигнал будет принят правильно (событие A), если будет передана «1» и она будет принята правильно или будет передан «0» и он будет принят правильно. Тогда $A = B_1 \cdot A + B_2 \cdot A$. В силу несовместности событий B_1 и B_2 имеем

$$P(A) = P(B_1 \cdot A) + P(B_2 \cdot A) =$$
$$= P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) =$$
$$= 0,7 \cdot 0,9 + 0,3 \cdot 0,95 = 0,915.$$

Пусть событие C означает, что принят сигнал «1». При этом условии вычислим вероятность того, что действительно была передана «1», т. е. вероятность $P(B_1/C)$. По формуле Байеса

$$P(B_1/C) = \frac{P(B_1) \cdot P(C/B_1)}{P(B_1) \cdot P(C/B_1) + P(B_2) \cdot P(C/B_2)} =$$
$$= \frac{0,7 \cdot 0,9}{0,7 \cdot 0,9 + 0,3 \cdot 0,05} = 0,9767441 \approx 0,977. \triangleright$$

ПРИМЕР 5.5. Девяносто шесть процентов изделий некоторого производства удовлетворяют стандарту. Используется упрощенная схема проверки качества, дающая положительный результат с вероятностью 0,98 для стандартных изделий, для нестандартных изделий — с вероятностью 0,05. Изделие выдержало проверку. Какова вероятность того, что оно действительно стандартно?

$\triangleleft$ Факт прохождения контроля обозначим через A. В отношении проверенного изделия можно выдвинуть два предположения: а) оно стандартно (событие B_1); б) оно нестандартно (событие B_2). По условию задачи необходимо найти вероятность $P(B_1/A)$. По формуле Байеса

$$P(B_1/A) = \frac{P(B_1) \cdot P(A/B_1)}{P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2)} =$$
$$= \frac{0,96 \cdot 0,98}{0,96 \cdot 0,98 + 0,04 \cdot 0,05} = \frac{2352}{2355} \approx 0,9988. \triangleright$$

ПРИМЕР 5.6. По каналу связи передается одна из последовательностей букв $AAAA$, $BBBB$, $CCCC$ с вероятностями соответственно 0,5, 0,4, 0,1. Каждая передаваемая буква принимается правильно с вероятностью 0,8 и с вероятностями 0,1 и 0,1 за две другие буквы. Предполагается, что искажаются буквы при передаче независимо друг от друга. Найти вероятность того, что передано $AAAA$, если принято $ABCA$.

◁ Для краткости записи формулы обозначим $AAAA$ через T_1, $BBBB$ через T_2, $CCCC$ через T_3. Тогда по формуле Байеса

$$P(T_1/ABCA) =$$

$$= \frac{P(T_1)\cdot P(ABCA/T_1)}{P(T_1)\cdot P(ABCA/T_1)+P(T_2)\cdot P(ABCA/T_2)+P(T_3)\cdot P(ABCA/T_3)} =$$

$$= \frac{0{,}5\cdot 0{,}8\cdot 0{,}1\cdot 0{,}1\cdot 0{,}8}{0{,}5\cdot 0{,}8\cdot 0{,}1\cdot 0{,}1\cdot 0{,}8+0{,}4\cdot 0{,}1\cdot 0{,}8\cdot 0{,}1\cdot 0{,}1+0{,}1\cdot 0{,}1\cdot 0{,}1\cdot 0{,}8\cdot 0{,}1} = \frac{8}{9}. \triangleright$$

ПРИМЕР 5.7. В одной урне лежит белый шар, в другой — черный. Урны внешне не различимы. В выбранную наугад урну положили белый шар. Шары перемешали и наугад выбрали один шар. Он оказался белым. Какова вероятность того, что в урне остался тоже белый шар?

◁ Обозначим выбор белого шара через A. В отношении цвета оставшегося шара можно сделать два предположения: B_1 — этот шар белый; B_2 — этот шар черный. Вероятности этих предположений $P(B_1) = P(B_2) = 1/2$. В условиях задачи требуется вычислить вероятность $P(B_1/A)$. По формуле Байеса

$$P(B_1/A) = \frac{P(B_1)\cdot P(A/B_1)}{P(B_1)\cdot P(A/B_1) + P(B_2)\cdot P(A/B_2)} =$$

$$= \frac{0{,}5\cdot 1}{0{,}5\cdot 1 + 0{,}5\cdot 0{,}5} = \frac{2}{3}. \triangleright$$

ПРИМЕР 5.8. Три стрелка производят по одному выстрелу в одну и ту же мишень. Вероятности попадания в мишень при одном выстреле для этих стрелков соответственно равны 0,8, 0,7 и 0,6. Какова вероятность того, что третий стрелок промахнулся, если в мишени оказалось две пробоины?

◁ Обозначим через A событие, состоящее в появлении двух пробоин в мишени. В отношении двух пробоин могут быть три предположения: B_1 — попали первый и второй стрелки, а третий не попал, вероятность чего равна $P(B_1) = 0{,}8\cdot 0{,}7\cdot 0{,}4 = 0{,}224$; B_2 — попали первый и третий, а второй не попал, вероятность чего равна $P(B_2) = 0{,}8\cdot 0{,}3\cdot 0{,}6 = 0{,}144$; B_3 — попали второй и третий, а первый не попал, вероятность чего равна $P(B_3) = 0{,}2\cdot 0{,}7\cdot 0{,}6 = 0{,}084$. Заметим, что $P(A/B_i) = 1$, $i = 1,2,3$. Тогда по формуле Байеса

$$P(B_1/A) = \frac{P(B_1)\cdot P(A/B_1)}{P(B_1)\cdot P(A/B_1)+P(B_2)\cdot P(A/B_2)+P(B_3)\cdot P(A/B_3)} =$$

$$= \frac{0{,}224}{0{,}224+0{,}144+0{,}084} = \frac{56}{113} \approx \frac{1}{2}. \triangleright$$

ПРИМЕРЫ (для самостоятельного решения).

1. Из условий примера 5.1 стало известно, что путник пришел в пункт A. Какова вероятность того, что он попал в A через пункт B_i?

2. Известно, что в партии из 10 изделий с равными вероятностями может оказаться от 0 до 2 изделий со скрытым дефектом. Проверили пять изделий, взятых наугад из этой партии. Среди проверенных не оказалось изделий с дефектами. Какова вероятность того, что в оставшейся половине партии нет изделий со скрытыми дефектами?

3. Партия транзисторов, среди которых 10% дефектных, поступила на проверку. Схема проверки такова, что дефект обнаруживается с вероятностью 0,95. Вероятность того, что исправный транзистор будет признан дефектным, равна 0,03. Какова вероятность того, что выбранный наугад для проверки транзистор, будет признан дефектным?

4. В цехе болты изготовляют три автоматических станка A, B и C. Станок A производит 25% болтов, станок B — 35% и станок C — 40% всех болтов. В продукции станков брак составляет соответственно 5%, 4% и 2%. Наугад взятый болт оказался дефектным. Какова вероятность того, что он был произведен машиной A?

5. Из урны с четырьмя белыми и двумя черными шарами два шара, взятые наугад, были перенесены в урну с двумя белыми и тремя черными шарами. Какова после этого вероятность вынуть белый шар из второй урны?

6. Из урны, содержащей пять белых и три черных шара, наугад без возвращения выбирают три шара. Третий шар оказался белым. Какова вероятность того, что первые два шара были разного цвета?

7. Первая урна содержит один черный и один красный шар, а во второй урне два черных и три красных шара. Из первой урны вынули наугад один шар и переместили его во вторую урну. Затем вынули из второй урны наугад один шар. Какова вероятность того, что были вынуты шары одного цвета? Какова вероятность того, что из первой урны был вынут красный шар, если из второй урны был вынут черный шар?

8. Стрелок попадает в цель с вероятностью 1/5 при каждом выстреле. Монету подбрасывают три раза, и стрелку предоставляется столько выстрелов, сколько раз выпал герб. Найдите вероятность поражения цели.

9. Из урны, в которой было четыре белых и три черных шара, убрали один из шаров неизвестного цвета. Для определения состава урны наугад выбрали два шара. Они оказались черными. Какова вероятность того, что первый вынутый шар был белым?

10. По каналу связи передается одна из двух команд управления в виде кодовых комбинаций 11111 или 00000, причем вероятности этих комбинаций равны соответственно 0,7 и 0,3. Из-за наличия помех вероятность правильного приема каждого символа 0 или 1 равна 0,6 независимо от правильности приема остальных символов. На приемном устройстве получена комбинация 10110. Какова вероятность того, что была передана комбинация 11111?

Ответы:

1. $P(B_1/A) = 3/17$, $P(B_2/A) = 6/17$, $P(B_3/A) = 8/17$;
2. $18/31 \approx 3/5$;
3. 0,122;
4. $25/69 \approx 0,36$;
5. 10/21;
6. 4/7;
7. 7/12, 2/5;
8. $0,271$;
9. 0,8;
10. $7/9 \approx 0,78$.

Занятие 6

СХЕМА НЕЗАВИСИМЫХ ИСПЫТАНИЙ. ФОРМУЛА ПУАССОНА. ПРОСТЕЙШИЙ ПОТОК СОБЫТИЙ

Опыты называются независимыми, если вероятность каждого исхода любого опыта не изменяется от того, какие исходы имели другие опыты.

Пусть производится n независимых опытов и известна $P(A) = p$ — вероятность появления события A в каждом из опытов, $P(\bar{A}) = 1 - p = q$. Тогда вероятность того, что в n независимых опытах событие A появится ровно k раз, равна по *формуле Бернулли*

$$P_n(k) = C_n^k p^k q^{n-k}. \qquad (6.1)$$

При большом числе опытов n формула Бернулли приводит к большому объему вычислений. Существуют приближенные формулы для вычисления вероятностей $P_n(k)$, которые дают тем большую точность, чем больше число n. Пусть число независимых опытов n велико (чем больше, тем лучше), а вероятность события p мала (чем меньше, тем лучше, но $p > 0$). Тогда

$$P_n(k) \approx \frac{\lambda^k e^{-\lambda}}{k!}, \qquad (6.2)$$

где $\lambda = np$. Это и есть *формула Пуассона*. Формула Пуассона дает приемлемую точность, если опытов производится хотя бы несколько десятков, а $p < 0,1$.

Значение $k = k_0$, при котором вероятность (6.1) принимает наибольшее значение, называют *наивероятнейшим числом появлений события*. Если $(n+1)p$ — число дробное, то k_0 принимает единственное значение, равное целой части числа k_0. Если величина $(n+1)p$ — целая, то k_0 принимает два значения: $k_0 = (n+1)p - 1$ и $k_0 = (n+1)p$.

Пусть в результате каждого из n независимых опытов может произойти одно из m попарно несовместных событий $A_1, A_2, ..., A_m$ с вероятностями $p_1, p_2, ..., p_m$ соответственно ($p_1 + p_2 + ... + p_m = 1$). Тогда вероятность того, что событие A_1 наступит k_1 раз, событие A_2 наступит k_2 раза, ..., событие A_m наступит k_m раз, причем $k_1 + k_2 + ... + k_m = n$, равна:

$$P_n(k_1, k_2, ..., k_m) = \frac{n!}{k_1!k_2!...k_m!} p_1^{m_1} p_2^{m_2} ... p_k^{m_k}. \tag{6.3}$$

ПРИМЕР 6.1. Предположим, что 30% студентов нашего института занимаются спортом. Какова вероятность того, что среди первых пяти встречных студентов окажется только один спортсмен? Какова вероятность того, что среди них есть хотя бы один спортсмен? Каково наиболее вероятное число спортсменов среди них?

$\triangleleft$ Так как студентов в институте много (несколько тысяч), то по мере опроса нескольких из них пропорции в оставшейся части практически не изменяются. Поэтому можно считать опрос каждого студента независимым опытом. Всего опытов производится $n = 5$, а вероятность положительного ответа $p = 0,3$. По формуле Бернулли имеем $P_5(1) = C_5^1 \cdot 0,3 \cdot (0,7)^4 = 0,36015$. Вероятность хотя бы одного правильного ответа проще вычислять, если перейти к противоположному событию:

$$P_5(k \geqslant 1) = 1 - P_5(0) = 1 - (0,7)^5 = 1 - 0,16807 = 0,83193.$$

Так как $(n+1)p = (5+1)0,3 = 1,8$ (целая часть числа равна 1), то наиболее вероятное число спортсменов среди пяти опрошенных $k_0 = 1$. $\triangleright$

ПРИМЕР 6.2. На каждый вопрос предлагается три ответа, среди которых следует выбрать один правильный. Задано пять вопросов. Какова вероятность того, что путем простого угадывания удастся правильно ответить на четыре вопроса? Какова вероятность угадать правильный ответ хотя бы на один вопрос?

$\triangleleft$ Выбор ответа на вопрос можно рассматривать как независимый опыт. Всего таких опытов производится $n = 5$, а вероятность успеха в каждом опыте равна $p = 1/3$. Тогда вероятность путем простого угадывания правильно ответить на четыре вопроса равна $P_5(4) = C_5^4 \cdot (1/3)^4 \cdot (2/3)^1 = 10/243 \approx 1/24$. Вероятность угадать хотя бы один правильный ответ равна

$$P_5(k \geqslant 1) = 1 - P_5(0) = 1 - (2/3)^5 = 1 - 32/243 \approx 7/8. \triangleright$$

ПРИМЕР 6.3. Вероятность попадания в цель при выстреле равна 0,3. Сколько нужно сделать выстрелов, чтобы вероятность поражения цели была больше 0,9?

$\triangleleft$ Каждый выстрел можно рассматривать как независимое испытание, и в каждом из них вероятность появления события (попадания в цель) равна $p = 0,3$. Цель будет поражена, если в n выстрелах будет хотя бы одно попадание, вероятность чего равна:

$$P_n(k \geqslant 1) = 1 - P_n(0) > 0,9 \Rightarrow 1 - (0,7)^n > 0,9 \Rightarrow$$

$$\Rightarrow (0,7)^n < 0,1 \Rightarrow n \geqslant 7. \triangleright$$

ПРИМЕР 6.4. В цехе шесть станков, которые работают независимо друг от друга. В течение рабочего дня (8 часов) каждый станок простаивает в сумме 2 часа. Какова доля времени, в течение которой в цехе работает не менее пяти станков?

$\triangleleft$ Наблюдение над состоянием каждого станка можно рассматривать как независимый опыт, число которых $n = 6$. Вероятность застать станок работающим равна $p = 6/8 = 3/4$ (в соответствии с геометрическим определением вероятности). Тогда вероятность застать работающими не менее пяти станков равна

$$P_6(k \geqslant 5) = P_6(5) + P_6(6) =$$

$$= C_6^5 \cdot (3/4)^5 \cdot (1/4)^1 + C_6^6 \cdot (3/4)^6 \cdot (3/4)^0 \approx 0,53. \triangleright$$

ПРИМЕР 6.5. Монету подбрасывают до тех пор, пока герб не выпадет три раза. Какова вероятность того, что до этого цифра выпадет пять раз?

◁ Всего должно состояться восемь подбрасываний монеты. Чтобы опыт закончился именно на восьмом броске, необходимо при восьмом подбрасывании получить герб (вероятность этого события составляет 1/2) и до этого при семи подбрасываниях герб должен выпасть ровно два раза (вероятность события по формуле Бернулли равна $P_7(2) = C_7^2 \cdot (1/2)^2 \cdot (1/2)^5 = 21/128$). В силу независимости опытов вычислили искомую вероятность:

$$(21/128) \cdot (1/2) = 21/256 \approx 1/13. \triangleright$$

ПРИМЕР 6.6. Вероятность того, что изделие при транспортировке с завода повредится, равна 0,0005. С завода отправлено четыре тысячи изделий. Какова вероятность того, что в пути повредится больше двух изделий?
◁ Транспортировку каждого изделия можно рассматривать как независимый опыт, число которых ($n = 4000$) велико. Вероятность же появления события в каждом опыте ($p = 0,0005$) мала. Это дает основание воспользоваться для вычислений формулой Пуассона. Заметим, что $\lambda = n \cdot p = 4000 \cdot 0,0005 = 2$. Нас интересует вероятность $P_{4000}(k > 2) = P_{4000}(3) + P_{4000}(4) + ... + P_{4000}(4000)$. Проще эту вероятность вычислить, если рассмотреть вероятность противоположного события:

$$\begin{aligned} P_{4000}(k > 2) &= 1 - P_{4000}(k \leqslant 2) = \\ &= 1 - P_{4000}(0) - P_{4000}(1) - P_{4000}(2) = \\ &= 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} - \frac{2^2}{2!}e^{-2} = \\ &= 1 - e^{-2} - 2e^{-2} - 2e^{-2} = 1 - 5e^{-2} \approx 0,31. \triangleright \end{aligned}$$

ПРИМЕР 6.7. Среди 300 изделий 15 бракованных. Для проверки наугад выбрали пять изделий. Какова вероятность того, что среди них нет бракованных? Сравнить точное значение вероятности с приближенным, найденным по формуле Бернулли.
◁ Пусть A — интересующее нас событие. Выбрать пять изделий из 300 можно C_{300}^5 способами. Событию A благоприятствуют те способы выбора, при которых пять изделий выбирается из 285 годных. Число таких способов равно C_{285}^5. Точное

значение искомой вероятности (с точностью до одной десятитысячной) равно $P(A) = C_{285}^5 / C_{300}^5 = 0,7724$. Так как партия изделий велика, то выбор одного за другим нескольких изделий не меняет заметно пропорции в этой партии. Поэтому можно считать, что производится пять независимых опытов и что вероятность выбора бракованного изделия в каждом опыте примерно равна $p = 15/300 = 0,05$. По формуле Бернулли:

$$P(A) = P_5(0) = C_5^0 \cdot (0,05)^0 \cdot (0,95)^5 = 0,7738. \triangleright$$

ПРИМЕР 6.8. Известно, что из каждых 1000 элементов в среднем 999 сохраняют свою работоспособность в течение гарантийного срока. Какова вероятность того, что из 3000 элементов все до единого сохранят свою работоспособность в течение гарантийного срока?
◁ Работу каждого элемента в течение гарантийного срока можно считать независимым опытом. Число опытов велико ($n = 3000$). Вероятность того, что элемент сохранит работоспособность в течение гарантийного срока, равна 0,999. Формула Бернулли из-за большого числа опытов для расчетов неприемлема. Для применения формулы Пуассона будем говорить не о работоспособных элементах, а об элементах вышедших из строя. Вероятность выхода из строя элемента $p = 0,001$. Тогда $\lambda = np = 3000 \cdot 0,001 = 3$. Все 3000 элементов сохранят свою работоспособность, если ни один из них не выйдет из строя. По формуле Пуассона:

$$P_{3000}(0) = \frac{3^0}{0!}e^{-3} \approx 0,05. \triangleright$$

Потоком событий называется последовательность событий, наступающих одно за другим в случайные моменты времени. Обозначим через $P_k(t)$ вероятность появления k событий на интервале времени длины t.

Поток событий называется *простейшим,* если выполняются следующие условия:

1) появление того или иного числа событий на интервале времени длины t зависит только от длины этого интервала и не зависит от его расположения на оси и от событий, происходящих вне этого интервала;

2) вероятность появления события за малый промежуток времени $\triangle t$ пропорциональна длине этого промежутка, т. е. $P_1(\triangle t) = \mu \cdot \triangle t$, где μ — некоторая постоянная;

3) вероятность появления двух и более событий за малый промежуток времени $\triangle t$ есть величина более высокого порядка малости по сравнению с $\triangle t$.

Из условий 1–3 следует, что $P_k(t) = \dfrac{(\mu t)^k}{k!} \exp(-\mu t)$.

Условие 1 означает, что события происходят независимо друг от друга и характеристики этого потока не изменяются со временем (поток стационарный). Условия 2 и 3 означают, что события происходят по одному, а не группами (поток ординарен). Величина μ называется *интенсивностью простейшего потока*, и равна среднему числу событий в единицу времени. Величина μt равна среднему числу событий за время t.

Приведенный результат можно упрощенно сформулировать следующим образом.

Пусть события происходят независимо друг от друга во времени (или в пространстве) и по одному. Тогда, если на отрезок времени (или пространства) приходится в среднем λ событий, то вероятность появления на этом отрезке k событий приблизительно равна

$$P_k = \frac{\lambda^k \exp(-\lambda)}{k!}. \ \triangleright$$

ПРИМЕР 6.9. Известно, что наборщик в среднем допускает одну ошибку на две страницы текста. В набранной книге взяли наугад страницу. Какова вероятность того, что на этой странице содержится более одной опечатки?

$\triangleleft$ Опечатки появляются по одной и независимо друг от друга. Условия простейшего потока приблизительно выполняются, и формула Пуассона приблизительно верна. На одну страницу приходится в среднем $\lambda = 1/2$ опечатки. Поэтому вероятность того, что на данной странице содержится более одной опечатки, равна

$$P(k > 1) = 1 - P(k = 0) - P(k = 1) = \\ = 1 - \frac{(1/2)^0}{0!} e^{-1/2} - \frac{(1/2)^1}{1!} e^{-1/2} \approx 0,1. \ \triangleright$$

ПРИМЕР 6.10. В тесто засыпали изюм из расчета по пяти изюмин на одну булку и тщательно перемешали тесто. Какова вероятность того, что взятая наугад булка содержит хотя бы одну изюмину?

$\triangleleft$ При тщательном перемешивании теста изюмины распределяются практически независимо друг от друга и по одной. Условия простейшего потока приблизительно выполнены, для оценки искомой вероятности можно использовать формулу Пуассона. Среднее число изюмин, приходящихся на одну булку, равно $\lambda = 5$, поэтому

$$P(k \geqslant 1) = 1 - P(k = 0) = 1 - e^{-5} \approx 0,994. \ \triangleright$$

ПРИМЕР 6.11. Стрелок попадает в «десятку» с вероятностью 0,4, в «девятку» — с вероятностью 0,3, в «восьмерку» — с вероятностью 0,2, в «семерку» — с вероятностью 0,1. Какова вероятность того, что при шести выстрелах стрелок трижды попадет в «десятку» и по одному разу в «девятку», «восьмерку» и «семерку»?

$\triangleleft$ Каждый выстрел можно считать независимым испытанием, в котором может появиться одно из несовместных событий: «10», «9», «8» или «7». Поэтому по формуле (6.3)

$$P_6(3,1,1,1) = \frac{6!}{3!1!1!1!} \cdot (0,4)^3 \cdot 0,3 \cdot 0,2 \cdot 0,1 = 0,04608 \approx 0,05. \ \triangleright$$

ПРИМЕРЫ (для самостоятельного решения).

1. Игральный кубик подброшен пять раз. Какова вероятность того, что два раза выпадет шесть очков?

2. Вероятность рождения мальчика равна 1/2. Какова вероятность того, что в семье с четырьмя детьми два мальчика и две девочки?

3. Первый стрелок попадает в цель с вероятностью 1/3 и производит четыре выстрела. Второй стрелок попадает в цель с вероятностью 1/2 и производит три выстрела. Для какого стрелка в этих условиях вероятнее попасть в цель дважды?

4. Прибор состоит из пяти блоков. Надежность (вероятность безотказной работы в течение заданного времени t) для каждого блока равна 0,9. Узлы выходят из строя независимо друг от друга. Найти вероятность того, что за время t:

а) откажет только один узел; б) откажет хотя бы один узел; в) откажут не менее двух узлов.

5. Вероятность выпуска детали со скрытым дефектом равна 0,02. Детали упаковывают в ящики по 100 штук. Какова вероятность того, что в данном ящике нет деталей со скрытым дефектом? Какова вероятность того, что в данном ящике больше двух деталей со скрытым дефектом? Каково наиболее вероятное число деталей со скрытым дефектом в данном ящике?

6. В крупной партии изделий 10% имеют низкое качество. Сколько нужно проверить изделий, чтобы с вероятностью не менее 0,95 обнаружить хотя бы одно изделие низкого качества?

7. В течение часа на коммутатор поступает в среднем 120 телефонных вызовов. Какова вероятность того, что в течение заданной минуты поступит четыре вызова?

8. Вероятность того, что при перевозке изделие повредится, равна $p = 0,005$. С завода отправлено четыреста изделий. Найдите вероятность того, что в пути повредится более двух изделий.

9. По каналу связи передается цифровой текст. Из-за помех каждая цифра может быть принята неправильно с вероятностью 0,0025. Какова вероятность того, что в тексте, состоящем из 800 цифр, все цифры будут приняты правильно?

10. При стрельбе из зенитного орудия в среднем только один выстрел из 200 достигает цели. Какова вероятность того, что при 100 выстрелах будет хотя бы одно попадание?

11. Дальтоники составляют 1% населения. Какова вероятность того, что среди пятидесяти студентов окажется по меньшей мере один дальтоник?

12. Двое бросают монету по пять раз каждый. Какова вероятность того, что у них выпадет одинаковое число гербов?

13. Частица в начальный момент времени находится в начале координат. Каждую последующую секунду она сдвигается вправо с вероятностью 1/3 или влево с вероятностью 2/3 независимо от того, как она двигалась в предыдущие секунды. Какова вероятность того, что через пять секунд частица окажется в отрезке $[-2; 0]$?

14. В биатлоне на каждом из трех огневых рубежей спортсмен должен поразить пять мишеней в пяти выстрелах. За каждую непораженную мишень спортсмен обязан пробежать штрафной круг. Пусть вероятность поражения мишени при одном выстреле равна 0,95. Какова вероятность того, что спортсмен все три огневых рубежа пройдет без штрафных кругов? Какова вероятность того, что после каждого огневого рубежа спортсмен будет пробегать один штрафной круг?

15. Вероятность попадания в цель при одном выстреле равна 1/3. Стрелок стреляет по цели до тех пор, пока не накопится три попадания. Какова вероятность того, что к этому моменту у стрелка будет два промаха?

Ответы:

1. $1250/7776 \approx 0,16$;
2. $3/8$;
3. Для второго;
4. а) $\approx 0,328$,
   б) $1-(0,9)^5 \approx 0,4$,
   в) $\approx 0,08$;
5. $e^{-2} \approx 0,14$,
   $1-5e^{-2} \approx 0,3$, $k_0 = 2$;
6. $n \geqslant 29$;
7. $(2e^{-2})/3 \approx 0,09$;
8. $1-5e^{-2} \approx 0,31$;
9. $e^{-2} \approx 0,136$;
10. $1-e^{-1/2} \approx 0,4$;
11. $1-e^{-0,5} \approx 0,4$;
12. $63/256 \approx 1/4$;
13. $80/243 \approx 1/3$;
14. $\approx 0,46$, $\approx 0,008$;
15. $8/81$.

Лекция 11

ВЫБОРОЧНЫЙ МЕТОД

§ 11.1. ПРЕДМЕТ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Вспомним, что теория вероятностей изучает математические модели случайных явлений, т. е. вероятностные пространства $(\Omega, \mathcal{F}, P)$. При этом предполагается, что соответствующая модель задана заранее или ее можно получить, описав некоторые ее особые свойства (например, классическая вероятностная модель получена из предположения о равнораспределенности элементарных исходов; геометрические вероятности — из предположения о том, что вероятность попадания в любой достаточно малый прямоугольник не зависит от его расположения).

Математическая статистика призвана определить, какая именно вероятностная модель описывает данное случайное явление, по результатам наблюдения за этим явлением. Сосредоточимся на классическом разделе математической статистики, посвященном определению распределения и числовых характеристик случайной величины по результатам наблюдения за ней.

Пусть осуществлено n реализаций случайного эксперимента и получены числовые значения $x_1, x_2, ..., x_n$ некоторой случайной величины X. Это означает, что получены значения n независимых случайных величин $x_1, x_2, ..., x_n$, распределенных так же, как наблюдаемая случайная величина X (причем указанные значения получены при одном и том же элементарном исходе w). Указанный набор чисел $(x_1, x_2, ..., x_n)$ называется *выборкой объема* n случайной величины X. Очевидно,

что другой наблюдатель при реализации своих экспериментов получит совсем другую выборку для той же случайной величины X. Поэтому на выборку можно смотреть двояко. С одной стороны, это набор из n чисел, а с другой — это n независимых случайных величин, распределенных так же, как наблюдаемая случайная величина X.

Принципиальный вопрос состоит в следующем. Можно ли по выборке, т. е. числам $x_1(\omega), x_2(\omega), ..., x_n(\omega)$, делать выводы о распределении и числовых характеристиках случайной величины X? Положительный ответ на этот вопрос дает усиленный закон больших чисел, по которому почти наверное

$$\lim_{n\to\infty} \frac{x_1(\omega) + x_2(\omega) + ... + x_n(\omega)}{n} = MX.$$

Следовательно, если объем n выборки $(x_1, x_2, ..., x_n)$ достаточно велик, то число $(x_1(\omega) + x_2(\omega) + ... + x_n(\omega))/n$ дает хорошее представление о математическом ожидании случайной величины X. Что касается других числовых характеристик, то речь о них пойдет в следующем разделе.

§ 11.2. ВЫБОРОЧНЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Пусть наблюдается случайная величина X и $(x_1, x_2, ..., x_n)$ — ее выборка. Как по выборке оценить функцию распределения $F(x) = P(X \leqslant x)$? Случайное событие $\{X \leqslant x\}$ осуществилось столько раз, сколько чисел x_i удовлетворяет неравенству $x_i \leqslant x$. Обозначим это число $\mu_n(x)$. Тогда частота события $\{X \leqslant x\}$ равна $\mu_n(x)/n$. По свойству устойчивости частот, выведенному из усиленного закона больших чисел (см. лекцию № 10), почти наверное

$$P(X \leqslant x) = \lim_{n\to\infty} \frac{\mu_n(x)}{n}.$$

Таким образом, хорошей оценкой для $F(x)$ является величина $\mu_n(x)/n$. Попытаемся понять смысл этой величины, для чего введем вспомогательную дискретную случайную величину $\hat{X}$ с рядом распределения

$\hat{X}$	x_1	x_2	...	x_n
P	$1/n$	$1/n$	...	$1/n$

Обозначим $\hat{F}(x)$ функцию распределения этой случайной величины. По свойству ряда распределения

$$\hat{F}(x) = P(\hat{X} \leqslant x) = \sum_{i:x_i \leqslant x} \frac{1}{n} = \frac{\mu_n(x)}{n}.$$

Отсюда видно, что в качестве оценки для функции распределения $F(x)$ наблюдаемой случайной величины X разумно рассматривать функцию распределения $\hat{F}(x)$ вспомогательной случайной величины $\hat{X}$. Назовем $\hat{F}(x)$ *выборочной функцией распределения* случайной величины X. При достаточно больших n

$$F(x) \approx \hat{F}(x).$$

Сказанное наводит на мысль использовать в качестве оценки той или иной числовой характеристики случайной величины X соответствующую числовую характеристику вспомогательной случайной величины $\hat{X}$. В этом состоит так называемый *выборочный метод*. Например, в качестве оценки для математического ожидания MX следует взять

$$M\hat{X} = \sum_{i=1}^{n} x_i \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Величина $\sum_{i=1}^{n} x_i/n$ называется *выборочным средним* случайной величины X и обозначается $\bar{x}$. Итак, хорошей оценкой для MX является выборочное среднее $\bar{x}$, т. е. при достаточно больших n

$$MX \approx \bar{x}$$

(вспомните, что точно такой же результат был получен непосредственно из усиленного закона больших чисел в начале лекции).

В качестве оценки для дисперсии DX возьмем

$$D\hat{X} = M(\hat{X} - M\hat{X})^2 = M(\hat{X} - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \frac{1}{n}.$$

При достаточно больших n

$$DX \approx \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Иногда вместо указанной оценки для дисперсии используют величину

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2,$$

которую будем называть *выборочной дисперсией*.

Аналогично хорошей оценкой для начального момента $M(X^k)$, $k = 1, 2, \ldots$, является *выборочный начальный момент*, равный

$$\frac{1}{n}\sum_{i=1}^{n} x_i^k;$$

а для центрального момента $M(X - MX)^k$, $k = 2, 3, \ldots$, хорошей оценкой является *выборочный центральный момент*:

$$\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^k.$$

§ 11.3. СВОЙСТВА ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК

Пусть θ — некоторая числовая характеристика наблюдаемой случайной величины X, однозначно определяемая по ее функции распределения (в качестве θ можно рассматривать MX или DX; θ может быть одним из параметров, задающих распределение X). Из предыдущего раздела понятно, что при оценивании θ по выборке $x_1, x_2, \ldots, x_n$ находится некоторая числовая функция $\hat{\theta}_n = \hat{\theta}_n(x_1, x_2, \ldots, x_n)$ от переменных $x_1, x_2, \ldots, x_n$.

Ясно, что $\hat{\theta}_n$ является случайной величиной, т. е. функцией от элементарного исхода w.

Чтобы более четко характеризовать качество оценки $\hat{\theta}_n$, введем следующие определения.

Определение 11.1. Оценка $\hat{\theta}_n$ является *состоятельной*, если почти наверное

$$\lim_{n\to\infty} \hat{\theta}_n = \theta.$$

Определение 11.2. Оценка $\hat{\theta}_n$ является *несмещенной*, если при любом n

$$M\hat{\theta}_n = \theta.$$

Утверждение 11.1. Выборочное среднее $\bar{x}$ является состоятельной и несмещенной оценкой математического ожидания MX.

Доказательство. По усиленному закону больших чисел почти наверное

$$\lim_{n\to\infty} \bar{x} = \lim_{n\to\infty} \frac{x_1 + x_2 + \ldots + x_n}{n} = MX,$$

но это и означает состоятельность оценки $\bar{x}$. Далее, если смотреть на $x_1, x_2, \ldots, x_n$ как на независимые случайные величины, распределенные как X, то получаем, что

$$M\bar{x} = M\frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}M(x_1 + x_2 + \ldots + x_n) =$$

$$= \frac{1}{n}(Mx_1 + Mx_2 + \ldots + Mx_n) = \frac{nMX}{n} = MX,$$

но это означает несмещенность оценки $\bar{x}$.

Утверждение 11.2. Выборочная дисперсия s^2 является состоятельной и несмещенной оценкой дисперсии DX.

Доказательство. Сначала заметим, что для произвольного числа a

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}[(x_i - a) - (\bar{x} - a)]^2 = \sum_{i=1}^{n}(x_i - a)^2 -$$

$$-2\sum_{i=1}^{n}(x_i - a)(\bar{x} - a) + n(\bar{x} - a)^2 = \sum_{i=1}^{n}(x_i - a)^2 - 2n(\bar{x} - a)^2 +$$

$$+ n(\bar{x} - a)^2 = \sum_{i=1}^{n}(x_i - a)^2 - n(\bar{x} - a)^2. \qquad (11.1)$$

Вспоминая определение s^2 и применяя (11.1) при $a = 0$, получаем, что

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) =$$

$$= \frac{n}{n-1}\left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2\right). \qquad (11.2)$$

Из независимости случайных величин $x_1, x_2, \ldots, x_n$ следует независимость случайных величин $x_1^2, x_2^2, \ldots, x_n^2$, поэтому по усиленному ЗБЧ почти наверное

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n} x_i^2}{n} = M(X^2), \ \lim_{n\to\infty}\bar{x} = MX.$$

Откуда ввиду (11.2) находим, что почти наверное

$$\lim_{n\to\infty} s^2 = \lim_{n\to\infty}\frac{n}{n-1}\left(\lim_{n\to\infty}\frac{\sum_{i=1}^{n} x_i^2}{n} - \lim_{n\to\infty}\bar{x}^2\right) =$$

$$= M(X^2) - (MX)^2 = DX,$$

что означает состоятельность оценки s^2.

Для доказательства несмещенности s^2 применим (11.1) при $a = MX$, тогда

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}(x_i - a)^2 - n(\bar{x} - a)^2\right],$$

причем

$$M\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n} M(x_i - a)^2 = \sum_{i=1}^{n} Dx_i = nDX,$$

$$M(\bar{x} - a)^2 = \frac{1}{n^2}M(x_1 + x_2 + \ldots + x_n - na)^2 =$$

$$= \frac{1}{n^2}D(x_1 + x_2 + \ldots + x_n) = \frac{nDX}{n^2} = \frac{DX}{n}$$

(мы воспользовались независимостью случайных величин $x_1, x_2, \ldots, x_n$).

Следовательно,

$$Ms^2 = \frac{1}{n-1}\left[M\sum_{i=1}^{n}(x_i-a)^2 - nM(\bar{x}-a)^2\right] =$$

$$= \frac{1}{n-1}\left(nDX - n\frac{DX}{n}\right) = DX,$$

что и требовалось доказать.

Замечание 11.1. Аналогично можно доказать, что выборочные начальные моменты случайной величины X являются состоятельными и несмещенными оценками соответствующих начальных моментов. Что касается выборочных центральных моментов, то они являются состоятельными оценками соответствующих центральных моментов, но не являются несмещенными оценками этих моментов (правда, величина смещения есть $O(1/n)$ при $n \to \infty$).

§ 11.4. ГРУППИРОВАНИЕ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЯ

В некоторых случаях элементы выборки $x_1, x_2, ..., x_n$ случайной величины X группируют, т. е. разбивают область возможных значений случайной величины X на непересекающиеся интервалы $\Delta_1, \Delta_2, ..., \Delta_m$ и указывают, сколько элементов выборки оказалось в интервале Δ_1, сколько в интервале Δ_2 и т. д. Метод группирования облегчает регистрацию статистических данных, позволяет сделать статистическую информацию более обозримой. Правда, при этом теряется точность статистических оценок тех или иных числовых характеристик случайной величины X.

Пусть $t_1, t_2, ..., t_m$ — середины интервалов $\Delta_1, \Delta_2, ..., \Delta_m$ соответственно. Пусть u_j — число элементов выборки $x_1, x_2, ..., x_n$, оказавшихся в интервале Δ_j, $j = 1, 2, ..., m$. Ясно, что $\sum_{j=1}^{n} n_j = n$.

Основное предположение при статистическом анализе группированных данных состоит в следующем: все n_j элементов выборки, оказавшихся в интервале Δ_j, считаются равными t_j, $j = 1, 2, ..., m$. В итоге вместо выборки $x_1, x_2, ..., x_n$ получаем выборку, состоящую из n элементов, в которой значение t_1 встречается n_1 раз, значение t_2 встречается n_2 раз и

т. д. По этой второй выборке строятся выборочные характеристики наблюдаемой случайной величины X. Так, выборочным средним величины X является

$$\bar{x} = \frac{\sum_{j=1}^{m} t_j n_j}{n},$$

а выборочной дисперсией

$$s^2 = \frac{\sum_{j=1}^{m}(t_j - \bar{x})^2 n_j}{n}.$$

Метод группирования позволяет оценить плотность вероятностей $p(x)$ в случае непрерывной случайной величины X. Предположим, что число наблюдений n достаточно велико, а длины интервалов $\Delta_1, \Delta_2, ..., \Delta_m$ достаточно малы. Тогда из свойств плотности вероятностей следует, что $P(X \in \Delta_j) \approx \approx p(t_j)\Delta_j$, $j = 1, 2, ..., m$. Но вероятность $P(X \in \Delta_j)$ мало отличается от частоты события $\{X \in \Delta_j\}$, которая равна n_j/n.

Итак,

$$\frac{n_j}{n} \approx p(t_j)\Delta_j.$$

Следовательно,

$$p(t_j) \approx \frac{n_j}{n\Delta_j}.$$

Рассмотрим систему координат на плоскости. На оси абсцисс изобразим интервалы $\Delta_1, \Delta_2, ..., \Delta_m$. Над каждым таким интервалом Δ_j построим прямоугольник (рис. 11.1) высоты $n_j/(n\Delta_j)$.

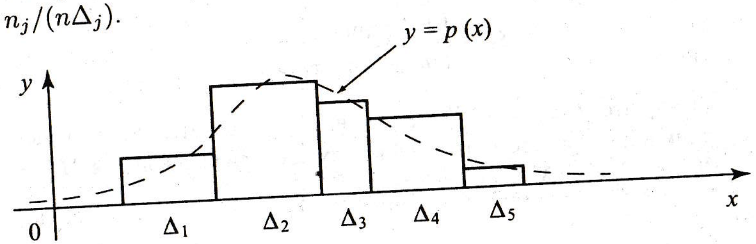


Рис. 11.1

Получающаяся при этом фигура называется *гистограммой*. В силу сказанного верхняя часть контура гистограммы дает представление о графике плотности вероятностей $y = = p(x)$.

Замечание 11.2. Как правило длины интервалов $\Delta_1, \Delta_2, \ldots, \Delta_m$ выбираются равными (пусть, например, числу l). Указанные интервалы $\Delta_1, \Delta_2, \ldots, \Delta_m$ могут не вобрать всю имеющуюся статистическую информацию. Для значений исходной выборки, лежащих левее Δ_1 (самого левого интервала), их величина считается равной $t_1 - l$. А для значений исходной выборки, лежащих правее Δ_m (самого правого интервала), их величина считается равной $t_m + l$.

Контрольные вопросы

1. В чем состоит отличие теории вероятностей и математической статистики?
2. Дайте определение выборки случайной величины.
3. Какое утверждение теории вероятностей лежит в основе классической математической статистики?
4. Докажите, что выборочная функция распределения является состоятельной несмещенной оценкой функции распределения.
5. Покажите, что выборочный второй начальный момент является несмещенной оценкой второго начального момента.
6. Объясните, почему на выборку случайной величины нужно смотреть двояко (на примере таких понятий, как состоятельность и несмещенность оценок).

Ответы

3. Усиленный закон больших чисел.
4. Введем при $i = 1, \ldots, n$ случайные величины

$$\xi_i = \begin{cases} 1, & \text{если} \quad x_i \leqslant x; \\ 0, & \text{если} \quad x_i > x. \end{cases}$$

Ясно, что $\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \xi_i$, т. е. $\hat{F}(x)$ является выборочным средним и поэтому состоятельной и несмещенной оценкой для $M\xi_i = P(x_i \leqslant x) = P(X \leqslant x) = F(x)$.

МЕТОДЫ ПОЛУЧЕНИЯ ОЦЕНОК

§ 12.1. МЕТОД МОМЕНТОВ

Пусть наблюдается случайная величина X. Предположим, что функция распределения случайной величины X принадлежит известному семейству функций распределения $F(x; \theta_1, \theta_2, \ldots, \theta_m)$, где $\theta_1, \theta_2, \ldots, \theta_m$ — параметры (например, семейство нормальных распределений $N(a, \sigma^2)$ с параметрами a, σ^2). Задача состоит в нахождении хороших оценок $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$ истинных значений параметров $\theta_1, \theta_2, \ldots, \theta_m$ по выборке $(x_1, x_2, \ldots, x_n)$ случайной величины X.

Рассмотрим метод моментов, предложенный К. Пирсоном. По выборке $(x_1, x_2, \ldots, x_n)$ найдем m выборочных начальных или центральных моментов случайной величины X. Пусть, например, нашли m выборочных начальных моментов

$$\hat{a}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \quad k = 1, 2, \ldots, m.$$

По функции распределения $F(x; \theta_1, \theta_2, \ldots, \theta_m)$ найдем для произвольных $\theta_1, \theta_2, \ldots, \theta_m$ m начальных моментов $a_k(\theta_1, \theta_2, \ldots, \theta_m)$, $k = 1, 2, \ldots, m$. Если X — непрерывная случайная величина, то вместо функции распределения $F(x; \theta_1, \theta_2, \ldots, \theta_m)$ удобно использовать плотность вероятностей $p(x; \theta_1, \theta_2, \ldots, \theta_m)$. Тогда начальные моменты находятся по формуле

$$a_k(\theta_1, \theta_2, \ldots, \theta_m) = \int\limits_{-\infty}^{+\infty} x^k p(x; \theta_1, \theta_2, \ldots, \theta_m)\,dx, \quad k = 1, 2, \ldots, m.$$

Как известно (см. лекцию № 11), $\hat{a}_k$ является состоятельной оценкой $a_k(\theta_1^*, \theta_2^*, ..., \theta_m^*)$, где $\theta_1^*, \theta_2^*, ..., \theta_m^*$ — истинные значения параметров $\theta_1, \theta_2, ..., \theta_m$. Это означает, что при достаточно больших n

$$a_k(\theta_1^*, \theta_2^*, ..., \theta_m^*) \approx \hat{a}_k, \quad k = 1, 2, ..., m.$$

Оценки по методу моментов находятся из предположения, что вместо приближенного равенства имеет место точное равенство, т. е.

$$\begin{cases} a_1(\theta_1, \theta_2, ..., \theta_m) = \hat{a}_1, \\ a_2(\theta_1, \theta_2, ..., \theta_m) = \hat{a}_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_m(\theta_1, \theta_2, ..., \theta_m) = \hat{a}_m. \end{cases} \tag{12.1}$$

Получаем систему из m уравнений с m неизвестными $\theta_1, \theta_2, ..., \theta_m$. Обозначим ее решение $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m$. Это и есть оценки по методу моментов для истинных значений параметров $\theta_1, \theta_2, ..., \theta_m$.

Как правило, система (12.1) разрешается однозначно и ее решение непрерывно зависит от $\hat{a}_1, \hat{a}_2, ..., \hat{a}_m$. Поскольку почти наверное

$$\lim_{n\to\infty} \hat{a}_k = a_k(\theta_1^*, \theta_2^*, ..., \theta_m^*)$$

и при замене в системе (12.1) правых частей $\hat{a}_k$ на $a_k(\theta_1^*, \theta_2^*, ..., \theta_m^*)$ ее решение будет $\theta_1^*, \theta_2^*, ..., \theta_m^*$, то почти наверное

$$\lim_{n\to\infty} \hat{\theta}_k = \theta_k^*, \quad k = 1, 2, ..., m.$$

Но это означает состоятельность оценок, полученных по методу моментов.

ПРИМЕР 12.1. Пусть распределение случайной величины X принадлежит семейству равномерных распределений $U(\theta_1, \theta_2)$, где $\theta_1 < \theta_2$. По выборке $(x_1, x_2, ..., x_n)$ случайной величины X, используя метод моментов, найти оценки истинных значений параметров θ_1, θ_2.

Поскольку у нас два неизвестных параметра θ_1 и θ_2, рассмотрим два выборочных момента: выборочное среднее $\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$ и выборочную дисперсию $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - x)^2$.

Ранее было установлено (см. пример 6.5 лекции № 6), что если $X \sim U(\theta_1, \theta_2)$, то

$$MX = \frac{\theta_1 + \theta_2}{2}, \quad DX = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Составим систему для нахождения оценок по методу моментов:

$$\begin{cases} \dfrac{\theta_1 + \theta_2}{2} = \bar{x}, \\ \dfrac{(\theta_2 - \theta_1)^2}{12} = s^2 \end{cases} \Leftrightarrow \begin{cases} \dfrac{\theta_1 + \theta_2}{2} = \bar{x}, \\ \dfrac{\theta_2 - \theta_1}{2} = \sqrt{3}\, s. \end{cases}$$

Откуда легко находится решение

$$\hat{\theta}_1 = \bar{x} - \sqrt{3}\, s, \quad \hat{\theta}_2 = \bar{x} + \sqrt{3}\, s.$$

Это и есть оценки для истинных значений параметров θ_1, θ_2, полученные по методу моментов.

Покажем их состоятельность. Из лекции № 11 известно, что почти наверное

$$\lim_{n\to\infty} \bar{x} = MX, \quad \lim_{n\to\infty} s^2 = DX.$$

Если истинные значения параметров θ_1 и θ_2 есть θ_1^* и θ_2^*, то $MX = \dfrac{\theta_1^* + \theta_2^*}{2}$, $DX = \dfrac{(\theta_2^* - \theta_1^*)^2}{12}$. Следовательно, почти наверное

$$\lim_{n\to\infty} \hat{\theta}_1 = \lim_{n\to\infty} (\bar{x} - \sqrt{3}\, s) = MX - \sqrt{3DX} = \\ = \frac{\theta_1^* + \theta_2^*}{2} - \sqrt{3}\,\frac{\theta_2^* - \theta_1^*}{2\sqrt{3}} = \theta_1^*.$$

Аналогично показывается состоятельность оценки $\hat{\theta}_2$.

§ 12.2. МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Пусть наблюдается случайная величина X. Известно, что распределение X принадлежит семейству распределений $F(x; \bar{\theta})$, где $\bar{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ — вектор параметров. Рассмотрим предложенный Р. Фишером метод максимального правдоподобия для нахождения оценок $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m$ истинных значений параметров $\theta_1, \theta_2, ..., \theta_m$ по выборке $(x_1, x_2, ..., x_n)$ случайной величины X.

а) *Дискретный случай.* Пусть X — дискретная случайная величина, принимающая значения $a_1, a_2, ..., a_r$. Знание функции распределения $F(x, \bar{\theta})$ равносильно знанию вероятностей $p(a_1; \bar{\theta})$, $p(a_2; \bar{\theta})$, ..., $p(a_r; \bar{\theta})$, с которыми эти значения принимает случайная величина. Тогда вероятность появления именно данной выборки $(x_1, x_2, ..., x_n)$ равна $p(x_1; \bar{\theta}) \cdot p(x_2; \bar{\theta}) \times ... \times p(x_n; \bar{\theta})$.

Назовем *функцией правдоподобия* следующую функцию от переменных $\theta_1, \theta_2, ..., \theta_m$:

$$L(\theta_1, \theta_2, ..., \theta_m) = p(x_1; \bar{\theta}) p(x_2; \bar{\theta}) \cdot ... \cdot p(x_n; \bar{\theta}).$$

Оценкой максимального правдоподобия $(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m)$ истинного значения вектора параметров $\bar{\theta}$ назовем такое значение $\bar{\theta}$, при котором функция правдоподобия достигает своего наибольшего значения.

Замечание 12.1. Точно так же в реальной жизни при нахождении виновника того или иного происшествия из возможных кандидатур выбирается та, для которой осуществленное деяние является наиболее типичным.

Часто вместо того, чтобы максимизировать функцию $L(\bar{\theta})$, максимизируют функцию $\ln L(\bar{\theta})$, что ввиду монотонности функции $y = \ln x$ является равносильным действием. А для нахождения максимума функции $\ln L(\bar{\theta})$, как известно, надо рассмотреть систему уравнений

$$\frac{\partial \ln L(\theta_1, \theta_2, ..., \theta_m)}{\partial \theta_k} = 0, \quad k = 1, 2, ..., m,$$

которая называется *системой уравнений правдоподобия.*

ПРИМЕР 12.2. Пусть случайная величина X имеет распределение Пуассона с неизвестным параметром a. Требуется по выборке $(x_1, x_2, ..., x_n)$ случайной величины X найти оценку максимального правдоподобия параметра a.

Случайная величина X принимает целые неотрицательные значения k с вероятностью

$$p(k; a) = \frac{a^k e^{-a}}{k!}.$$

Поэтому функция правдоподобия имеет вид

$$L(a) = \prod_{i=1}^{n} \frac{a^{x_i} e^{-a}}{x_i!} = \frac{a^{\sum_{i=1}^{n} x_i} e^{-na}}{x_1! x_2! \cdot ... \cdot x_n!}.$$

Следовательно,

$$\ln L(a) = \sum_{i=1}^{n} x_i \ln a - na - \ln(x_1! x_2! \cdot ... \cdot x_n!).$$

Составим уравнение правдоподобия:

$$\frac{d \ln L(a)}{d a} = 0 \iff \frac{\sum_{i=1}^{n} x_i}{a} - n = 0.$$

Отсюда находим оценку максимального правдоподобия

$$\hat{a} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}.$$

б) *Непрерывный случай.* Пусть X — непрерывная случайная величина. Вместо функции распределения $F(x, \bar{\theta})$ удобно рассматривать плотность вероятностей $p(x; \bar{\theta})$. Предположим, что случайная величина X измеряется прибором с ценой деления $\triangle x$. Тогда вместо непрерывной случайной величины X получаем дискретную случайную величину, а вместо выборки $(x_1, x_2, ..., x_n)$ — мало отличающуюся от нее (если $\triangle x$ мало) выборку $(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$. Вероятность получения именно такой выборки для полученной дискретной случайной величины приблизительно равна (вспомните свойства плотности вероятностей):

$$p(\tilde{x}_1; \bar{\theta}) \triangle x p(\tilde{x}_2; \bar{\theta}) \triangle x \cdot ... \cdot p(\tilde{x}_n; \bar{\theta}) \triangle x = \prod_{i=1}^{n} p(\tilde{x}_i; \bar{\theta}) (\triangle x)^n.$$

Чтобы максимизировать эту функцию по $\bar{\theta}$, надо максимизировать функцию $\prod_{i=1}^{n} p(\tilde{x}_i; \bar{\theta})$. Устремляя $\triangle x$ к 0, приходим к необходимости максимизировать по $\bar{\theta}$ функцию

$$L(\theta_1, \theta_2, ..., \theta_m) = \prod_{i=1}^{n} p(x_i; \bar{\theta}),$$

которая называется *функцией правдоподобия.*

Оценкой максимального правдоподобия $(\hat\theta_1, \hat\theta_2, ..., \hat\theta_m)$ истинного значения вектора параметров $\bar\theta$ назовем такое значение $\bar\theta$, при котором функция правдоподобия достигает своего наибольшего значения. Как и в дискретном случае, оценка максимального правдоподобия находится из *системы уравнений правдоподобия*

$$\frac{\partial \ln L(\theta_1, \theta_2, ..., \theta_m)}{\partial \theta_k} = 0, \quad k = 1, 2, ..., m.$$

ПРИМЕР 12.3. Пусть случайная величина X имеет нормальное распределение с неизвестными параметрами a и σ^2. По выборке $(x_1, x_2, ..., x_n)$ случайной величины X найти оценки максимального правдоподобия этих параметров.

Положим для удобства $\sigma^2 = b$. Тогда плотность вероятностей случайной величины X имеет вид

$$p(x; a, b) = \frac{1}{\sqrt{2\pi b}} e^{-(x-a)^2/(2b)}, \quad x \in R.$$

Поэтому функция правдоподобия определяется соотношением

$$L(a, b) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi b}} e^{-(x_i-a)^2/(2b)} = \frac{1}{(2\pi b)^{n/2}} e^{-\sum\limits_{i=1}^{n}(x_i-a)^2/(2b)}.$$

Следовательно,

$$\ln L(a, b) = -\frac{n}{2}\ln(2\pi b) - \frac{\sum\limits_{i=1}^{n}(x_i - a)^2}{2b}.$$

Система уравнений правдоподобия имеет вид

$$\begin{cases} \dfrac{\partial \ln L(a,b)}{\partial a} = 0, \\ \dfrac{\partial \ln L(a,b)}{\partial b} = 0 \end{cases} \Leftrightarrow \begin{cases} \dfrac{\sum\limits_{i=1}^{n}(x_i - a)}{b} = 0, \\ \dfrac{-n}{2b} + \dfrac{\sum\limits_{i=1}^{n}(x_i - a)^2}{2b^2} = 0, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \sum_{i=1}^{n}(x_i - na) = 0, \\ b = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - a)^2. \end{cases}$$

Отсюда находим оценки максимального правдоподобия

$$\hat a = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar x, \qquad \hat b = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar x)^2.$$

Поскольку $MX = a$, $DX = \sigma^2$, то найденные оценки совпадают с теми, которые дает выборочный метод.

Рассмотрим случай, когда случайная величина X является непрерывной и ее плотность вероятностей принадлежит семейству положительных плотностей вероятностей $p(x;\theta)$, где θ — одномерный параметр. Объясним, почему оценка максимального правдоподобия $\hat\theta$ близка к истинному значению θ^* параметра θ (хотя и не является состоятельной в принятом у нас смысле). Оценка максимального правдоподобия максимизирует по θ функцию

$$\ln L(\theta) = \sum_{i=1}^{n} \ln p(x_i; \theta),$$

а следовательно, и функцию

$$\frac{1}{n}\sum_{i=1}^{n} \ln p(x_i; \theta).$$

Случайные величины $x_1, x_2, ..., x_n$ независимы и одинаково распределены (как случайная величина X). Отсюда следует, что независимы и одинаково распределены (как случайная величина $\ln p(X;\theta)$) случайные величины $\ln p(x_1;\theta)$, $\ln p(x_2;\theta)$, ..., $\ln p(x_n;\theta)$. Следовательно, по усиленному закону больших чисел при достаточно больших n

$$\frac{1}{n}\sum_{i=1}^{n} \ln p(x_i; \theta) \approx M \ln p(X; \theta).$$

Из этого соотношения вытекает, что близки друг к другу те значения θ, при которых достигаются наибольшие значения левой и правой частей. Но левая часть достигает наибольшего значения при $\theta = \hat\theta$. А правая часть (см. лемму 12.1) достигает наибольшего значения при $\theta = \theta^*$. Итак, $\hat\theta \approx \theta^*$.

Лемма 12.1. Пусть случайная величина X имеет плотность вероятностей $p(x;\theta^*)$. Тогда при любом θ

$$M\ln p(X;\theta^*) \geqslant M\ln p(X;\theta).$$

Доказательство. Требуется показать, что

$$M\ln\frac{p(X;\theta)}{p(X;\theta^*)} \leqslant 0. \tag{12.2}$$

Воспользуемся неравенством $\ln x \leqslant x-1$, справедливым при любом положительном x. Тогда

$$M\ln\frac{p(X;\theta)}{p(X;\theta^*)} \leqslant M\left(\frac{p(X;\theta)}{p(X;\theta^*)}-1\right).$$

Покажем, что правая часть равна нулю (тем самым будет доказано соотношение (12.2)). Действительно, поскольку плотность вероятностей случайной величины X равна $p(x;\theta^*)$, то

$$M\left(\frac{p(X;\theta)}{p(X;\theta^*)}-1\right) = \int\limits_{-\infty}^{+\infty}\left(\frac{p(x;\theta)}{p(x;\theta^*)}-1\right)p(x;\theta^*)dx =$$

$$= \int\limits_{-\infty}^{+\infty}p(x;\theta)dx - \int\limits_{-\infty}^{+\infty}p(x;\theta^*)dx = 1-1=0$$

(воспользовались тем, что интеграл от плотности вероятностей по всей числовой оси равен единице). Лемма доказана.

Контрольные вопросы

1. Какое предположение о распределении наблюдаемой случайной величины делается при применении методов моментов и максимального правдоподобия?

2. Перечислите условия, при которых оценка по методу моментов является состоятельной.

3. Каков вероятностный смысл функции правдоподобия?

4. Объясните, чем вызван переход от функции правдоподобия к ее логарифму.

5. Докажите неравенство $\ln x \leqslant x-1$, $x\in(0,+\infty)$, используемое при доказательстве леммы 12.1.

Ответы

4. Это вызвано тем, что многие плотности вероятностей выражаются с помощью показательной функции.

5. Функция $f(x)=x-1-\ln x$ убывает при $x\in(0,1]$ и возрастает при $x\in[1,+\infty)$, поэтому достигает своего наименьшего значения при $x=1$, причем $f(1)=0$.

Лекция 13

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

§ 13.1. ОПРЕДЕЛЕНИЕ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА. ПРИМЕР ДЛЯ БОЛЬШОЙ ВЫБОРКИ

До сих пор при нахождении оценок числовых характеристик наблюдаемой случайной величины, как правило, устанавливалось, что при большом объеме выборки оценка мало отличается от оцениваемой числовой характеристики. При этом основным инструментом для установления таких фактов являлся усиленный закон больших чисел. Но ведь можно поставить вопрос о точности находимых оценок. И здесь на помощь приходит центральная предельная теорема.

Пусть $(x_1, x_2, ..., x_n)$ — выборка случайной величины X. Чтобы говорить о точности оценки той или иной числовой характеристики θ случайной величины X, необходимо вместо одной функции от выборки рассмотреть две числовые функции $\hat{\theta}_1 = \hat{\theta}_1(x_1, x_2, ..., x_n)$ и $\hat{\theta}_2 = \hat{\theta}_2(x_1, x_2, \; ..., x_n)$, причем $\hat{\theta}_1 < \hat{\theta}_2$. Говорят, что они образуют *доверительный интервал* $(\hat{\theta}_1, \hat{\theta}_2)$ с *доверительной вероятностью* p, если

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = p.$$

Как правило, доверительная вероятность p выбирается близкой к 1. Поэтому наличие доверительного интервала говорит о том, что с вероятностью, близкой к единице, оцениваемая числовая характеристика θ случайной величины X находится между $\hat{\theta}_1$ и $\hat{\theta}_2$ и, значит, числовая характеристика θ найдена с точностью $\hat{\theta}_2 - \hat{\theta}_1$. С одной стороны, видно, что ситуация отличается от той, которая характерна, например

для математического анализа, когда точность оценки является величиной безусловной, а не связанной с какой-либо случайностью. С другой стороны, события, вероятности которых имеют порядок 10^{-4}, можно считать крайне редкими в нашей жизни, поэтому построение доверительных интервалов с доверительной вероятностью, отличающейся от единицы на величину порядка 10^{-4}, имеет большой практический смысл.

В качестве примера построим доверительный интервал для неизвестного математического ожидания (обозначим его a) наблюдаемой случайной величины X при известной дисперсии DX (обозначим ее σ^2). Если $(x_1, x_2, ..., x_n)$ — выборка случайной величины X, то, как известно, $x_1, x_2, ..., x_n$ — независимые и одинаково распределенные (как X) случайные величины, поэтому в силу центральной предельной теоремы при $n \to \infty$

$$\frac{x_1 + x_2 + ... + x_n - na}{\sigma\sqrt{n}} \xrightarrow{D} N(0;1).$$

Разделив числитель и знаменатель левой части на n, получаем, что при $n \to \infty$

$$\sqrt{n}\frac{\bar{x} - a}{\sigma} \xrightarrow{D} N(0;1). \tag{13.1}$$

Определение 13.1. Пусть функция распределения $F(x)$ некоторой случайной величины непрерывна и строго возрастает (от 0 до 1) на некотором промежутке. Тогда для любого числа $p \in (0,1)$ существует единственное решение x уравнения $F(x) = p$, которое обозначается x_p и называется *квантилью уровня* p распределения $F(x)$ (рис. 13.1).

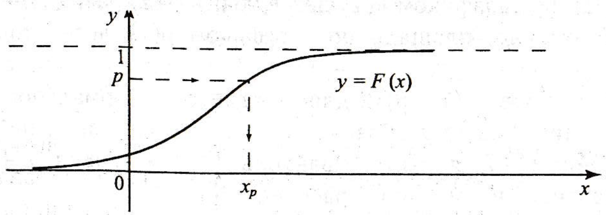


Рис. 13.1

Понятие квантили часто используется в математической статистике. По сути квантиль связана с обратной функцией к функции распределения, если эта обратная функция существует. Созданы специальные таблицы математической статистики, которые позволяют находить значение функции распределения $F(x)$ при известном x и квантили x_p при известном p.

При помощи понятия квантили можно решать не только уравнения $F(x) = p$ при известном p, но и более сложные уравнения. Пусть, например, случайная величина ξ распределена по стандартному нормальному закону, т. е. $\xi \sim N(0;1)$. Найдем такое число z, что

$$P(|\xi| < z) = p, \qquad (13.2)$$

где p — известная величина. Другими словами, нам надо найти на числовой оси такую положительную точку z, что площадь под графиком плотности вероятности $p(x)$ случайной величины ξ (рис. 13.2) между точками $(-z)$ и z равна p.

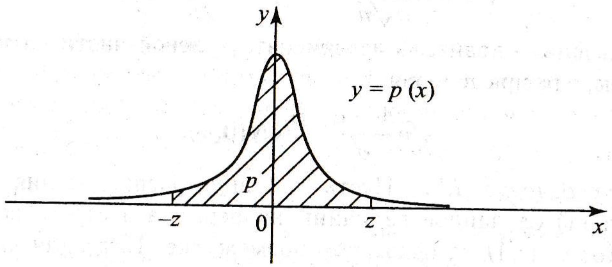


Рис. 13.2

В силу симметричности графика $p(x)$ относительно оси Oy площадь под графиком $p(x)$ левее точки $(-z)$ равна $(1-p)/2$. Следовательно, площадь под графиком $p(x)$ левее точки z равна $p + (1-p)/2 = (1+p)/2$. Другими словами, z есть квантиль уровня $(1+p)/2$ для стандартного нормального распределения, поэтому обозначим ее $z_{(p+1)/2}$. Итак, решением уравнения (13.2) является квантиль $z_{(p+1)/2}$ уровня $(p+1)/2$ стандартного нормального распределения.

Возвратимся к нашей задаче. Ввиду соотношения (13.1) при больших n можно считать, что

$$P\left(\left|\sqrt{n}\frac{\bar{x}-a}{\sigma}\right| < z\right) \approx P(|\xi| < z),$$

где $\xi \sim N(0,1)$. Следовательно, решением уравнения

$$P\left(\left|\sqrt{n}\frac{\bar{x}-a}{\sigma}\right| < z\right) = p$$

приближенно является $z_{(p+1)/2}$. Но

$$\left|\sqrt{n}\frac{\bar{x}-a}{\sigma}\right| < z \;\Leftrightarrow\; |\bar{x}-a| < \frac{z\sigma}{\sqrt{n}} \;\Leftrightarrow\; \bar{x} - \frac{z\sigma}{\sqrt{n}} < a < \bar{x} + \frac{z\sigma}{\sqrt{n}}.$$

Таким образом,

$$P\left(\bar{x} - \frac{\sigma z_{(1+p)/2}}{\sqrt{n}} < a < \bar{x} + \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}\right) = p,$$

а это означает, что искомый доверительный интервал для математического ожидания a случайной величины X с доверительной вероятностью p имеет вид

$$\left(\bar{x} - \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}, \bar{x} + \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}\right),$$

где $z_{(1+p)/2}$ — квантиль уровня $z_{(1+p)/2}$ для стандартного нормального распределения. Как мы помним, состоятельной оценкой для a является выборочное среднее $\bar{x}$. Точность этой оценки равна (с вероятностью p) $\frac{\sigma z_{(1+p)/2}}{\sqrt{n}}$. Отсюда видно, что улучшение точности в m раз достигается за счет увеличения объема выборки в m^2 раз.

Прежде, чем привести другие примеры доверительных интервалов, введем некоторые важные в математической статистике распределения.

§ 13.2. ОСНОВНЫЕ РАСПРЕДЕЛЕНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Определение 13.2. Пусть $\xi_1, \xi_2, ..., \xi_n$ — независимые случайные величины, распределенные по стандартному нормальному закону $N(0;1)$. *χ^2-распределением с n степенями свободы* называется распределение случайной величины $\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$.

Тот факт, что случайная величина η имеет такое распределение, обозначают так: $\eta \sim \chi_n^2$. График плотности вероятностей $p(x)$ такой случайной величины (при $n \geqslant 3$) изображен на рис. 13.3.

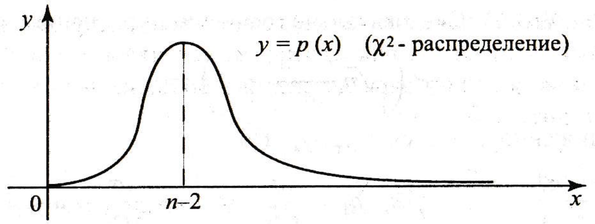


Рис. 13.3

Определение 13.3. Пусть случайные величины ξ и η независимы, причем $\xi \sim N(0;1)$, $\eta \sim \xi_n^2$. *Распределением Стьюдента (t-распределением) с n степенями свободы* называют распределение случайной величины

$$\frac{\xi}{\sqrt{\eta/n}}.$$

Если случайная величина ζ имеет такое распределение, то это обозначают так: $\zeta \sim t_n$. График плотности вероятностей $p(x)$ такой случайной величины изображен на рис. 13.4. Если $\eta \sim \chi_n^2$, то в силу закона больших чисел $\eta/n \approx 1$, поэтому t-распределение при достаточно больших n $(n \geqslant 30)$ мало отличается от стандартного нормального распределения $N(0;1)$.

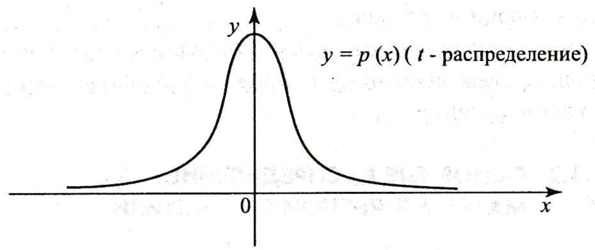


Рис. 13.4

Определение 13.4. Пусть случайные величины ξ и η независимы, причем $\xi \sim \xi_n^2$, $\eta \sim \xi_m^2$. *Распределением Фишера (F-распределением) с параметрами n, m* называется распределение случайной величины

$$\zeta = \frac{(\xi/n)}{(\eta/m)}.$$

Если случайная величина ξ имеет такое распределение, то пишут: $\zeta \sim F_{n,m}$. График плотности вероятностей $p(x)$ такой случайной величины при достаточно больших n и m изображен на рис. 13.5.

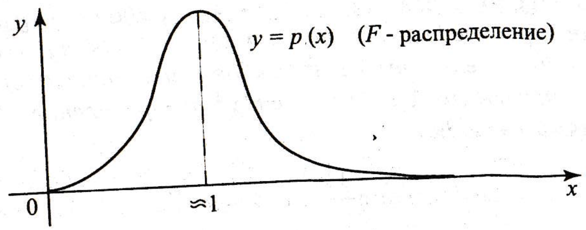


Рис. 13.5

Для всех указанных распределений созданы таблицы значений этих распределений и их квантилей.

Предположим теперь, что $(x_1, x_2, ..., x_n)$ — *нормальная выборка*, т. е. выборка случайной величины X, имеющей нормальное распределение $N(a;\sigma^2)$. Это означает, что $x_1, x_2, ...,$ x_n — независимые случайные величины, распределенные по закону $N(a;\sigma^2)$. В лекции № 9 было показано, что: а) сумма независимых случайных величин, распределенных по нормальному закону, также распределена нормально; б) случайная величина, полученная из нормально распределенной случайной величины при помощи линейного преобразования, также распределена нормально. Поэтому

$$\frac{x_1 + x_2 + ... + x_n - na}{\sigma\sqrt{n}} \sim N(0;1) \qquad (13.3)$$

(сами покажите, что математическое ожидание и дисперсия случайной величины в левой части равны 0 и 1 соответственно). Соотношение (13.3) показывает, что центральная предельная теорема в случае нормально распределенных слагаемых верна при любом числе слагаемых, а не только, когда их число стремится к бесконечности. Именно поэтому для нормально распределенной наблюдаемой случайной величины X можно строить точные доверительные интервалы для тех или иных ее числовых характеристик.

Разделив числитель и знаменатель левой части (13.3) на n получаем, что

$$\sqrt{n}\frac{\bar{x}-a}{\sigma} \sim N(0;1). \tag{13.4}$$

Р. Фишер установил следующий замечательный результат.

Теорема 13.1. *Если $(x_1, x_2, ..., x_n)$ — выборка случайной величины, распределенной по нормальному закону $N(a;\sigma^2)$, то случайные величины $\bar{x}$ и s^2 (выборочная дисперсия) независимы, причем $(n-1)s^2/\sigma^2$ имеет χ^2-распределение с $(n-1)$ степенями свободы:*

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}. \tag{13.5}$$

Следствие 13.1. В условиях теоремы 13.1

$$\sqrt{n}\frac{\bar{x}-a}{s} \sim t_{n-1}. \tag{13.6}$$

Доказательство. В силу независимости $\bar{x}$ и s^2 также независимы случайные величины $\xi = (\bar{x}-a)/(\sigma\sqrt{n})$ и $\eta = (n-1)s^2/\sigma^2$, откуда, учитывая (13.4), (13.5), получаем по определению 13.4, что случайная величина

$$\frac{\xi}{\sqrt{\eta/(n-1)}} = \sqrt{n}\,\frac{\bar{x}-a}{s}$$

имеет распределение Стьюдента с $(n-1)$ степенями свободы.

Соотношения (13.4)–(13.6) позволяют находить различные доверительные интервалы в случае нормальной выборки.

§ 13.3. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ В СЛУЧАЕ НОРМАЛЬНОЙ ВЫБОРКИ

Пусть $(x_1, x_2, ..., x_n)$ — выборка случайной величины X, распределенной по нормальному закону $N(a;\sigma^2)$.

а) *Доверительный интервал для математического ожидания MX в случае известной дисперсии DX.*

Пусть p — доверительная вероятность. Подберем число z так, чтобы

$$P\left(\left|\sqrt{n}\frac{\bar{x}-a}{\sigma}\right| < z\right) = p \tag{13.7}$$

Положим $\xi = \sqrt{n}(\bar{x}-a)/\sigma$. Ввиду соотношения (13.4) $\xi \sim N(0;1)$. Поэтому уравнение (13.7) равносильно уравнению (13.2), решением которого является число $z = z_{(1+p)/2}$ — квантиль уровня $(1+p)/2$ стандартного нормального распределения. Как уже было показано, неравенство $|\sqrt{n}(\bar{x}-a)/\sigma| < z$ эквивалентно неравенству

$$\bar{x} - \frac{\sigma z}{\sqrt{n}} < a < \bar{x} + \frac{\sigma z}{\sqrt{n}}.$$

Итак,

$$P\left(\bar{x} - \frac{\sigma z_{(1+p)/2}}{\sqrt{n}} < a < \bar{x} + \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}\right) = p.$$

Следовательно, с доверительной вероятностью p доверительный интервал для MX при известной дисперсии $DX = \sigma^2$ имеет вид

$$\left(\bar{x} - \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}, \bar{x} + \frac{\sigma z_{(1+p)/2}}{\sqrt{n}}\right).$$

б) *Доверительный интервал для математического ожидания MX в случае неизвестной дисперсии DX.*

Пусть p — доверительная вероятность. Подберем число u так, чтобы

$$P\left(\left|\sqrt{n}\frac{\bar{x}-a}{s}\right| < u\right) = p. \tag{13.8}$$

В силу соотношения (13.6) случайная величина $\eta = \sqrt{n}(\bar{x}-a)/s$ имеет t-распределение с $(n-1)$ степенями свободы. Ввиду симметричности графика плотности вероятностей $p(x)$ случайной величины η относительно оси Oy решением уравнения

$$P(|\eta| < u) = p$$

является число $u = u_{(1+p)/2}$ — квантиль уровня $(1+p)/2$ распределения Стьюдента с $(n-1)$ степенями свободы. Итак, решением уравнения (13.8) является $u = u_{(1+p)/2}$. Неравенство $|\sqrt{n}(\bar{x}-a)/s| < u$ эквивалентно неравенству

$$\bar{x} - \frac{su}{\sqrt{n}} < a < \bar{x} + \frac{su}{\sqrt{n}}.$$

Следовательно, с доверительной вероятностью p доверительный интервал для MX при неизвестной дисперсии DX имеет вид

$$\left(\bar{x} - \frac{su_{(1+p)/2}}{\sqrt{n}}, \bar{x} + \frac{su_{(1+p)/2}}{\sqrt{n}}\right),$$

где $u = u_{(1+p)/2}$ — квантиль уровня $(1+p)/2$ распределения Стьюдента с $(n-1)$ степенями свободы.

в) *Доверительный интервал для дисперсии* DX.

Подберем числа v_1 и v_2 так, чтобы

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leqslant v_1\right) = \frac{1-p}{2}, \tag{13.9}$$

$$P\left(\frac{(n-1)s^2}{\sigma^2} \geqslant v_2\right) = \frac{1-p}{2}, \tag{13.10}$$

где p — доверительная вероятность. Тогда, очевидно,

$$P\left(v_1 < \frac{(n-1)s^2}{\sigma^2} < v_2\right) = p.$$

По теореме 13.1 случайная величина $(n-1)s^2/\sigma^2$ имеет χ^2-распределение с $(n-1)$ степенями свободы, поэтому решением уравнения (13.9) является число $v_1 = v_{(1-p)/2}$ — квантиль уровня $(1-p)/2$ χ^2-распределения с $(n-1)$ степенями свободы, а решением уравнения (13.10) является число $v_2 = v_{(1+p)/2}$ — квантиль уровня $(1+p)/2$ того же распределения. Далее,

$$v_1 < \frac{(n-1)s^2}{\sigma^2} < v_2 \quad \Leftrightarrow \quad \frac{(n-1)s^2}{v_2} < \sigma^2 < \frac{(n-1)s^2}{v_1}.$$

Итак, с доверительной вероятностью p доверительный интервал для DX имеет вид

$$\left(\frac{(n-1)s^2}{v_{(1+p)/2}}, \frac{(n-1)s^2}{v_{(1-p)/2}}\right),$$

где $v_{(1-p)/2}$, $v_{(1+p)/2}$ — квантили уровней $(1-p)/2$ и $(1+p)/2$ соответственно χ^2-распределения с $(n-1)$ степенями свободы.

Контрольные вопросы

1. Дайте определение доверительного интервала.

2. Чем отличается вероятностный подход к определению точности оценки от принятого в математическом анализе?

3. Дайте определение квантили.

4. Объясните, почему решением x уравнения $p = P(\xi > x)$ является квантиль уровня $(1-p)$ распределения случайной величины ξ.

5. Дайте определение основных распределений математической статистики.

6. Объясните, почему доверительный интервал для математического ожидания при известной дисперсии находится точно, если наблюдаемая случайная величина распределена нормально, а в произвольном случае — приблизительно и лишь при больших выборках.

Ответы

4. $P(\xi \leqslant x) = 1 - P(\xi > x) = 1 - p \;\Rightarrow\; x$ — квантиль уровня $(1-p)$ распределения случайной величины ξ.

5. Это объясняется тем, что центральная предельная теорема для нормально распределенных слагаемых справедлива при любом числе слагаемых, а в произвольном случае — когда это число стремится к бесконечности.

Лекция 14

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

§ 14.1. ОСНОВНЫЕ ПОЛОЖЕНИЯ

Пусть наблюдается случайная величина X. Предположим, что относительно распределения этой случайной величины выдвинута некоторая гипотеза. Если по этой гипотезе распределение X описывается однозначно, то гипотеза называется *простой*. Если же гипотеза описывает некоторый класс распределений, то гипотеза называется *сложной*. Пусть, например, известно, что X имеет нормальное распределение. Тогда гипотеза

$$H_1: \quad MX = a, \; DX = \sigma^2$$

является простой, поскольку по математическому ожиданию и дисперсии нормальное распределение восстанавливается однозначно. Гипотеза

$$H_2: \quad MX > a$$

является сложной, поскольку даже математическое ожидание X неизвестно в случае, когда гипотеза верна.

Задача состоит в проверке некоторой гипотезы H_0, касающейся распределения случайной величины X, по ее выборке $(x_1, x_2, ..., x_n)$. Как правило, наряду с выдвигаемой гипотезой рассматривается альтернативная гипотеза H_1. В этом случае статистические выводы носят более определенный характер.

В чем состоит процедура (критерий) проверки гипотезы H_0 против альтернативы H_1? Сначала надо найти некоторую *статистику* $t(x_1, x_2, ..., x_n)$ (это числовая функция от переменных $x_1, x_2, ..., x_n$), значения которой ощутимо отличаются

в случае, когда верна гипотеза H_0, и в случае, когда верна альтернатива H_1. Эта статистика называется *статистикой критерия*. Пусть A_0 — область на числовой оси, где концентрируются значения статистики t в случае, когда верна гипотеза H_0, а A_1 — область, где концентрируются значения t в случае, когда верна альтернатива H_1. Было бы здорово, если бы области A_0 и A_1 не пересекались, а их объединение составило всю числовую прямую. Тогда при любом значении выборки $(x_1, x_2, ..., x_n)$ можно было вынести безошибочное решение относительно того, верна гипотеза H_0 или нет. К сожалению, на практике не все так просто. С одной стороны, статистика t и при гипотезе H_0, и при альтернативе H_1 может принимать практически любые значения, поэтому построить безошибочный критерий проверки гипотезы H_0 против альтернативы H_1 невозможно.

С другой стороны, с достаточно большой вероятностью статистика t в случае, когда верна гипотеза H_0, принимает значения в достаточно узкой области A_0, а в случае, когда верна альтернатива, в достаточно узкой области A_1. Помимо A_0 и A_1 есть еще область, куда значения t попадают с «неохотой» и в случае справедливости гипотезы, и в случае справедливости альтернативы. Предположим, что H_0 и H_1 — простые гипотезы и $p_0(x)$, $p_1(x)$ — плотности вероятностей статистики $t(x_1, x_2, ..., x_n)$ в случае справедливости H_0, H_1. На рис. 14.1 заштрихованы области A_0 и A_1, где концентрируются значения статистики t в случае справедливости H_0 и H_1.

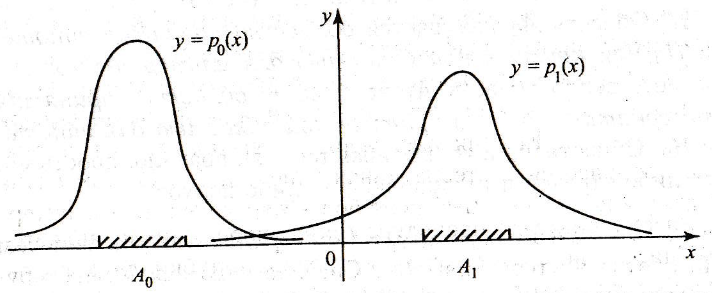


Рис. 14.1

При попадании значения статистики t в оставшуюся часть прямой вынести определенное решение сложно. Но *основное требование*, которое предъявляется к проверке статистических гипотез, состоит в том, чтобы решение относительно справедливости H_0 или H_1 было вынесено и это решение носило однозначный характер. Поэтому следует расширить области A_0 и A_1 так, чтобы, во-первых, они не пересекались и, во-вторых, составляли вместе всю числовую прямую. Для рассмотренного частного случая это расширение выглядит так, как показано на рис. 14.2 (область A_0 заштрихована).

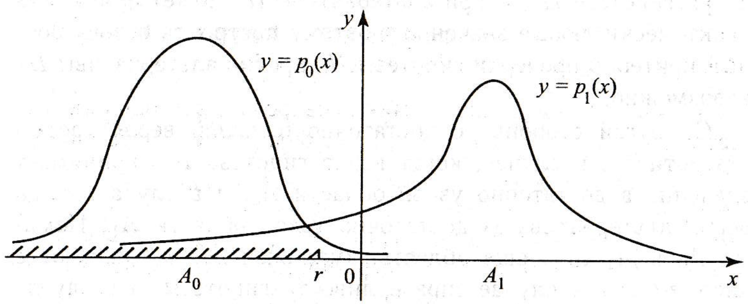


Рис. 14.2

Итак, после выбора статистики критерия t вся числовая прямая разбивается на две части A_0 и A_1 и принимается гипотеза H_0, если $t(x_1, x_2, ..., x_n) \in A_0$, или принимается альтернатива H_1 (отклоняется гипотеза H_0), если $t(x_1, x_2, ..., x_n) \in$ $\in A_1$. Область A_0 называется *областью принятия гипотезы* H_0 или *критической областью для альтернативы* H_1, соответственно область A_1 называется *областью принятия альтернативы* H_1 или *критической областью для гипотезы* H_0. Области A_0 и A_1 выбираются, как правило, простыми. Это либо отрезки и их дополнения, либо полуоси.

Рассматриваемая процедура принятия решения не является безошибочной. Более точно, возможны ошибки двух родов. *Ошибка первого рода* состоит в том, что гипотеза H_0 будет отклонена, хотя на самом деле она верна. *Ошибка второго рода* состоит в том, что альтернатива H_1 будет отклонена, хотя на самом деле она верна (гипотеза H_0 принимается, хотя она неверна). Указанным ошибкам соответствуют вероятности их совершения. Ошибка первого рода совершается с вероятностью $P_{H_0}\ (t \in A_1)$, где значок H_0 рядом с вероятностью показывает, что эта вероятность вычисляется в предположении, что верна гипотеза H_0 (в случае, когда H_0 — сложная гипотеза будет не одна, а множество различных вероятностей, соответствующих различным распределениям, описываемым гипотезой H_0). Ошибка второго рода совершается с вероятностью $P_{H_1}\ (t \in A_0)$, где значок H_1 рядом с вероятностью показывает, что она вычисляется в предположении, что верна альтернатива H_1. Иногда под ошибками первого и второго рода понимают именно эти вероятности $P_{H_0}\ (t \in A_1)$ и $P_{H_1}\ (t \in A_0)$.

Естественно желание — минимизировать эти ошибки. Как нетрудно понять из рис. 14.2, если будем уменьшать ошибку первого рода (площадь под графиком $y = p_0(x)$ правее *порогового* значения r), сдвигая r вправо, то возрастет ошибка второго рода (площадь под графиком $y = p_1(x)$ левее порогового значения r). Поэтому ограничиваются тем, что задают так называемый *уровень значимости* ε (положительное число, близкое к нулю) и предполагают, что ошибка первого рода не превосходит ε, т. е.

$$P_{H_0}\ (t \in A_1) \leqslant \varepsilon. \qquad (14.1)$$

При этом ошибку второго рода стараются сделать минимальной. А это в силу сказанного означает, что на самом деле вместо неравенства необходимо в случае простой гипотезы H_0 рассматривать равенство

$$P_{H_0}\ (t \in A_1) = \varepsilon. \qquad (14.2)$$

Именно это соотношение позволяет находить пороговые значения, разделяющие критические области A_0 и A_1.

§ 14.2. ПРОВЕРКА ГИПОТЕЗЫ О ЗНАЧЕНИИ ВЕРОЯТНОСТИ УСПЕХА В СХЕМЕ БЕРНУЛЛИ

Рассмотрим схему Бернулли с вероятностью успеха в одном испытании, равной p. Пусть проведено n испытаний (n велико) и x_i — число успехов в i-м испытании, $i = 1, 2, ..., n$. Требуется по числам $x_1, x_2, ..., x_n$ проверить гипотезу

$$H_0: \ p = p_0$$

против альтернативы

$$H_1: \ p > p_0.$$

Наблюдается случайная величина X, равная числу успехов в одном испытании, и $(x_1, x_2, ..., x_n)$ — ее выборка. Ряд распределения случайной величины X имеет вид

X	0	1
P	q	p

.

При этом $MX = p$, $DX = pq$ $(q = 1 - p)$. Следовательно, в гипотезе H_0 и альтернативе H_1 идет речь о значении математического ожидания X.

Итак, следует проверить гипотезу

$$H_0: \ MX = p_0$$

против альтернативы

$$H_1: \ MX > p_0.$$

Как известно, хорошей оценкой MX является статистика $\bar{x} = (x_1 + x_2 + ... + x_n)/n$. Поскольку $MX \approx \bar{x}$, то если $\bar{x}$ близко к p_0, то надо принимать гипотезу H_0, а если $\bar{x}$ существенно больше, чем p_0, то надо принимать альтернативу H_1. Отсюда ясно, что критические области A_0 и A_1 имеют вид, изображенный на рис. 14.3 (область A_0 заштрихована).

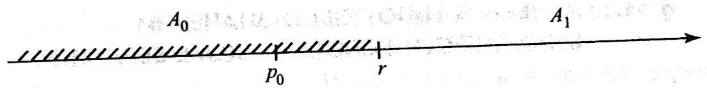


Рис. 14.3

Осталось найти пороговое значение r. Соотношение (14.2) принимает вид

$$P_{H_0}(\bar{x} > r) = \varepsilon, \tag{14.3}$$

где ε — уровень зависимости, задаваемый изначально. Если гипотеза H_0 верна, то в силу центральной предельной теоремы при $n \to \infty$

$$\frac{x_1 + x_2 + ... + x_n - np_0}{\sqrt{np_0q_0}} \xrightarrow{D} N(0;1),$$

где $q_0 = 1 - p_0$. Откуда, после деления числителя и знаменателя левой части на n, получаем, что

$$\sqrt{n}\,\frac{\bar{x} - p_0}{\sqrt{p_0q_0}} \xrightarrow{D} N(0;1). \tag{14.4}$$

Перепишем соотношение (14.3) в следующем виде:

$$P_{H_0}\left(\sqrt{n}\,\frac{\bar{x} - p_0}{\sqrt{p_0q_0}} > \sqrt{n}\,\frac{r - p_0}{\sqrt{p_0q_0}}\right) = \varepsilon. \tag{14.5}$$

Ввиду (14.4), соотношение (14.5) при достаточно больших n по сути означает, что

$$P\left(N(0;1) > \sqrt{n}\,\frac{r - p_0}{\sqrt{p_0q_0}}\right) = \varepsilon, \tag{14.6}$$

где $N(0;1)$ здесь служит обозначением случайной величины, распределенной по стандартному нормальному закону. Решением z уравнения

$$P\left(N(0;1) > z\right) = \varepsilon$$

является квантиль $z_{1-\varepsilon}$ уровня $(1 - \varepsilon)$ стандартного нормального распределения. Поэтому из уравнения (14.6) вытекает, что

$$\sqrt{n}\,\frac{r - p_0}{\sqrt{p_0q_0}} = z_{1-\varepsilon} \ \Rightarrow \ r = p_0 + \frac{\sqrt{p_0q_0}\,z_{1-\varepsilon}}{\sqrt{n}}.$$

Итак, пороговое значение r найдено. При попадании статистики $\bar{x}$ левее точки r принимается гипотеза H_0, а в противном случае — альтернатива H_1.

Оставим гипотезу H_0 прежней, а альтернативу H_1 заменим на

$$H_2: \ p < p_0,$$

которая иначе записывается следующим образом:

$$H_2: \ MX < p_0.$$

Статистика $\bar{x}$ в случае, когда верна альтернатива H_2 существенно меньше p_0, поэтому критические области A_0 и A_2 имеют вид, изображенный на рис. 14.4.

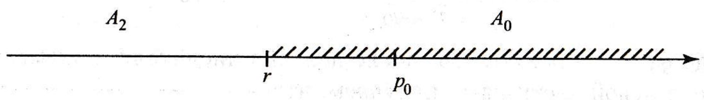


Рис. 14.4

Уравнение (14.2) для нахождения порогового значения r примет вид:

$$P_{H_0}(\bar{x} \leqslant r) = \varepsilon \ \Leftrightarrow \ P_{H_0}\left(\sqrt{n}\,\frac{\bar{x} - p_0}{\sqrt{p_0 q_0}} \leqslant \sqrt{n}\,\frac{r - p_0}{\sqrt{p_0 q_0}}\right) = \varepsilon.$$

Если верна гипотеза H_0, то при больших n статистика $\sqrt{n}\,(\bar{x} - p_0)/\sqrt{p_0 q_0}$ имеет практически стандартное нормальное распределение, поэтому последнее уравнение эквивалентно уравнению

$$P\left(N(0;1) \leqslant \sqrt{n}\,\frac{r - p_0}{\sqrt{p_0 q_0}}\right) = \varepsilon.$$

Решением z уравнения

$$P\,(N(0;1) \leqslant z) = \varepsilon$$

является квантиль z_ε уровня ε для стандартного нормального распределения. Итак,

$$\sqrt{n}\,\frac{r - p_0}{\sqrt{p_0 q_0}} = z_\varepsilon \ \Rightarrow \ r = p_0 + \frac{\sqrt{p_0 q_0}\, z_\varepsilon}{\sqrt{n}}.$$

Заметим, что при малых ε квантиль z_ε отрицательна, поэтому пороговое значение r действительно меньше p_0, как изображено на рис. 14.4. Если $\bar{x} \in A_0$, то принимаем гипотезу H_0; если же $\bar{x} \in A_2$, то принимаем альтернативу H_2.

В заключение построим критерий проверки гипотезы H_0 против альтернативы

$$H_3: \ p \neq p_0.$$

При выполнении гипотезы H_0 статистика $\bar{x}$ близка к p_0, а при выполнении альтернативы H_3 статистика $\bar{x}$ существенно отличается от p_0, поэтому критические области A_0 и A_3 будут иметь вид, изображенный на рис. 14.5.

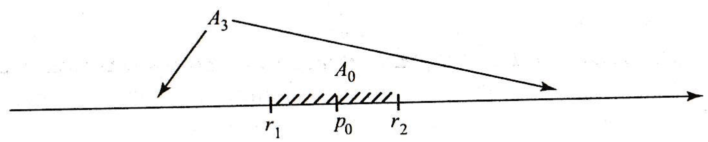


Рис. 14.5

Уравнение (14.2) для нахождения пороговых значений r_1 и r_2 примет вид:

$$P_{H_0}\,(\bar{x} \in (-\infty, r_1) \cup (r_2, +\infty)) = \varepsilon.$$

Это уравнение выполняется, если

$$P_{H_0}\,(\bar{x} \in (-\infty, r_1)) = \frac{\varepsilon}{2},$$

$$P_{H_0}\,(\bar{x} \in (r_2, +\infty)) = \frac{\varepsilon}{2}.$$

Поступая как прежде, находим, что

$$r_1 = p_0 + \frac{\sqrt{p_0 q_0}\, z_{\varepsilon/2}}{\sqrt{n}},$$

$$r_2 = p_0 + \frac{\sqrt{p_0 q_0}\, z_{1-\varepsilon/2}}{\sqrt{n}},$$

где $z_{\varepsilon/2}$, $z_{1-\varepsilon/2}$ — квантили стандартного нормального распределения уровней $\varepsilon/2$ и $1 - \varepsilon/2$ соответственно.

Контрольные вопросы

1. Что такое статистика критерия?
2. Дайте определение критических областей для гипотезы и альтернативы.
3. Дайте определение ошибок первого и второго рода.
4. Что такое уровень значимости?
5. Как находить границы критических областей для гипотезы и альтернативы?
6. Если в вашем распоряжении имеются два различных критерия для проверки одной и той же гипотезы (при одинаковом уровне значимости), какой из них вы предпочтете?

Ответы

6. Тот критерий, у которого ошибка второго рода меньше.

Лекция 15

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ ДЛЯ НОРМАЛЬНЫХ ВЫБОРОК

§ 15.1. СЛУЧАЙ ОДНОЙ НОРМАЛЬНОЙ ВЫБОРКИ

Предположим, что наблюдается случайная величина X, подчиненная нормальному закону $N(a;\sigma^2)$, и $(x_1, x_2, ..., x_n)$ — выборка этой случайной величины. Напомним (см. лекцию № 13) основные факты о распределении выборочного среднего $\bar{x}$ и выборочной дисперсии s^2 нормальной выборки.

Утверждение 15.1. Если $(x_1, x_2, ..., x_n)$ — выборка случайной величины X, имеющей нормальное распределение $N(a;\sigma^2)$, то при любом натуральном n:

$$\sqrt{n}\,\frac{\bar{x}-a}{\sigma} \sim N(0;1), \tag{15.1}$$

$$(n-1)\,\frac{s^2}{\sigma^2} \sim \chi^2_{n-1}, \tag{15.2}$$

$$\sqrt{n}\,\frac{\bar{x}-a}{s} \sim t_{n-1}. \tag{15.3}$$

Это утверждение позволяет осуществлять проверку различных статистических гипотез при любом объеме нормальной выборки (не обязательно большом).

а) *Проверка гипотезы о значении математического ожидания при известной дисперсии.*

Пусть наблюдается случайная величина $X \sim N(a;\sigma^2)$, причем параметр a неизвестен, а параметр σ^2 известен. Проверим гипотезу

$$H_0: \ MX = a_0$$

Занятие 7

ЗАКОНЫ РАСПРЕДЕЛЕНИЯ И ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

С каждым случайным экспериментом связано множество его возможных исходов $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$. *Случайной величиной* называется функция $\xi = \xi(\omega)$, определенная на множестве элементарных исходов эксперимента и принимающая действительные или комплексные значения. Если множество исходов эксперимента конечно, то приведенное определение является точным. В общем случае функция $\xi(\omega)$ полагается измеримой.

Случайную величину называют *дискретной*, если она может принимать отделенные друг от друга значения с определенными вероятностями. Множество возможных значений дискретной случайной величины конечно или счетно, т. е. их можно занумеровать с помощью ряда натуральных чисел.

Случайная величина называется *непрерывной*, если ее возможные значения составляет целый интервал (конечный или бесконечный).

Случайная величина считается заданной, если указано, какие значения она может принимать и каковы вероятности этих значений. Всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями, называется *законом распределения случайной величины.*

Случайные величины будем обозначать большими латинскими буквами $(X, Y, Z, ...)$, а отдельные значения этих величин соответствующими малыми буквами $(x, y, z, ...)$.

Соответствие между возможными значениями дискретной случайной величины и вероятностями этих значений можно задать в виде формулы. Если это затруднительно, то можно просто перечислить то и другое в виде таблицы:

X	x_1	x_2	x_3	...	x_k	...
P	p_1	p_2	p_3	...	p_k	...

Эту таблицу называют рядом распределения. Здесь $p_k = P(X = x_k)$ и $\sum_k P_k = 1$.

Ряд распределения можно изобразить графически. Для этого в каждой точке x_i на горизонтальной оси откладывают вдоль вертикальной оси отрезок, равный p_i. Полученную в результате фигуру называют многоугольником распределения (рис. 7.1).

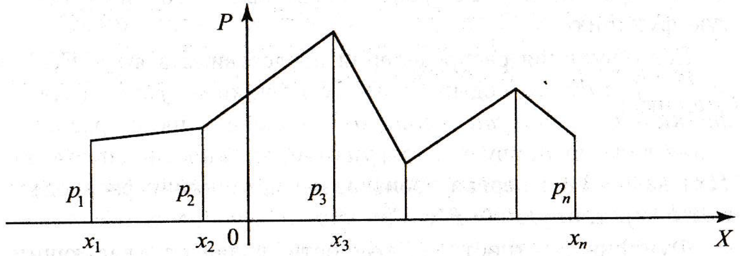


Рис. 7.1

Функцией распределения случайной величины X называют функцию

$$F(x) = P(X \leqslant x),$$

определяющую для каждого значения x вероятность того, что случайная величина X в результате опыта примет значение меньшее или равное x.

Функция распределения обладает следующими свойствами:

1) $0 \leqslant F(x) \leqslant 1, \quad -\infty < x < +\infty$;
2) $F(-\infty) = 0, \quad F(+\infty) = 1$;
3) $F(x)$ — неубывающая функция на всей числовой оси;

4) $F(x)$ непрерывна справа, т. е. $\lim F(x) = F(x_0)$ при $x \to x_0 + 0$.

Функция распределения непрерывной случайной величины является непрерывной функцией. Если x — точка непрерывности для $F(x)$, то $P(X = x) = \lim_{\triangle x \to 0}[F(x + \triangle x) - F(x)] = 0$. Это значит, что *для непрерывной случайной величины вероятность любого отдельно взятого значения равна нулю.*

Функцию распределения можно задать и для непрерывной, и для дискретной случайной величины. Для дискретной случайной величины функция распределения представляет собой, как это следует из определения, функцию накопленных вероятностей:

$$F(x) = \sum_{x_i \leqslant x} P(X = x_i),$$

где суммирование распространяется на все значения индекса i, для которых $x_i < x$. Ее график представляет собой ступенчатую функцию.

Если функция распределения представима в виде $F(x) = \int_{-\infty}^{x} f(x)dx$, то подынтегральную функцию $f(x)$ называют *функцией плотности вероятности*. Если функция распределения дифференцируема, то функцией плотности вероятности $f(x)$ называется первая производная от функции распределения $F(x)$, т. е. $f(x) = F'(x)$.

Функция плотности вероятности обладает следующими свойствами:

1) $f(x) \geqslant 0, \quad -\infty < x < +\infty$;

2) $F(x) = \int_{-\infty}^{+\infty} f(x)dx = 1$.

Последнее условие называется *условием нормировки*. Геометрически это условие означает, что площадь, заключенная между осью абсцисс и графиком функции плотности вероятности, равна единице.

Через функцию распределения и функцию плотности вероятности можно выразить вероятность попадания случайной величины в интервал (a, b):

$$P(a < X < b) = F(b) - F(a) = \int_a^b f(x)dx. \tag{7.1}$$

Для случайной величины X, распределенной по нормальному закону распределения с математическим ожиданием m и дисперсией δ^2, вероятность (7.1) определяется по формуле:

$$P(a < X < b) = \Phi_0\left(\frac{b - m}{\sigma}\right) - \Phi_0\left(\frac{a - m}{\sigma}\right), \tag{7.2}$$

где $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}dt$ — функция Лапласа, таблица которой находится в конце книги (см. приложение 1). Заметим, что функция Лапласа нечетная, т. е. $\Phi_0(-x) = -\Phi_0(x)$. Поэтому таблица дана только для положительных x. Кроме того, функция Лапласа быстро растет и при значениях $x > 5$ практически неотличима от 0,5. Поэтому таблица дана только для $x \in [0, 5]$. Из формулы (7.2) следует, что

$$P(|X - m| < \alpha) = 2\Phi_0\left(\frac{\alpha}{\sigma}\right). \tag{7.3}$$

Математическим ожиданием (или средним значением) дискретной случайной величины X называется число

$$M(X) = \sum_i x_i p_i, \tag{7.4}$$

равное сумме произведений возможных значений x_i на соответствующие им вероятности p_i. Если дискретная случайная величина имеет бесконечно много значений, то требуется абсолютная сходимость ряда (7.4). Если ряд (7.4) не сходится абсолютно, то математическое ожидание такой случайной величины не существует.

Математическим ожиданием непрерывной случайной величины, имеющей функцию плотности вероятности $f(x)$, называется число

$$M(X) = \int_{-\infty}^{+\infty} x f(x)dx,$$

если интеграл абсолютно сходится. Если интеграл не сходится абсолютно, то говорят, что математическое ожидание не существует.

Свойства математического ожидания:

1) $M(C) = C$, где C — постоянная величина;

2) $M(X \pm Y) = M(X) \pm M(Y)$;

3) $M(X \cdot Y) = M(X) \cdot M(Y)$, если случайные величины независимы;

$M(C \cdot X) = C \cdot M(X)$, где C — постоянная величина.

Дисперсией случайной величины называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X) = M[X - M(X)]^2. \tag{7.5}$$

Для вычисления дисперсии иногда удобно использовать формулу:

$$D(X) = M[X^2] - [M(X)]^2. \tag{7.6}$$

Свойства дисперсии:

1) дисперсия постоянной величины равна нулю: $D(C) = 0$;

2) постоянный множитель можно выносить за знак дисперсии с возведением в квадрат, т. е. $D(CX) = C^2 D(X)$, где C — постоянная величина;

3) если случайные величины X и Y *независимы*, то

$$D(X \pm Y) = D(X) + D(Y).$$

Дисперсия имеет размерность квадрата случайной величины. Это лишает наглядности дисперсию как числовую характеристику. Поэтому для характеристики разброса значений случайной величины используют *среднее квадратическое отклонение*, которое равно положительному значению корня квадратного из дисперсии: $\sigma(X) = \sqrt{D(X)}$.

Пусть в каждом из независимых опытов вероятность появления события равна p. Тогда k — число появлений события в n независимых опытах, имеет следующие числовые характеристики:

$$M(k) = n \cdot p, \quad D(k) = n \cdot p \cdot q.$$

ПРИМЕР 7.1. Монета подбрасывается пять раз. Написать закон распределения случайной величины, равной числу выпавших гербов минус число выпавших цифр. Построить многоугольник распределения и функцию распределения этой случайной величины.

◁ Если $N(\text{г})$ — число выпавших гербов, а $N(\text{ц})$ — число выпавших цифр, то случайная величина $X = N(\text{г}) - N(\text{ц})$ может принимать значения $x_1 = 0 - 5 = -5$, $x_2 = 1 - 4 = -3$, $x_3 = 2 - 3 = -1$, $x_4 = 3 - 2 = 1$, $x_5 = 4 - 1 = 3$, $x_6 = 5 - 0 = 5$. Подбрасывания монеты можно рассматривать как независимые испытания, вероятность выпадения герба в каждом из которых равна $p = 1/2$. Поэтому применима формула Бернулли:

$$P(X=5) = P(X=-5) = P_5(0) = C_5^0(1/2)^0(1/2)^5 = 1/32;$$

$$P(X=3) = P(X=-3) = P_5(1) = C_5^1(1/2)^1(1/2)^4 = 5/32;$$

$$P(X=1) = P(X=-1) = P_5(2) = C_5^2(1/2)^2(1/2)^3 = 10/32.$$

Зная возможные значения случайной величины и вероятности этих значений, можно написать закон ее распределения. Ряд распределения имеет вид:

X	-5	-3	-1	1	3	5
P	1/32	5/32	10/32	10/32	5/32	1/32

Для наглядности ряд распределения можно изобразить графически в виде многоугольника распределения (рис. 7.2).

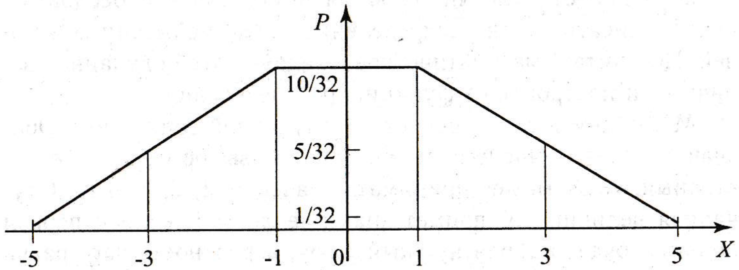


Рис. 7.2

При построении функции распределения $F(x)$ нужно для каждого x найти вероятности $P(X < x)$:

$$F(x) = P(X < x) = \begin{cases} 0 & \text{при } x \leqslant -5, \\ 1/32 & \text{при } -5 < x \leqslant -3, \\ 6/32 & \text{при } -3 < x \leqslant -1, \\ 16/32 & \text{при } -1 < x \leqslant 1, \\ 26/32 & \text{при } 1 < x \leqslant 3, \\ 31/32 & \text{при } 3 < x \leqslant 5, \\ 1 & \text{при } 5 < x. \end{cases}$$

График этой функции распределения изображен на рис. 7.3. ▷

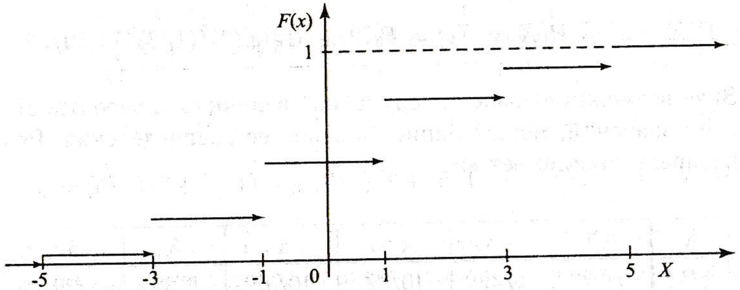


Рис. 7.3

ПРИМЕР 7.2. Некто имеет на связке пять ключей. При отмыкании замка он последовательно испытывает ключи, пока не подберет нужный. Полагая выбор ключей бесповторным, написать закон распределения числа испытанных ключей. Подсчитать математическое ожидание этой случайной величины и построить ее функцию распределения.

◁ Мы имеем дело с дискретной случайной величиной. Обозначим через X число попыток. Так как выбор ключей бесповторный, то X может принимать значения: 1, 2, 3, 4, 5. Случайная величина X примет значение $x_1 = 1$, если с первой попытки будет выбран нужный ключ, вероятность чего равна 1/5, так как выбор любого из пяти ключей равновозможен. Значение $x_2 = 2$ случайная величина примет, если при первой попытке ключ будет выбран ошибочно (вероятность чего равна 4/5) и при второй попытке будет выбран нужный ключ из оставшихся четырех (вероятность этого равна 1/4). Поэтому:

$$P(X = 2) = (4/5) \cdot (1/4) = 1/5;$$

$$P(X = 3) = (4/5) \cdot (3/4) \cdot (1/3) = 1/5;$$

$$P(X = 4) = (4/5) \cdot (3/4) \cdot (2/3) \cdot (1/2) = 1/5;$$

$$P(X = 4) = (4/5) \cdot (3/4) \cdot (2/3) \cdot (1/2) \cdot 1 = 1/5.$$

Случайная величина X имеет закон распределения

X	1	2	3	4	5
P	1/5	1/5	1/5	1/5	1/5

.

Вычислим среднее значение случайной величины X:

$$M(X) = 1 \cdot (1/5) + 2 \cdot (1/5) + 3 \cdot (1/5) + 4 \cdot (1/5) + 5 \cdot (1/5) = 3. \triangleright$$

ПРИМЕР 7.3. Вероятность попадания в цель при одном выстреле равна p. Написать закон распределения случайной величины, равной числу выстрелов до первого попадания в цель. Подсчитать математическое ожидание этой случайной величины.

◁ Случайная величина X примет значение 1, если будет попадание при первом же выстреле, вероятность чего равна p. Понадобится два выстрела ($X = 2$), если при первом выстреле будет промах (вероятность чего равна $q = 1 - p$) и при втором выстреле будет попадание (вероятность чего равна p). Поэтому в силу независимости событий $P(X = 2) = q \cdot p$. Аналогично, $P(X = 3) = q \cdot q \cdot p = q^2 \cdot p, \ldots$; $P(X = n) = q \cdot q \cdot q \cdot \ldots \cdot q \cdot p = q^{n-1} \cdot p$. Случайная величина X имеет закон распределения:

X	1	2	3	...	n	...
P	p	$q \cdot p$	$q^2 \cdot p$	...	$q^{n-1} \cdot p$	...

.

Этот закон распределения называют *геометрическим*. Вычислим среднее число выстрелов до первого попадания в цель.

$$M(X) = 1\cdot p + 2\cdot q\cdot p + 3\cdot q^2\cdot p + \ldots + n\cdot q^{n-1}\cdot p + \ldots =$$
$$= p(1 + 2q + 3q^2 + \ldots + nq^{n-1} + \ldots).$$

Для вычисления суммы полученного ряда перегруппируем слагаемые следующим образом:

$$M(X) = p\left[(1+q+q^2+\ldots+q^{n-1}+\ldots)+(q+q^2+\ldots+q^{n-1}+\ldots)+\right.$$
$$\left.+(q^2+q^3+\ldots+q^{n-1}+\ldots)+\ldots\right] = p\left[(1+q+q^2+\ldots+q^{n-1}+\ldots)+\right.$$
$$\left.+q(1+q+q^2+\ldots+q^{n-1}+\ldots)+q^2(1+q+q^2+\ldots+q^{n-1}+\ldots)+\ldots\right] =$$
$$= p(1+q+q^2+\ldots+q^{n-1}+\ldots)(1+q+q^2+\ldots+q^{n-1}+\ldots) =$$
$$= p(1+q+q^2+\ldots+q^{n-1}+\ldots)^2.$$

В скобке стоит сумма бесконечной убывающей геометрической прогрессии

$$1+q+q^2+\ldots+q^{n-1}+\ldots = \frac{1}{1-q} = \frac{1}{p}.$$

Поэтому $M(x) = p\cdot\frac{1}{p^2} = \frac{1}{p}$. ▷

ПРИМЕР 7.4. В партии из 12 деталей содержится три детали низкого качества. Наугад выбраны четыре детали. Написать закон распределения и найти математическое ожидание числа деталей низкого качества среди выбранных.

◁ Пусть X — число деталей низкого качества среди выбранных четырех. Это дискретная случайная величина с возможными значениями от 0 до 3. Четыре детали из 12 можно выбрать $n = C_{12}^4 = 495$ равновозможными способами. Значению $X = 0$ благоприятствует $C_9^4 = 126$ способов выбора деталей. Значению $X = 1$ благоприятствует $C_3^1\cdot C_9^3 = 252$ способа. Значению $X = 2$ благоприятствует $C_3^2\cdot C_9^2 = 108$ способов, и значению $X = 3$ благоприятствует $C_3^3\cdot C_9^1 = 9$ способов. Поэтому случайная величина X имеет закон распределения:

X	0	1	2	3
P	126/495	252/495	108/495	9/495

Определим среднее число деталей низкого качества в выборке:

$$M(X) = 1\cdot 252/495 + 2\cdot 108/495 + 3\cdot 9/495 = 1. ▷$$

ПРИМЕР 7.5. Поезда в метро следуют с интервалом 2 минуты. Пассажир в случайный момент времени приходит на платформу и ожидает ближайший поезд. Найти функцию распределения времени ожидания пассажира. Найти математическое ожидание и дисперсию времени ожидания.

◁ Естественно предположить, что равновозможен приход пассажира в любой момент времени между прибытием поездов. Формально это означает, что плотность вероятности на отрезке $[0, 2]$ постоянна. А так как длина отрезка равна 2, то на этом отрезке функция плотности вероятности равна 1/2. Итак, время ожидания X имеет равномерный закон распределения с функцией плотности вероятности $f(x) = 1/2$ при $x \in [0,2]$ и $f(x) = 0$ при $x \notin [0,2]$. Тогда при $x \leqslant 0$

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} 0dx = 0;$$

при $0 < x \leqslant 2$

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{x} 1/2dx = x/2;$$

при $2 < x$

$$F(x) = \int_{-\infty}^{0} 0dx + \int_{0}^{2} 1/2dx + \int_{2}^{x} 0dx = 1.$$

Итак, $F(x) = \begin{cases} 0, & x \leqslant 0, \\ x/2, & 0 < x \leqslant 2, \\ 1, & 2 < x. \end{cases}$

Вычислим числовые характеристики случайной величины:

$$M(X) = \int_0^2 x\cdot 1/2\,dx = 1, \quad D(X) = \int_0^2 (x-1)^2\cdot 1/2\,dx = 1/3. ▷$$

ПРИМЕР 7.6. Случайная величина X имеет функцию распределения

$$F(x)=\begin{cases}0, & x\leqslant 0,\\ x^2/9, & 0<x\leqslant 3,\\ 1, & 3<x.\end{cases}$$

Найти: а) функцию плотности вероятности; б) глядя на график функции распределения (рис. 7.4), указать основные особенности случайной величины (например, интервал возможных значений, наиболее вероятные значения); в) $M(X)$, $D(X)$, $P(X<1)$, $P(1<X<2)$.

$\triangleleft$ а) Функция плотности вероятности равна первой производной от функции распределения. Поэтому

$$f(x)=F'(x)=\begin{cases}0, & x\leqslant 0,\\ 2x/9, & 0<x\leqslant 3,\\ 0, & 3<x.\end{cases}$$

б)

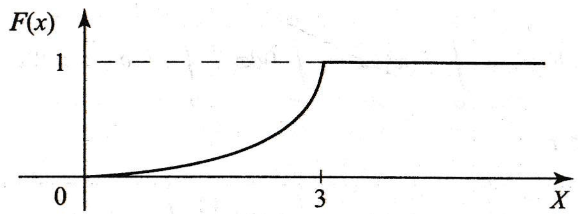


Рис. 7.4

Случайная величина X принимает значения только из промежутка $[0,3]$, причем ее значения тем вероятнее, чем ближе они к 3.

в) $M(X)=\int\limits_0^3 x(2x/9)dx=2;\quad D(X)=\int\limits_0^3 (x-2)^2(2x/9)dx=1/2;$

$P(X<1)=F(1)=1/9;$

$P(1<X<2)=F(2)-F(1)=4/9-1/9=1/3.$ $\triangleright$

ПРИМЕР 7.7. Пусть равновозможно прибытие автомобиля к перекрестку в любой момент цикла работы светофора (рис. 7.5). Найти функцию распределения времени ожидания автомобиля.

Рис. 7.5

Обозначим время ожидания у светофора через X. Это неотрицательная случайная величина. Вероятность того, что время ожидания будет меньше x, равна вероятности прибыть к светофору в момент времени из интервала (A,B). Поэтому $F(x)=P(X\leqslant x)=(x+30)/70$ при $0<x<40$ и $F(x)=1$ при $x\geqslant 40$. Функция распределения времени ожидания изображена на рис. 7.6.

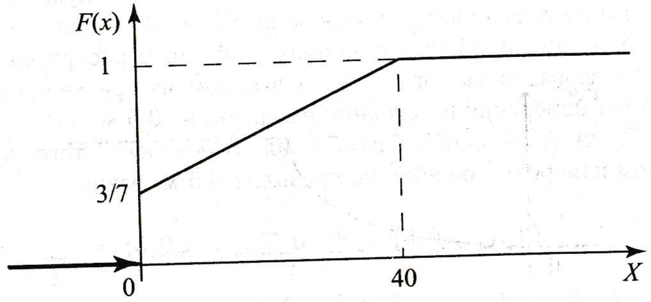


Рис. 7.6

Из графика функции распределения видно, что нулевое время ожидания, имея вероятность $3/7$, соответствует точке скачка функции, равного этой величине. $\triangleright$

ПРИМЕР 7.8. На круговом экране локатора равновозможно появление пятна в каждой точке экрана. Радиус экрана равен R. Найти закон распределения расстояния от центра экрана до пятна. Найти математическое ожидание и дисперсию этого расстояния.

$\triangleleft$ Обозначим через X расстояние от центра экрана до пятна. Это расстояние будет меньше x, если пятно попадет внутрь круга радиуса x. Вероятность этого по геометрическому определению вероятности равна отношению площади круга радиуса x к площади всего экрана локатора. Поэтому функция распределения случайной величины X имеет вид $F(x) = P(X < x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}$ при $0 < x < R$. Тогда функция плотности вероятности равна $f(x) = \frac{2x}{R^2}$ при $0 < x < R$. Математическое ожидание и дисперсия равны:

$$M(X) = \int_0^R x \cdot \frac{2x}{R^2} dx = \frac{2R}{3};$$

$$D(X) = \int_0^R \left(x - \frac{2R}{3}\right)^2 \cdot \frac{2x}{R^2} dx = \frac{R^2}{18}. \triangleright$$

ПРИМЕР 7.9. Дальномер имеет систематическую ошибку 0,1 м и среднюю квадратическую ошибку 0,4 м. Полагая, что ошибки измерений имеют нормальный закон распределения, найти вероятность того, что ни в одном из трех измерений ошибка измерения расстояния не превысит 0,5 м?

$\triangleleft$ Пусть X — ошибка измерения. Вероятность того, что в одном измерении ошибка не превысит 0,5 м, равна

$$P(|X| < 0,5) = P(-0,5 < X < 0,5) =$$

$$= \Phi_0\left(\frac{0,5-0,1}{0,4}\right) - \Phi_0\left(\frac{-0,5-0,1}{0,4}\right) = \Phi_0(1) - \Phi_0(-3,75) =$$

$$= \Phi_0(1) + \Phi_0(3,75) = 0,3413 + 0,4998 \approx 0,83. \triangleright$$

Характеристической функцией $\varphi(z)$ случайной величины X называется комплекснозначная функция, определенная при $z \in R$ соотношением

$$\varphi(z) = M\,e^{izX} = M\left[\cos(zX) + i\sin(zX)\right].$$

Если $F(x)$ — функция распределения случайной величины X, то

$$\varphi(z) = \int_{-\infty}^{+\infty} e^{izx} dF(x).$$

Существование интеграла, определяющего характеристическую функцию, вытекает из непрерывности функции e^{izx} и ее ограниченности: $|e^{izx}| \leqslant 1$.

Некоторые свойства характеристических функций:

1) $\varphi(0) = 1$, $|\varphi(z)| \leqslant 1$ для всех вещественных z;

2) если существует $M|X^n|$ — момент порядка n, то функция $\varphi(z)$ имеет n непрерывных производных и

$$\varphi^{(n)}(0) = i^n M X^n;$$

3) характеристическая функция однозначно определяет распределение случайной величины;

4) если X_1 и X_2 — независимые случайные величины, а $\varphi_1(z)$ и $\varphi_2(z)$ — их характеристические функции, то характеристическая функция суммы $X_1 + X_2$ равна произведению $\varphi_1(z) \cdot \varphi_2(z)$.

Если $X_1, X_2, ..., X_n$ — независимые случайные величины, то характеристическая функция суммы $X_1+X_2+...+X_n$ равна произведению характеристических функций слагаемых.

Пусть X — целочисленная неотрицательная случайная величина, для которой $P(X = k) = p_k$, $k \geqslant 0$. *Производящей функцией* $\psi(z)$ величины X называется функция

$$\psi(z) = M(z^X) = \sum_{k=0}^{\infty} p_k z^k, \tag{7.7}$$

определенная при комплексных z, для которых $|z| \leqslant 1$.

Заметим, что из $\psi'(z) = M(z^X) = \sum_{k=0}^{\infty} p_k k z^{k-1}$ следует, что

$$\psi'(1) = M(1^X) = \sum_{k=0}^{\infty} p_k k = M(X). \tag{7.8}$$

Так как $\psi''(z) = \sum\limits_{k=0}^{\infty} p_k k(k-1) z^{k-2}$, то

$$\psi''(1) = \sum_{k=0}^{\infty} p_k (k^2 - k) = M(X^2) - M(X).$$

Поэтому

$$D(X) = \psi''(1) + \psi'(1) - [\psi'(1)]^2. \qquad (7.9)$$

Формулы (7.8) и (7.9) позволяют вычислять математическое ожидание и дисперсию случайной величины, если известна ее производящая функция.

ПРИМЕР 7.10. Пусть $P(X = k) = pq^k$, $0 < p < 1$, $q = 1 - p$, $k = 0, 1, 2, 3, \ldots$. Найти математическое ожидание и дисперсию этой случайной величины.

◁ Найдем сначала производящую функцию. По определению (7.7)

$$\psi(z) = \sum_{k=0}^{\infty} p_k z^k = pz^0 + pq^1 z^1 + pq^2 z^2 + pq^3 z^3 + \ldots =$$

$$= p(1 + qz + q^2 z^2 + q^3 z^3 + \ldots).$$

Так как z и q по модулю меньше единицы, то в скобке находится сумма бесконечной убывающей геометрической прогрессии. Поэтому

$$\psi(z) = \frac{p}{1 - qz}.$$

Найдем первую и вторую производные производящей функции:

$$\psi'(z) = \frac{pq}{(1 - qz)^2}, \qquad \psi''(z) = \frac{2pq^2}{(1 - qz)^3}.$$

По формулам (7.8) и (7.9)

$$M(X) = \psi'(1) = \frac{pq}{(1 - q)^2} = \frac{pq}{p^2} = \frac{q}{p},$$

$$D(X) = \psi''(1) + \psi'(1) - [\psi'(1)]^2 = \frac{2pq^2}{(1 - q)^3} + \frac{q}{p} - \frac{q^2}{p^2} =$$

$$= \frac{2q^2 + pq - q^2}{p^2} = \frac{q(p + q)}{p^2} = \frac{q}{p^2}. \ ▷$$

ПРИМЕР 7.11. Случайные величины X и Y независимы и имеют пуассоновские законы распределения с параметрами λ_1 и λ_2 соответственно:

$$P(X = k) = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, \quad P(Y = k) = \frac{\lambda_2^k}{k!} e^{-\lambda_2}, \quad k = 0, 1, 2, 3, \ldots.$$

Найти закон распределения случайной величины $X + Y$.

◁ Найдем характеристические функции случайных величин X и Y:

$$\varphi_1(z) = Me^{izX} = \sum_{k=0}^{\infty} e^{izk} \frac{\lambda_1^k}{k!} e^{-\lambda_1} = e^{-\lambda_1} \sum_{k=0}^{\infty} \frac{(\lambda_1 e^{iz})^k}{k!} =$$

$$= \exp(-\lambda_1) \cdot \exp(\lambda_1 e^{iz}) = \exp[\lambda_1 (e^{iz} - 1)].$$

Аналогично

$$\varphi_2(z) = Me^{izX} = \exp[\lambda_2 (e^{iz} - 1)].$$

Сумме независимых случайных величин соответствует произведение характеристических функций слагаемых. Поэтому $X + Y$ имеет характеристическую функцию

$$\exp[\lambda_1 (e^{iz} - 1)] \cdot \exp[\lambda_2 (e^{iz} - 1)] = \exp[(\lambda_1 + \lambda_2)(e^{iz} - 1)].$$

Это означает, что $X + Y$ имеет пуассоновский закон распределения с параметром $\lambda_1 + \lambda_2$. ▷

Полученный результат известен как факт *устойчивости пуассоновского закона распределения*. Этот результат можно обобщить на сумму любого конечного числа пуассоновских случайных величин.

ПРИМЕР 7.12. Случайная величина X равномерно распределена на отрезке $[-a, a]$. Найти характеристическую функцию этой случайной величины.

◁ Так как все значения на отрезке $[-a, a]$ равновозможны, то плотность вероятности случайной величины X имеет вид $f(x) = \frac{1}{2a}$ при $x \in [-a, a]$ и $f(x) = 0$ при остальных x. Поэтому

$$\varphi(z) = \frac{1}{2a} \int_{-a}^{a} e^{izx} dx = \frac{1}{2a} \cdot \frac{e^{izb} - e^{iza}}{iz} = \frac{\sin az}{az}. \ ▷$$

ПРИМЕРЫ (для самостоятельного решения).

1. Из условий примера 7.2 написать закон распределения числа попыток и найти математическое ожидание этого числа в предположении повторного выбора ключей.

2. На электронное реле воздействует случайное напряжение, имеющее плотность вероятности $f(x) = \frac{x}{\sigma^2} \cdot \exp\left\{-\frac{x^2}{2\sigma^2}\right\}$, $x \geqslant 0$. Реле срабатывает всякий раз, когда напряжение на его входе превышает 3 В. Какова вероятность срабатывания реле?

3. В кошельке было пять монет по 10 копеек и три монеты по 50 копеек. Из кошелька вынули наугад четыре монеты. Найдите закон распределения случайной величины X, которая равна сумме вынутых копеек.

4. Время безотказной работы предохранителя имеет показательный закон распределения с функцией плотности вероятности $f(x) = 0,002e^{-0,002x}$ при $x \geqslant 0$, $f(x) = 0$ при $x < 0$. Найдите функцию распределения времени безотказной работы. Найдите вероятность того, что предохранитель безотказно проработает 1000 часов.

5. Случайная величина X равна сумме выпавших очков на двух игральных кубиках. Напишите ее закон распределения и найдите ее математическое ожидание.

6. В урне лежат два черных и три белых шара. Из этой урны вынимаются один за другим без возвращения шары до тех пор, пока не будет вынут черный шар. Найдите среднее число вынутых при этом шаров.

7. Случайная величина X имеет функцию плотности вероятности $f(x) = 0,5\sin x$ при $x \in [0, \pi]$ и $f(x) = 0$ при остальных x. Найдите: а) функцию распределения величины X; б) математическое ожидание этой величины; в) вероятность попадания в интервал $(0, \pi/3)$.

8. Монету бросают до первого выпадения герба, либо до тех пор, пока цифра не выпадет четыре раза подряд. Найдите среднее число бросков монеты.

9. Случайная величина X имеет функцию плотности вероятности $f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$ (закон распределения Коши). Какова вероятность того, что при трех независимых наблюдениях этой случайной величины будут наблюдаться значения только из интервала $(-1, 1)$?

10. Случайная величина X — погрешность измерительного прибора распределена по нормальному закону распределения с дисперсией $\sigma^2 = 25$ мВ^2. Систематическая погрешность прибора отсутствует. Найдите вероятность того, что при пяти независимых измерениях ошибка измерения хотя бы один раз превзойдет по модулю 10 мВ.

11. Взвешивание производится без систематической ошибки, а случайные ошибки подчинены нормальному закону распределения со средним квадратическим отклонением $\sigma = 20$ мг. Найдите вероятность того, что взвешивание будет произведено с ошибкой, не превосходящей по абсолютной величине 50 мг.

12. Пусть $X = 1$ с вероятностью $1/2$, $X = -1$ с вероятностью $1/2$. Найдите характеристическую функцию этой случайной величины.

13. Случайная величина X равномерно распределена на отрезке $[a, b]$. Найдите характеристическую функцию этой случайной величины.

14. Пусть случайная величина X имеет распределение Пуассона, т. е. $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, 3, \ldots$ С помощью производящей функции найдите математическое ожидание и дисперсию этой случайной величины.

Ответы:

1. $M(X) = 5$; 2. $\exp\{-9/2\sigma^2\}$;
3. $P(X = 40) = 1/14$, $P(X = 80) = 3/7$, $P(X = 120) = 3/7$, $P(X = 160) = 1/14$;
4. $F(x) = 1 - e^{-0,002x}$ при $x \geqslant 0$, $F(x) = 0$ при $x < 0$, $P = e^{-2} \approx 0,17$;
5. $M(\xi) = 7$; 6. 2;
7. а) $F(x) = 0$ при $x < 0$, $F(x) = 0,5(1 - \cos x)$ при $x \in [0, \pi]$, $F(x) = 1$ при $\pi < x$, б) $M(\xi) = \pi/2$, в) $P(0 < \xi < \pi/3) = 1/4$;
8. 15/8; 9. 1/8; 10. $\approx 0,21$; 11. $\approx 0,98$;
12. $\varphi(z) = \cos z$; 13. $\varphi(z) = \dfrac{1}{b-a} \cdot \dfrac{e^{izb} - e^{iza}}{iz}$;
14. $M(X) = D(X) = \lambda$.

Занятие 8

ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

Пусть $h(x)$ — однозначная функция. Функцией случайной величины X называется такая случайная величина $H = h(X)$, которая принимает значение $h_i = h(x_i)$ каждый раз, когда величина X принимает значение x_i.

1. Пусть X — дискретная случайная величина. Если функция $h(x)$ в области возможных значений X монотонна, то величина H примет значение $h_i = h(x_i)$ тогда и только тогда, когда $X = x_i$. Следовательно, возможными значениями H будут значения $h_i = h(x_i)$, и этим значениям соответствуют вероятности $p_i = P(H = h_i) = P(X = x_i)$.

2. Если $h(x)$ немонотонна и существует несколько значений $x_1, x_2, \ldots, x_m$, при которых $h(x) = h_i$, то

$$P[h(x) = h_i] = P[X = x_1 \text{ или } X = x_2 \text{ или } \ldots \text{ или } X = x_m] =$$

$$= \sum_{i=1}^{m} P(X = x_i).$$

Следовательно, для нахождения закона распределения случайной величины $H = h(X)$ нужно вычислить все ее значения, расположить их в порядке возрастания, отбрасывая повторяющиеся, и каждому из полученных значений h_i приписать вероятность, равную сумме вероятностей тех значений X, для которых $h(x) = h_i$.

3. Пусть функция $h(x)$ монотонна и непрерывна вместе со своей производной в области возможных значений случайной величины X. Если X имеет непрерывную функцию плотности

вероятности $f(x)$, то $g(h)$ — плотность вероятности случайной величины $H = h(X)$, определяется соотношением

$$g(h) = f[x(h)] \left| \frac{dx(h)}{dh} \right|, \tag{8.1}$$

где $x = x(h)$ обратная к $h = h(x)$ функция.

4. Пусть функция $h(x)$ немонотонна и непрерывна вместе со своей производной в области возможных значений случайной величины X, которая имеет непрерывную функцию плотности вероятности $f(x)$.

Тогда плотность вероятности $g(h)$ случайной величины $H = h(X)$ определяется соотношением

$$g(h) = f[x_1(h)] \left| \frac{dx_1(h)}{dh} \right| + f[x_2(h)] \left| \frac{dx_2(h)}{dh} \right| + \ldots$$

$$\ldots + f[x_m(h)] \left| \frac{dx_m(h)}{dh} \right|, \tag{8.2}$$

где $x_i = x_i(h)$ функции, обратные к функции $h = h(x)$ соответственно на каждом из m интервалов монотонности.

Замечание 8.1. Пусть X — некоторая случайная величина с функцией распределения $F(x)$. Для вычисления математического ожидания случайной величины $h(X)$ не обязательно предварительно находить ее функцию распределения, а можно подсчитать его непосредственно:

$$M[h(X)] = \int h(x) \cdot dF(x). \tag{8.3}$$

Для дискретной случайной величины X соответствующая формула имеет вид

$$M[h(X)] = \sum h(x_i) \cdot P(X = x_i). \tag{8.4}$$

ПРИМЕР 8.1. Дискретная случайная величина X имеет закон распределения:

X	-2	0	2	4
P	0,4	0,2	0,3	0,1

Найти закон распределения случайной величины $H = 9 - X^2$.
$\lhd$ Вероятность возможного значения $h_1 = -7$ равна вероятности события $X = 4$, т. е. равна 0,1. Вероятность возможного значения $h_2 = 5$ равна сумме вероятностей несовместных событий $X = -2$ и $X = 2$, т. е. равна $0,4 + 0,3 = 0,7$. Вероятность значения $h_3 = 9$ равна вероятности события $X = 0$, т. е. равна 0,2. Искомое распределение имеет вид:

H	-7	5	9
P	$0,1$	$0,7$	$0,2$

$\rhd$

ПРИМЕР 8.2. Случайная величина X имеет нормальный закон распределения $N(0;1)$. Найти закон распределения случайной величины $Y = aX + b$, где a и b — некоторые постоянные.
$\lhd$ Функция $y = ax+b$ монотонна на всей числовой оси и имеет обратную функцию $x = (y-b)/a$. Так как плотность вероятности случайной величины X имеет вид $f(x) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}$, то по формуле (8.1) случайная величина Y имеет функцию плотности вероятности $g(y) = \frac{1}{\sqrt{2\pi}}e^{-(y-b)^2/(2a^2)} \cdot \frac{1}{a} = N(b;a^2)$. $\rhd$

ПРИМЕР 8.3. Ошибка округления X распределена равномерно в интервале $(0;0,5)$. «Цена» ошибки Y пропорциональна квадрату ошибки, т. е. $Y = aX^2$, где a — некоторая постоянная. Найти функцию распределения Y и ее математическое ожидание.
$\lhd$ Так как ошибка округления X распределена равномерно в интервале $(0;0,5)$, то ее плотность вероятности $f(x) = \frac{1}{0,5} = 2$ при $x \in (0;0,5)$ и $f(x) = 0$ при остальных x. Возможные значения случайной величины Y заключены в $(0;0,25a)$. Функция $y = ax^2$ имеет на $(0;0,5)$ обратную функцию $x = \sqrt{\frac{y}{a}}$. Поэтому по формуле (8.1) случайная Y величина имеет функцию плотности вероятности

$$g(y) = \frac{2d\left(\sqrt{y/a}\right)}{dy} = \frac{1}{\sqrt{ya}}$$

при $y \in (0;0,25a)$ и $g(y) = 0$ при остальных значениях y. Средняя цена ошибки равна $M(Y) = \int_0^{0,25a} y\frac{1}{\sqrt{ya}}dy = \frac{a}{12}$. $\rhd$

ПРИМЕР 8.4. Колесо радиусом R с осью в начале координат приводится во вращение, которое затухает под действием сил трения. В результате некоторая фиксированная точка A на ободе колеса останавливается на случайном расстоянии H от горизонтальной оси (рис. 8.1). Найти распределение этой случайной величины считая, что угол поворота колеса X имеет равномерное распределение в промежутке $[-\pi/2;\pi/2]$.
$\lhd$ Так как радиус колеса равен единице, то $H = \sin X$. На $[-\pi/2;\pi/2]$ функция $H = \sin X$ монотонно возрастает и имеет обратную функцию $X = \arcsin H$.

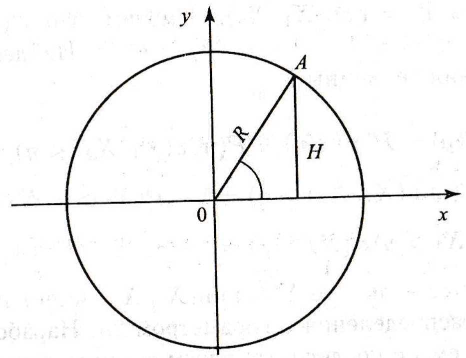


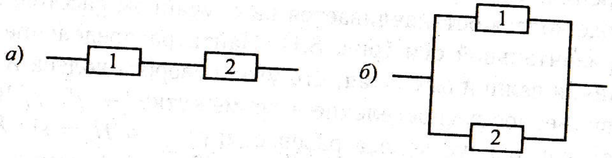
Рис. 8.1

Функция плотности вероятности величины H, согласно формуле (8.1), имеет вид

$$g(h) = \frac{1}{\pi}(\arcsin h)' = \frac{1}{\pi}\cdot\frac{1}{\sqrt{1-h^2}},$$

где $1/\pi$ — плотность вероятности случайной величины X, которая равномерно распределена на $[-\pi/2;\pi/2]$ $\rhd$.

ПРИМЕР 8.5. Время безотказной работы X каждого элемента имеет показательный закон распределения ($F(x) = P(\xi < x) = 1 - e^{-\lambda x}$, $x \geqslant 0$, $\lambda > 0$, $M(\xi) = 1/\lambda$). Считая, что элементы выходят из строя независимо друг от друга, найти среднее время безотказной работы («наработку на отказ») для каждой из систем:



$\triangleleft$ Обозначим время безотказной работы i-го элемента через X_i. Первая система выходит из строя вместе с первым отказавшим элементом, поэтому время безотказной работы первой системы $Y = \min(X_1, X_2)$. Заметим, что $P(X_i > x) = 1 - P(X < x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$. Найдем функцию распределения величины Y:

$$F(y) = P(Y < y) = P(\min(X_1, X_2) < y) =$$
$$= 1 - P(\min(X_1, X_2) > y) = 1 - P(X_1 > y, X_2 > y) =$$
$$= 1 - P(X_1 > y)P(X_2 > y) = 1 - e^{-\lambda y}e^{-\lambda y} = 1 - e^{-2\lambda y}.$$

Оказалось, что величина $Y = \min(X_1, X_2)$ имеет показательный закон распределения с параметром 2λ. Наработка на отказ для системы с последовательным соединением элементов:

$$M(\eta) = \int_0^\infty y dF(y) = \int_0^\infty y \cdot 2\lambda e^{-2\lambda y} dy = \frac{1}{2\lambda},$$

т. е. в два раза меньше наработки на отказ одного элемента.

Система б) работает безотказно, пока в рабочем состоянии находится хотя бы один из двух элементов. Поэтому ее время безотказной работы $Z = \max(X_1, X_2)$. Найдем функцию распределения величины Z:

$$F(z) = P(Z < z) = P(\max(X_1, X_2) < z) = P(X_1 < z, X_2 < z) =$$
$$= F(z) \cdot F(z) = (1 - e^{-\lambda z})^2.$$

Наработка на отказ для системы с параллельным соединением элементов (такое соединение при одновременно работающих

элементах называют нагруженным или «горячим» резервированием) определяется по формуле

$$M(\xi) = \int_0^\infty z dF(z) = \int_0^\infty z(1 - 2e^{-\lambda z} + e^{-2\lambda z})' dz =$$
$$= \int_0^\infty z(2\lambda e^{-\lambda z} - 2\lambda e^{-2\lambda z}) dz = \frac{3}{2\lambda}. \triangleright$$

ПРИМЕР 8.6. Пусть случайная величина X имеет показательный закон распределения с функцией плотности вероятности $f(x) = e^{-x}$, $0 \leqslant x$. Найти закон распределения случайной величины $H = e^{-X}$.

$\triangleleft$ Функция $h = e^{-x}$ имеет обратную функцию $x = -\ln h$. Заметим, что при неотрицательных x величина h принимает значения от 0 до 1. По формуле (8.1) вычисляем функцию плотности вероятности случайной величины H:

$$g(h) = e^{\ln h} \left| -\frac{1}{h} \right| = h \cdot \frac{1}{h} = 1.$$

Это означает, что случайная величина имеет равномерный закон распределения на отрезке $[0; 1]$. $\triangleright$

ПРИМЕР 8.7. Случайная величина X имеет пуассоновский закон распределения с параметром λ. Вычислить математическое ожидание случайной величины $1/(X + 1)$.

$\triangleleft$ Используя формулу (8.4), получаем

$$M\left(\frac{1}{X+1}\right) = \sum_{k=0}^{\infty} \frac{1}{k+1} \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} =$$
$$= \frac{e^{-\lambda}}{\lambda}\left(e^{\lambda} - 1\right) = \frac{1 - e^{-\lambda}}{\lambda}. \triangleright$$

Рассмотрим смешанную случайную величину X, функция распределения которой имеет точки разрыва $x_1, x_2, ..., x_m$ со скачками соответственно $p_1, p_2, ..., p_m$. Это означает, что X, помимо возможных значений нулевой вероятности, имеет значения $x_1, x_2, ..., x_m$ с отличными от нуля вероятностями p_1, $p_2, ..., p_m$.

В этом случае плотность распределения вероятностей в точках $x_1, x_2, ..., x_m$ обращается в бесконечность, т. е. формально не существует. Эту трудность можно обойти, если воспользоваться *дельта-функцией* $\delta(x)$, которая понимается как производная (в обобщенном смысле) от функции единичного скачка $\eta(x) = 0$ при $x \leqslant 0$ и $\eta(x) = 1$. Наглядно $\delta(x)$ можно представить себе плотностью распределения «масс», при которой в точке $x = 0$ сосредоточена единичная масса, а масса во всех остальных точках равна нулю. Поэтому

$$\int_{-\infty}^{\infty} \varphi(x)\delta(x)dx = \varphi(0)$$

для всех непрерывных функций.

Функцию распределения смешанной случайной величины можно разложить на непрерывную и скачкообразную компоненты:

$$F(x) = \tilde{F}(x) + \sum_{k=1}^{m} p_k \eta(x - x_k),$$

где $\tilde{F}(x)$ — непрерывная функция, которая не убывает и изменяется от 0 до $1 - \sum_{k=1}^{m} p_k$. Тогда

$$f(x) = \tilde{f}(x) + \sum_{k=1}^{m} p_k \delta(x - x_k)$$

и, например, формула (8.1) примет вид

$$g(h) = f[x(h)]\frac{dx(h)}{dh} + \sum_{k=1}^{m} p_k \delta(h - h_k),$$

где $h_k = h(x_k)$.

ПРИМЕР 8.8. Случайная величина X имеет закон распределения Коши с функцией плотности вероятности $f(x) = \frac{1}{\pi(1+x^2)}$. Случайная величина $Y = \varphi(X)$, где функция $\varphi(x)$ задана графически (рис. 8.2). Найти плотность вероятности величины Y.

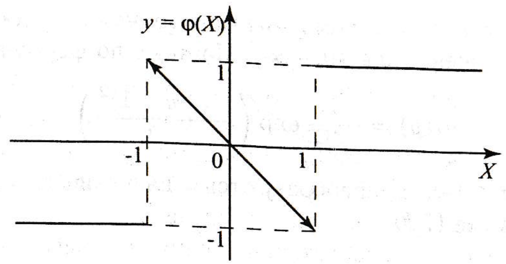


Рис. 8.2

◁ Из графика функции $\varphi(x)$ видно, что значения $X \in (-\infty, -1]$ преобразуются в значение $Y = -1$. Поэтому

$$P(Y = -1) = P(-\infty < X \leqslant 1) = \frac{1}{\pi}\int_{-\infty}^{-1} \frac{1}{1+x^2}dx =$$

$$= \frac{1}{\pi}\operatorname{arctg}(x)\Big|_{-\infty}^{-1} = \frac{1}{\pi}\left(-\frac{\pi}{4} + \frac{\pi}{2}\right) = \frac{1}{4}.$$

Аналогично $P(Y = 1) = 1/4$. На интервале $(-1, 1)$ функция $y = -x$. Обратная функция: $x = -y$. Поэтому по формуле (8.1)

$$\tilde{g}(y) = \frac{1}{\pi[1 + (-y)^2]} \cdot |-1| = \frac{1}{\pi(1+y^2)}.$$

Окончательно с учетом значений, имеющих ненулевые вероятности, получаем

$$g(y) = \frac{1}{4}\delta(y+1) + \frac{1}{4}\delta(y-1) + \frac{1}{\pi(1+y^2)}$$

при $y \in [-1, 1]$ и $g(y) = 0$ при $y \notin [-1, 1]$. ▷

ПРИМЕР 8.9. Случайная величина X имеет стандартное нормальное распределение $N(0, 1)$, т. е. функцию плотности вероятности $f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$. Найдите плотность вероятности случайной величины $Y = \varphi(X)$, где $\varphi(x) = -x - 1$ при $x < -1$, $\varphi(x) = 0$ при $x \in [-1, 1]$ и $\varphi(x) = x - 1$ при $x > 1$.

$\lhd$ Значения $x < -1$ с помощью монотонной функции отображаются в значения $y \in (0, +\infty)$. Поэтому по формуле (8.1)

$$g_1(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(-y-1)^2}{2}\right).$$

Значения $x \in [-1, 1]$ преобразуются в значение $Y = 0$. Поэтому по формуле (7.1)

$$P(Y=0) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-x^2/2} dx = 2\Phi_0(1) \approx 0,683.$$

Значения $x > 1$ с помощью монотонной функции преобразуются в значения $y \in (0, +\infty)$. По формуле (8.1)

$$g_2(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y+1)^2}{2}\right).$$

Заметим, что $g_1(y) = g_2(y)$. С учетом этого

$$g(y) = \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{(y+1)^2}{2}\right) + 2\Phi_0(1) \cdot \delta(y). \rhd$$

ПРИМЕР 8.10. Через точку, наугад взятую на окружности единичного радиуса с центром в начале координат, проводится касательная к окружности. Найти плотность распределения длины отрезка касательной, заключенного между осями координат.

$\lhd$ В силу симметрии окружности достаточно рассмотреть положения касательной с точкой касания в первой четверти (рис. 8.3).

Пусть отрезки осей, которые отсекает касательная, равны соответственно a и b. Если обозначить через X угол между осью абсцисс и направлением из начала координат на точку касания, то $\frac{1}{a} = \cos X$, откуда $a = \frac{1}{\cos X}$. Аналогично $\frac{1}{b} = \cos(90° - X) = \sin X$ и $b = \frac{1}{\sin X}$.

Тогда длина отрезка касательной, заключенного между осями координат, определяется соотношением

$$H = \sqrt{\frac{1}{\cos^2 X} + \frac{1}{\sin^2 X}} = \sqrt{\frac{\sin^2 X + \cos^2 X}{\cos^2 X \sin^2 X}} = \frac{1}{\cos X \sin X} = \frac{2}{\sin 2X}.$$

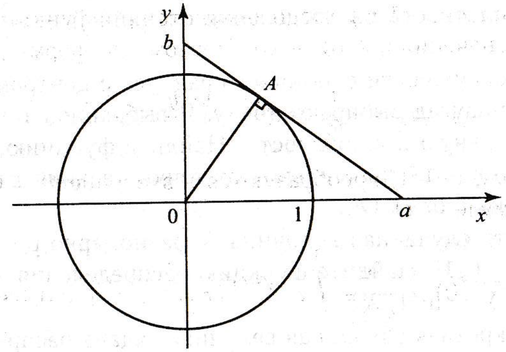


Рис. 8.3

Все положения точки A на дуге от 0° до 90° равновозможны. Следовательно, X имеет равномерное на $[0, \pi/2]$ распределение с функцией плотности вероятности $f(x) = 2/\pi$. По формуле (8.1), с учетом того, что $x = \frac{1}{2} \arcsin \frac{2}{h}$, получаем плотность вероятности

$$g(h) = \frac{2}{\pi} \cdot \frac{1}{\sqrt{1 - 4/h^2}} \cdot \frac{1}{2} \left| -\frac{2}{h^2} \right| = \frac{2}{\pi h \sqrt{h^2 - 4}}.$$

Заметим, что минимальное значение $H = 2$ получается при $X = \frac{\pi}{4}$. $\rhd$

ПРИМЕРЫ (для самостоятельного решения).

1. Случайная величина X имеет закон распределения:

X	-1	0	2	3
P	$0,1$	$0,4$	$0,3$	$0,2$

Найдите закон распределения случайной величины $Y = (X-1)^2$.

2. Случайная величина X распределена равномерно на отрезке $[a, b]$: $f(x) = \frac{1}{b-a}$, $0 < a \leqslant x \leqslant b$. Найдите плотность вероятности случайной величины $Y = X^2$.

3. Случайная величина X имеет плотность вероятности $f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$, $x \geqslant 0$ (закон распределения Релея).

Найдите плотность распределения случайной величины $Y = \sigma / X$.

4. На окружности единичного радиуса с центром в начале координат наугад выбирают точку. В выбранной точке проводят касательную к окружности. Найдите функцию распределения длины этой касательной от точки касания до точки ее пересечения с осью Ox.

5. Пусть случайная величина X равномерно распределена на отрезке $[0, 1]$. Найдите функцию распределения случайной величины $Y = \ln(1/X)$.

6. Дискретная случайная величина задана распределением

X	1	2	4
P	0,3	0,5	0,2

Найдите математическое ожидание случайной величины $Y = X^2 - X + 1$.

7. Пусть X — число выпавших гербов при трех подбрасываниях монеты. Найдите математическое ожидание случайной величины $Y = X^2$.

8. Случайная величина X равномерно распределена на отрезке $[0, 1]$. Найдите математическое ожидание случайной величины $Y = \sin(\pi X)$.

9. Непрерывная случайная величина X имеет функцию плотности вероятности: $f(x) = 0,5 \sin x$ при $x \in [0, \pi]$ и $f(x) = 0$ при остальных x. Найдите плотность вероятности и математическое ожидание случайной величины $Y = 2X$.

10. Пусть случайная величина X имеет стандартное нормальное распределение с функцией плотности вероятности $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Найдите функцию плотности вероятности случайной величины $Y = \frac{1}{2} X^2$.

11. Случайная величина X имеет функцию плотности вероятности $f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$, $x \geqslant 0$. Найдите закон распределения случайной величины $Y = \varphi(x)$, где $\varphi(x) = 0,5x$ при $x \in [0, 4]$ и $\varphi(x) \equiv 2$ при $x \in (4, +\infty)$.

12. Случайная величина X имеет стандартное нормальное распределение $N(0, 1)$, т. е. функцию плотности вероятности: $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$. Найдите плотность вероятности случайной величины $Y = \varphi(X)$, где $\varphi(x) = 0$ при $x < 0$ и $\varphi(x) = 2x$ при $x \geqslant 0$.

13. Две вершины треугольника совпадают с концами диаметра единичного круга, а третья вершина треугольника располагается в случайной точке внутри круга. Полагая равновозможными все положения случайной точки внутри круга, найдите функцию плотности вероятности площади треугольника S и математическое ожидание этой площади.

Ответы:

1.

X	1	4
P	0,7	0,3

;

2. $g(h) = \frac{1}{2(b-a)\sqrt{y}}$, $a^2 \leqslant x \leqslant b^2$;
3. $g(y) = y^{-3} \cdot \exp\left(-\frac{1}{2y^2}\right)$, $y \geqslant 0$;
4. $F(x) = \frac{2}{\pi} \operatorname{arctg} x$ при $x > 0$, $F(x) = 0$ при $x \leqslant 0$;
5. $F(x) = 1 - e^{-x}$;
6. 4,4;
7. 3;
8. $2/\pi$;
9. $f(y) = 0,25 \sin(y/2)$ при $y \in [0, 2\pi]$, $f(y) = 0$ при остальных y; $M(Y) = \pi$;
10. $g(x) = \frac{1}{\sqrt{\pi y}} e^{-y}$;
11. $g(y) = 2\lambda e^{-2\lambda x} + e^{-4\lambda} \delta(y - 2)$ при $y \in [0, 2]$ и $g(y) = 0$ при $y > 2$ и при $y < 0$;
12. $g(x) = \frac{1}{2}\delta(y) + \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{y^2}{8}\right)$;
13. $f(x) = \frac{4}{\pi}\sqrt{1 - x^2}$, $M(S) = \frac{3}{4\pi}$.