Р. С. ГУТЕР
Б. В. ОВЧИНСКИЙ

ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

ГЛАВНОЕ УПРАВЛЕНИЕ ВЫСШИХ И СРЕДНИХ ПЕДАГОГИЧЕСКИХ УЧЕБНЫХ ЗАВЕДЕНИЙ МИНИСТЕРСТВА ПРОСВЕЩЕНИЯ РСФСР

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ЗАОЧНЫЙ ПЕДАГОГИЧЕСКИЙ ИНСТИТУТ

Р. С. ГУТЕР и Б. В. ОВЧИНСКИЙ

ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

ИЗДАТЕЛЬСТВО «ПРОСВЕЩЕНИЕ» МОСКВА — 1967

ПРЕДИСЛОВИЕ

Наша книга предназначена в качестве учебного пособия по курсу теории вероятностей для студентов педагогических институтов. В ее основу положены лекционные курсы, читавшиеся обоими авторами студентам технических вузов, а также студентам-заочникам Московского государственного заочного педагогического института.

Содержание книги полностью соответствует обязательной части программы курса теории вероятностей для педагогических институтов. Что касается порядка расположения материала, то здесь мы сочли возможным несколько отступить от схемы, предложенной авторами программы. В основу нашей книги положено построение лекционного курса, выгоды которого подтверждены нашей практикой преподавания и которое уже использовалось ранее в соответствующей главе нашей предыдущей книги (см. Р. С. Гутер и Б. В. Овчинский, Элементы численного анализа и математической обработки результатов опыта, Физматгиз, 1962). Оттуда же заимствован целый ряд примеров.

Книга рассчитана, в первую очередь, на студентов-заочников. Чтобы облегчить читателям самостоятельную работу над книгой, мы сочли необходимым снабдить ее большим числом подробно разобранных примеров. Для той же цели в конце каждой главы помещены вопросы для самопроверки. Их основное назначение — дать возможность читателю проверить, насколько правильно понято им основное идейное содержание изученного материала. Вопросов типа «докажите такую-то теорему» или «выведите такую-то формулу» мы старались избегать.

В заключение мы считаем нужным отметить, что в создании этой книги значительную роль сыграл Н. Я. Виленкин, которому мы выражаем свою благодарность. Мы благодарим также А. С. Солодовникова и И. М. Яглома, внимательно прочитавших рукопись, и редактора книги Ю. А. Гастева за ряд полезных замечаний.

2-2-3

ГЛАВА I

СОБЫТИЕ И ВЕРОЯТНОСТЬ

§ 1. ОСНОВНЫЕ ПОНЯТИЯ. КЛАССИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Под *событием* мы будем понимать всякое явление, которое происходит или не происходит. Легко понять, что эта фраза отнюдь не может служить точным определением в том смысле, как мы понимаем математическое определение, однако мы вынуждены ею ограничиться.

Для большей ясности приведем некоторые примеры. Так, например, событием является выпадение герба при бросании монеты, выпадение того или иного числа очков (например, шестерки) при бросании игральной кости, попадание в цель при выстреле, нахождение молекулы газа в заранее выделенном объеме и т. п.

Различные события мы будем обозначать буквами A, B, C, ...

Событие называют *достоверным*, если оно непременно должно произойти. Так, достоверным является выпадение не более шести очков при бросании обычной игральной кости, появление белого шара при извлечении из урны, содержащей только белые шары, и т. д.

Наоборот, событие называют *невозможным*, если оно заведомо не наступит. Примерами невозможных событий являются извлечение более четырех тузов из обычной карточной колоды, появление красного шара из урны, содержащей лишь белые и черные шары, и т. д.

Пусть A — некоторое событие. Под событием, *противоположным* ему, будем понимать событие, состоящее в том, что A не наступило. Его обозначают через $\bar{A}$. Если, скажем, событие A состоит в появлении красной масти при вытаскивании карты из колоды, то $\bar{A}$ означает появление черной.

События A и B называют *несовместными*, если наступление одного из них исключает возможность наступления другого. Так, появление любого возможного числа очков при бросании игральной кости (событие A) несовместно с появлением иного числа (событие B). Выпадение четного числа очков несовместно с выпадением нечетного числа. Наоборот, выпадения четного числа очков (событие A) и числа очков, кратного трем (событие B), не будут несовместными, ибо выпадение шести очков означает наступление и события A и события B, так что наступление одного из них не исключает наступления другого. Легко понять, что события A и $\overline{A}$ всегда несовместны.

Рассмотрим некоторую совокупность событий A, B, ..., L. Эти события принято называть *единственно возможными*, если в результате каждого испытания хотя бы одно из них наверное наступит. Говорят также, что рассматриваемые события образуют *полную группу событий*. Так, например, при бросании игральной кости полную группу образуют события, состоящие в выпадении одного, двух, трех, четырех, пяти и шести очков.

Теперь мы можем перейти к рассмотрению важнейшего понятия *вероятности события*.

Рассмотрим систему конечного числа событий A_1, A_2, ..., A_n, относительно которой сделаем следующие предположения:

1. *Эти события попарно несовместны*; иначе говоря, для любых двух событий A_i и A_k (i, $k = 1, 2, \ldots, n$, $i \neq k$) появление одного из них исключает появление другого.

2. *События* A_1, A_2, ..., A_n *единственно возможны*, т. е. какое-либо одно из них непременно должно наступить.

3. *События* A_1, A_2, ..., A_n *равновозможны*. Это означает, что не существует никаких объективных причин, вследствие которых одно из них должно бы наступать чаще, чем какое-либо другое.

Пусть имеется событие A, которое наступает при появлении некоторых из наших «элементарных» событий A_1, A_2, ..., A_n и не наступает при появлении других. Мы будем говорить в таком случае, что те из «элементарных» событий A_i, при наступлении которых наступает также событие A, *благоприятствуют* событию A.

Допустим, что из общего числа n рассматриваемых событий A_1, A_2, ..., A_n событию A благоприятствует m из них. Тогда *вероятностью события* A *называется отношение числа событий; благоприятствующих событию* A, *к общему числу всех равновозможных событий*. Если, как это принято, обозначить вероятность события A через $P(A)$, то мы получаем, по определению,

$$P(A) = \frac{m}{n}.$$

Поясним приведенное нами определение примером. Рассмотрим бросание игральной кости и обозначим через A_1, A_2, ..., A_6 события, состоящие в выпадении соответственно одного, двух, ..., шести очков. Легко проверить, что эти события удовлетворяют всем сделанным выше предположениям.

Отсюда следует, что

$$P(A_1) = P(A_2) = \ldots = P(A_6) = \frac{1}{6},$$

потому что каждому из этих событий благоприятствует только оно само, так что здесь $m = 1$, а $n = 6$.

Если событие A означает появление четного числа очков, то ему благоприятствуют события A_2, A_4, A_6, состоящие в появлении двух, четырех и шести очков. Поэтому для события A имеем $m = 3$, так что $P(A) = \frac{3}{6} = \frac{1}{2}$.

Пусть событие B состоит в появлении числа очков, кратного трем. Тогда событию B благоприятствуют «элементарные» события A_3 и A_6, откуда следует, что для события B имеем $m = 2$. Поэтому $P(B) = \frac{2}{6} = \frac{1}{3}$.

Приведенное нами определение вероятности носит название *классического определения*. При всей его ясности и простоте у него есть недостатки, на которых мы подробно остановимся в дальнейшем (см. § 8). Сейчас мы хотим отметить некоторые простые свойства вероятности, основываясь на приведенном определении.

Прежде всего легко заметить, что для любого события A число благоприятствующих событий m удовлетворяет неравенствам $0 \leqslant m \leqslant n$. Поэтому *вероятность любого события A подчинена условиям*

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

Далее, обозначим через E некоторое достоверное событие. Ему, очевидно, должны благоприятствовать все «элементарные» события A_i, так что для него должно быть $m = n$. Поэтому *вероятность достоверного события равна единице*:

$$P(E) = 1.$$

Если, наоборот, U — невозможное событие, то из самого определения следует, что здесь $m = 0$, так что *вероятность невозможного события равна нулю*:

$$P(U) = 0.$$

Рассмотрим несколько примеров, разъясняющих введенное понятие вероятности.

Пример 1. В урне находятся три синих, восемь красных и девять белых шаров одинакового размера и веса, неразличимых на ощупь. Шары тщательно перемешаны. Какова вероятность появления синего, красного и белого шаров при одном вынимании шара из урны?

Так как появление любого шара можно считать равновозможным, то мы имеем всего $n = 3 + 8 + 9 = 20$ элементарных событий. Если через A, B, C обозначить события, состоящие в появлении соответственно синего, красного и белого шаров, а через m_1, m_2, m_3 — числа благоприятствующих этим событиям случаев, то ясно, что $m_1 = 3$, $m_2 = 8$, $m_3 = 9$. Поэтому

$$P(A) = \frac{3}{20} = 0,15; \quad P(B) = \frac{8}{20} = 0,40; \quad P(C) = \frac{9}{20} = 0,45.$$

Пример 2. Одновременно брошены две монеты. Какова вероятность появления m гербов ($m = 0, 1, 2$)?

Рассмотрим возможные при бросании двух монет исходы. Очевидно, их можно описать схемой

ГГ, ГР, РГ, РР,

где Г означает выпадение герба, а Р — надписи. Таким образом, возможны четыре элементарных события. Поскольку монеты предполагаются однородными и имеющими геометрически правильную форму, то нет никаких оснований предполагать, что одна из сторон какой-либо монеты выпадает чаще других. Поэтому все четыре случая следует считать равновозможными. Но тогда, обозначив через P_m вероятность выпадения m гербов, легко получим:

$$P_0 = \frac{1}{4}, \quad P_1 = \frac{2}{4} = \frac{1}{2}, \quad P_2 = \frac{1}{4}.$$

Пример 3. Одновременно бросаются две игральные кости, на гранях которых нанесены очки 1, 2, 3, 4, 5, 6. Какова вероятность того, что сумма очков, выпавших на двух костях, равна восьми?

Так как любое из возможного числа очков на одной кости может сочетаться с любым числом очков на другой, то общее число различных случаев равно $n = 6 \cdot 6 = 36$. Легко убедиться в том, что все эти случаи попарно несовместны, равновозможны и образуют полную группу событий. Для ответа на вопрос следует подсчитать, в каком числе случаев сумма очков равна восьми. Это будет, если число очков на брошенных костях равно

$$2 + 6, \ 3 + 5, \ 4 + 4, \ 5 + 3 \text{ или } 6 + 2,$$

причем первое слагаемое означает число очков на первой, а второе — на второй кости. Отсюда видно, что событию A, состоящему в том, что сумма очков, выпавших на двух костях, равна восьми, благоприятствует $m = 5$ случаев. Поэтому

$$P(A) = \frac{5}{36}.$$

§ 2. СЛОЖНЫЕ ВЕРОЯТНОСТИ. ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ. УСЛОВНЫЕ ВЕРОЯТНОСТИ

Непосредственный подсчет случаев, благоприятствующих данному событию, может оказаться затруднительным. Поэтому для определения вероятности события бывает выгодно представить данное событие в виде комбинации некоторых других, более простых событий. При этом, однако, надо знать правила, которым подчиняются вероятности при комбинации событий. Именно к этим правилам и относятся упомянутые в названии параграфа теоремы.

Первая из них относится к подсчету вероятности того, что осуществится хотя бы одно из нескольких событий.

Теорема сложения. *Пусть A и B — два несовместных события. Тогда вероятность того, что осуществится хотя бы одно из этих событий, равна сумме их вероятностей*, т. е.

$$P\,(A \text{ или } B) = P\,(A) + P\,(B). \tag{1}$$

Доказательство. Пусть A_1, A_2, ..., A_n — полная группа n попарно несовместных равновозможных событий. Если $P\,(A_1) = p_1 = \frac{m_1}{n}$, $P\,(B) = p_2 = \frac{m_2}{n}$, то среди этих n элементарных событий имеется ровно m_1 событий, благоприятствующих A, и ровно m_2 событий, благоприятствующих B. Так как события A и B несовместны, то никакое из событий A_i не может благоприятствовать обоим этим событиям. Событию (A или B), состоящему в том, что наступает хотя бы одно из этих двух событий, благоприятствует, очевидно, как каждое из событий A_i, благоприятствующих A, так и каждое из событий A_i, благоприятствующих B. Поэтому общее число событий, благоприятствующих событию (A или B), равно сумме $m_1 + m_2$, откуда следует

$$P\,(A \text{ или } B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = p_1 + p_2 = P\,(A) + P\,(B),$$

что и требовалось доказать*.

Нетрудно видеть, что теорема сложения, сформулированная выше для случая двух событий, легко переносится на случай любого конечного числа их. Именно, *если* A, B, C, ..., L — *несовместные события, то*

$$P\,(A \text{ или } B \text{ или } C \text{ или} \ldots \text{ или } L) = P\,(A) + P\,(B) + P\,(C) + \\ + \ldots + P\,(L). \tag{2}$$

* Заметим, что $P\,(A) + P\,(B) \leqslant 1$, так как $m_1 + m_2 \leqslant n$.

Для случая трех событий, например, можно написать:

P (A или B или C) $=$ P [(A или B) или C] $=$ P (A или B) $+$ $+$ P (C), откуда уже вытекает наше утверждение. Для большего числа событий следует воспользоваться методом математической индукции.

Важным следствием теоремы сложения является утверждение: *если события A_1, A_2, ..., A_n несовместны и единственно возможны, то*

$$P(A_1) + P(A_2) + \ldots + P(A_n) = 1. \qquad (3)$$

Действительно, событие (A_1 или A_2 или ... или A_n), по предположению, достоверно, и его вероятность, как было указано в § 1, равна единице. В частности, если A и $\bar{A}$ означают два взаимно противоположных события, то

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

т. е. *сумма вероятностей двух взаимно противоположных событий равна единице.*

Проиллюстрируем теорему сложения примерами.

Пример 1. При стрельбе по мишени вероятность сделать отличный выстрел равна 0,3, а вероятность сделать выстрел на оценку «хорошо» равна 0,4. Какова вероятность получить за сделанный выстрел оценку не ниже «хорошо»?

Если событие A означает получение оценки «отлично», а событие B — получение оценки «хорошо», то

$$P(A \text{ или } B) = P(A) + P(B) = 0{,}7.$$

Пример 2. В урне, содержащей n шаров белого, красного и черного цветов, находятся k белых шаров и l красных. Какова вероятность вынуть шар не черного цвета?

Пусть событие A состоит в появлении белого, а событие B — красного шара. Появление шара не черного цвета означает появление белого либо красного шара. Так как по определению вероятности

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

то по теореме сложения вероятность появления шара не черного цвета равна:

$$P(A \text{ или } B) = \frac{k}{n} + \frac{l}{n} = \frac{k+l}{n}.$$

Эту задачу можно решить и так. Пусть событие C состоит в появлении черного шара. Число черных шаров равно $n - (k + l)$, так что

$$P(C) = \frac{n-k-l}{n}.$$

Появление шара не черного цвета является противоположным событием $\overline{C}$, поэтому на основании указанного выше следствия из теоремы сложения имеем:

$$P(\overline{C}) = 1 - P(C) = 1 - \frac{n-k-l}{n} = \frac{k+l}{n},$$

как и раньше.

Пример 3. В денежно-вещевой лотерее на серию в 1000 билетов приходится 120 денежных и 80 вещевых выигрышей. Какова вероятность какого-либо выигрыша на один лотерейный билет?

Если обозначить через A событие, состоящее в выпадении денежного выигрыша, а через B — вещевого, то из определения вероятности следует:

$$P(A) = \frac{120}{1000} = 0{,}12; \quad P(B) = \frac{80}{1000} = 0{,}08.$$

Интересующее нас событие есть (A или B), поэтому из теоремы сложения вытекает

$$P(A \text{ или } B) = P(A) + P(B) = 0{,}12 + 0{,}08 = 0{,}20.$$

Таким образом, вероятность какого-либо выигрыша равна 0,2.

Прежде чем перейти к следующей теореме, необходимо ознакомиться с новым важным понятием — понятием *условной вероятности*. Для этой цели мы начнем с рассмотрения следующего примера.

Пусть на складе имеется 400 электрических лампочек, изготовленных на двух различных заводах, причем на первом изготовлено 75% всех лампочек, а на втором — 25%. Допустим, что среди лампочек, изготовленных первым заводом, 83% удовлетворяют условиям определенного стандарта, а для продукции второго завода этот процент равен 63%. Определим вероятность того, что случайно взятая со склада лампочка окажется удовлетворяющей условиям стандарта.

Заметим, что общее число имеющихся стандартных лампочек состоит из $400 \cdot 0{,}75 \cdot 0{,}83 = 249$ лампочек, изготовленных первым заводом, и $400 \cdot 0{,}25 \cdot 0{,}63 = 63$ лампочек, изготовленных вторым заводом, т. е. равно 312. Так как выбор любой лампочки следует считать равновозможным, то мы имеем 312 благоприятствующих случаев из 400, так что

$$P(B) = \frac{312}{400} = 0{,}78,$$

где событие B состоит в том, что выбранная нами лампочка стандартна.

При этом подсчете не делалось никаких предположений о том, к продукции какого завода принадлежит выбранная нами лампочка. Если же какие-либо предположения такого рода сделать, то очевидно, что интересующая нас вероятность может измениться. Так, например, если известно, что выбранная лампочка изготовлена на первом заводе (событие A), то вероятность того, что она стандартна, будет уже не 0,78, а 0,83.

Такого рода вероятность, т. е. вероятность события B при условии, что имеет место событие A, называют *условной вероятностью события B при условии наступления события A* и обозначают $P_A(B)$.

Если мы в предыдущем примере обозначим через A событие, состоящее в том, что выбранная лампочка изготовлена на первом заводе, то мы можем написать $P_A(B) = 0{,}83$.

Теперь мы можем сформулировать важную теорему, относящуюся к подсчету вероятности совмещения событий.

Теорема умножения. *Вероятность совмещения событий A и B равна произведению вероятности одного из событий на условную вероятность другого в предположении, что первое имело место*:

$$P(A \text{ и } B) = P(A)\, P_A(B). \tag{4}$$

При этом под *совмещением* событий A и B понимается *наступление каждого из них*, т. е. наступление как события A, так и события B.

Доказательство. Рассмотрим полную группу из n равновозможных попарно несовместных событий $A_1, A_2, \ldots, A_n$, каждое из которых может быть благоприятствующим или неблагоприятствующим как для события A, так и для события B.

Разобьем все эти события на четыре различные группы следующим образом. К первой группе отнесем те из событий $A_1, A_2, \ldots, A_n$, которые благоприятствуют и событию A, и событию B; ко второй и третьей группам отнесем такие события A_i, которые благоприятствуют одному из двух интересующих нас событий и не благоприятствуют другому; например, ко второй группе — те, которые благоприятствуют A, но не благоприятствуют B, а к третьей — те, которые благоприятствуют B, но не благоприятствуют A; наконец, к четвертой группе отнесем те из событий A_i, которые не благоприятствуют ни A, ни B.

Так как нумерация событий $A_1, A_2, \ldots, A_n$ не играет роли, то можно предположить, что это разбиение на четыре группы выглядит так:

I группа: $A_1, A_2, \ldots, A_k$;
II группа: $A_{k+1}, A_{k+2}, \ldots, A_{k+l}$;
III группа: $A_{k+l+1}, A_{k+l+2}, \ldots, A_{k+l+m}$;
IV группа: $A_{k+l+m+1}, \ldots, A_n$.

Таким образом, среди n равновозможных и попарно несовместных событий $A_1, A_2, \ldots, A_n$ имеется k событий, благоприятствующих и событию A, и событию B, l событий, благоприятствующих событию A, но не благоприятствующих событию B, m событий, благоприятствующих B, но не благоприятствующих A, и, наконец, $n - (k + l + m)$ событий, не благоприятствующих ни A, ни B.

Заметим, между прочим, что какая-либо из рассмотренных нами четырех групп (и даже не одна) может не содержать ни одного события. В этом случае соответствующее число, означающее количество событий в такой группе, будет равно нулю.

Произведенное нами разбиение на группы позволяет сразу написать:

$$P(A \text{ и } B) = \frac{k}{n}, \quad P(A) = \frac{k+l}{n}, \quad P(B) = \frac{k+m}{n},$$

ибо совмещению событий A и B благоприятствуют события первой группы и только они. Общее число событий, благоприятствующих A, равно общему числу событий в первой и второй группах, а благоприятствующих B — общему числу событий в первой и третьей группах.

Подсчитаем теперь вероятность $P_A(B)$, т. е. вероятность события B при условии, что событие A имело место. Теперь события, входящие в третью и четвертую группы, отпадают, так как их появление противоречило бы наступлению события A, и число возможных случаев оказывается равным уже не n, а $k + l$. Из них событию B благоприятствуют *лишь события первой группы*, так что мы получаем:

$$P_A(B) = \frac{k}{k+l}.$$

Для доказательства теоремы достаточно теперь написать очевидное тождество

$$\frac{k}{n} = \frac{k+l}{n} \cdot \frac{k}{k+l}$$

и заменить в нем все три дроби вычисленными выше вероятностями. Мы придем к утверждавшемуся в теореме равенству

$$P(A \text{ и } B) = P(A)\, P_A(B).$$

Ясно, что написанное нами выше тождество имеет смысл лишь при $k + l \neq 0$, что справедливо всегда, если только A не есть невозможное событие.

Так как события A и B равноправны, то, поменяв их местами, получим другую форму теоремы умножения:

$$P(A \text{ и } B) = P(B)\, P_B(A). \tag{5}$$

Впрочем, это равенство можно получить тем же путем, что и предыдущее, если заметить, что $P_B(A) = \frac{k}{k+m}$, и воспользоваться тождеством $\frac{k}{n} = \frac{k+m}{n} \cdot \frac{k}{k+m}$.

Сравнивая правые части двух выражений для вероятности P (A и B), получим равенство

$$P(A)\,P_A(B) = P(B)\,P_B(A), \tag{6}$$

которое часто бывает полезным.

Рассмотрим теперь примеры, иллюстрирующие теорему умножения.

Пример 4. В продукции некоторого предприятия признаются годными (событие A) 96% изделий. К первому сорту (событие B) оказываются принадлежащими 75 изделий из каждой сотни годных. Определить вероятность того, что произвольно взятое изделие принадлежит к первому сорту.

Искомая вероятность есть вероятность совмещения событий A и B. По условию имеем $P(A) = 0{,}96$ и $P_A(B) = 0{,}75$. Поэтому теорема умножения дает:

$$P(A \text{ и } B) = 0{,}96 \cdot 0{,}75 = 0{,}72.$$

Пример 5. Вероятность попадания в цель при отдельном выстреле (событие A) равна 0,2. Какова вероятность поразить цель, если 2% взрывателей дают отказы (т. е. в 2% случаев выстрела не произойдет)?

Пусть событие B состоит в том, что выстрел произойдет, а $\overline{B}$ означает противоположное событие. Тогда по условию $P(\overline{B}) = 0{,}02$ и согласно следствию из теоремы сложения $P(B) = 1 - P(\overline{B}) = 0{,}98$. Далее, по условию, $P_B(A) = 0{,}2$.

Поражение цели означает совмещение событий A и B (выстрел произойдет и даст попадание); поэтому по теореме умножения

$$P(A \text{ и } B) = P(B) \cdot P_B(A) = 0{,}196.$$

Важный частный случай теоремы умножения можно получить, если воспользоваться понятием *независимости событий*.

Два события называются независимыми, если вероятность одного из них не изменяется в результате того, наступило или не наступило другое.

Примерами независимых событий являются выпадение различного числа очков при повторном бросании игральной кости или той или иной стороны монет при повторном бросании монеты, так как очевидно, что вероятность выпадения герба при втором бросании равна $\frac{1}{2}$ независимо от того, выпал или не выпал герб в первом.

Аналогично, вероятность вынуть во второй раз белый шар из урны с белыми и черными шарами, если вынутый первым шар предварительно возвращен, не зависит от того, белый или черный шар был вынут в первый раз. Поэтому результаты первого и второго вынимания независимы между собой. Наоборот, если шар, вынутый первым, не возвращается в урну, то результат второго вынимания зависит от первого, ибо состав шаров, находящихся в урне после первого вынимания, меняется в зависимости от его исхода. Здесь мы имеем пример *зависимых* событий.

Пользуясь обозначениями, принятыми для условных вероятностей, можно записать условие независимости событий A и B в виде

$$P_A (B) = P_{\overline{A}} (B) = P (B)$$

или

$$P_B(A) = P_{\overline{B}} (A) = P (A).$$

Воспользовавшись этими равенствами, мы можем привести теорему умножения для независимых событий к следующей форме.

Если события A и B независимы, то вероятность их совмещения равна произведению вероятностей этих событий:

$$P (A \text{ и } B) = P (A) P (B). \tag{7}$$

Действительно, достаточно в первоначальном выражении теоремы умножения положить $P_A (B) = P (B)$, что вытекает из независимости событий, и мы получим требуемое равенство.

Рассмотрим теперь несколько событий: A, B, ..., L. Будем называть их *независимыми в совокупности*, если вероятность появления любого из них не зависит от того, произошли ли какие-либо другие рассматриваемые события или нет*.

В случае событий, независимых в совокупности, теорема умножения может быть распространена на любое конечное число их, благодаря чему ее можно сформулировать так:

Вероятность совмещения событий A, B, ..., L, независимых в совокупности, равна произведению вероятностей этих событий:

$$P (A \text{ и } B \text{ и } \dots \text{ и } L) = P (A) P (B) \dots P (L). \tag{8}$$

Это утверждение также доказывается по индукции.

Пример 6. Рабочий обслуживает три автоматических станка, к каждому из которых нужно подойти для устранения неисправности, если станок остановится. Вероятность того, что первый станок не остановится в течение часа, равна 0,9. Та же вероятность для второго станка равна 0,8, а для третьего — 0,7. Определить

* Отметим, что для того, чтобы события были независимыми в совокупности, недостаточно, чтобы они были лишь попарно независимы.

вероятность того, что в течение часа рабочему не потребуется подойти ни к одному из обслуживаемых им станков.

Если считать станки работающими независимо друг от друга, то в силу теоремы умножения искомая вероятность совмещения трех событий равна произведению

$$0{,}9 \cdot 0{,}8 \cdot 0{,}7 = 0{,}504.$$

Пример 7. Вероятность сбить самолет винтовочным выстрелом $p = 0{,}004$. Какова вероятность уничтожения неприятельского самолета при одновременной стрельбе из 250 винтовок?

Вероятность того, что при одиночном выстреле самолет не будет сбит, по теореме сложения равна $1 - p = 0{,}996$. Тогда можно подсчитать с помощью теоремы умножения вероятность того, что самолет не будет сбит при 250 выстрелах, как вероятность совмещения событий. Она равна $0{,}996^{250}$. После этого мы можем снова воспользоваться теоремой сложения и найти вероятность того, что самолет будет сбит, как вероятность противоположного события:

$$1 - 0{,}996^{250} \approx 0{,}62.$$

Отсюда видно, что хотя вероятность сбить самолет одиночным винтовочным выстрелом ничтожно мала, тем не менее при стрельбе из 250 винтовок вероятность сбить самолет оказывается уже весьма ощутимой. Она существенно возрастает, если число винтовок увеличить. Так, при стрельбе из 500 винтовок вероятность сбить самолет, как легко подсчитать, равна $1 - 0{,}996^{500} \approx 0{,}87$, а при стрельбе из 1000 винтовок — даже $1 - 0{,}996^{1000} \approx 0{,}98$.

Пример 8. Можно разобрать теперь предыдущий пример в самом общем виде. Пусть вероятность наступления некоторого события при отдельном испытании равна p. Решим два следующих вопроса:

а) какова вероятность P, что это событие наступит хотя бы один раз при N независимых испытаниях?

б) сколько требуется произвести испытаний, чтобы вероятность наступления события хотя бы один раз была не меньше $1 - \varepsilon$?

Повторяя рассуждения предыдущего примера, получим, что решение задачи а) дается формулой

$$P = 1 - p^N.$$

Для решения задачи б) требуется решить неравенство $P \geqslant 1 - \varepsilon$, что дает

$$N \geqslant \frac{\ln \varepsilon}{\ln (1 - p)}.$$

Доказанная выше теорема умножения позволяет несколько расширить теорему сложения, распространив ее на случай совместных событий. Если события A и B совместны, то вероятность наступления хотя бы одного из них не равна сумме их вероятностей. Например, если событие A означает выпадение четного числа очков при бросании игральной кости, а событие B — выпадение числа очков, кратного трем, то событию (A или B) благоприятствует выпадение 2, 3, 4 и 6 очков, т. е.

$$P(A \text{ или } B) = \frac{2}{3}.$$

С другой стороны, $P(A) = \frac{1}{2}$ и $P(B) = \frac{1}{3}$, т. е. $P(A) + P(B) = \frac{5}{6}$. Таким образом, в этом случае

$$P(A \text{ или } B) \neq P(A) + P(B).$$

Отсюда видно, что в случае совместных событий теорема сложения вероятностей должна быть изменена. Как мы сейчас увидим, ее можно сформулировать таким образом, чтобы она была справедлива и для совместных, и для несовместных событий, так что ранее рассмотренная теорема сложения окажется частным случаем новой.

Расширенная теорема сложения. *Пусть A и B — произвольные события. Вероятность того, что осуществится хотя бы одно из этих двух событий, равна сумме их вероятностей без вероятности их совмещения, т. е.*

$$P(A \text{ или } B) = P(A) + P(B) - P(A \text{ и } B). \tag{9}$$

Доказательство. Пусть $A_1, A_2, \ldots, A_n$ — полная группа n попарно несовместных событий. Если $P(A) = \frac{m_1}{n}$, то событию A благоприятствует m_1 из n элементарных событий. Допустим, что среди них есть k событий, благоприятствующих также и событию B, а $m_1 - k$ ему не благоприятствуют. Тогда среди n элементарных событий имеется ровно k событий, благоприятствующих и A и B. Поэтому если $P(B) = \frac{m_2}{n}$, то среди m_2 событий, благоприятствующих B, имеется k событий, благоприятствующих A, и $m_2 - k$ событий, которые A не благоприятствуют.

Все элементарные события, которые благоприятствуют событию (A или B), должны благоприятствовать либо A, либо B, либо и A и B. Таким образом, общее число таких событий равно:

$$(m_1 - k) + (m_2 - k) + k = m_1 + m_2 - k,$$

а вероятность:

$$P(A \text{ или } B) = \frac{m_1 + m_2 - k}{n} = \frac{m_1}{n} + \frac{m_2}{n} - \frac{k}{n} =$$
$$= P(A) + P(B) - P(A \text{ и } B),$$

что и требовалось доказать.

Применяя формулу (9) к рассмотренному выше примеру выпадения числа очков при бросании игральной кости, получим:

$$P(A \text{ или } B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3},$$

что совпадает с результатом непосредственного подсчета.

Очевидно, что формула (1) является частным случаем (9). Действительно, если события A и B несовместны, то $k = 0$ и вероятность совмещения $P(A \text{ и } B) = 0$.

Пример 9. В электрическую цепь включены последовательно два предохранителя. Вероятность выхода из строя первого предохранителя равна 0,6, а второго 0,2. Определить вероятность прекращения питания в результате выхода из строя хотя бы одного из этих предохранителей.

Так как события A и B, состоящие в выходе из строя первого и второго из предохранителей, совместны, то искомая вероятность определится по формуле (9):

$$P(A \text{ или } B) = 0{,}6 + 0{,}2 - 0{,}6 \cdot 0{,}2 = 0{,}68.$$

§ 3. ПОЛНАЯ ВЕРОЯТНОСТЬ. ФОРМУЛА БЕЙЕСА

При вычислении вероятностей сложных событий часто приходится одновременно применять теоремы сложения и умножения. Рассмотрим вначале следующий пример.

Пример 1. Имеется три одинаковые на вид урны с различным составом белых и черных шаров. Пусть в первой урне находится m_1 белых и n_1 черных шаров, во второй урне — соответственно m_2 белых и n_2 черных и, наконец, в третьей — m_3 белых и n_3 черных шаров. Выбирается наугад* одна из урн и из нее вынимается один шар. Требуется определить вероятность того, что вынутый шар окажется белым.

Сделаем сначала предположение, что шар вынут из первой урны. Можно сказать, что это предположение означает наступление события H_1, или осуществление гипотезы H_1. Так как выбор любой урны равновероятен, то вероятность этой гипотезы равна $P(H_1) = \frac{1}{3}$. Из предположения о составе шаров следует, что вероятность вынуть белый шар (событие A) из первой урны равна:

$$P_{H_1}(A) = \frac{m_1}{m_1 + n_1}.$$

---

* То есть выбор каждой урны равновозможен.

Рассмотрим сложное событие, состоящее в том, что выбрана первая урна и вынутый из нее шар оказался белым. Тогда вероятность такого события, в силу теоремы умножения, будет равна:

$$P(H_1 \text{ и } A) = P(H_1) P_{H_1}(A) = \frac{1}{3} \frac{m_1}{m_1 + n_1}$$

(см. формулу (4) предыдущего параграфа). Точно так же вероятность вынуть белый шар из второй урны есть вероятность сложного события, состоящего в совмещении события H_2 (выбрана вторая урна) и события A (вынут белый шар), в результате чего эта вероятность равна:

$$P(H_2 \text{ и } A) = P(H_2) P_{H_2}(A) = \frac{1}{3} \frac{m_2}{m_2 + n_2},$$

а для третьей урны:

$$P(H_3 \text{ и } A) = P(H_3) P_{H_3}(A) = \frac{1}{3} \frac{m_3}{m_3 + n_3}.$$

Событие A означает извлечение белого шара независимо от того, из какой именно урны он был вынут. Тогда, учитывая, что события H_1, H_2, H_3 несовместны, ибо выбирается лишь одна урна, мы можем воспользоваться для нахождения вероятности события теоремой сложения, которая дает:

$$\begin{aligned} P(A) &= P[(H_1 \text{ и } A) \text{ или } (H_2 \text{ и } A) \text{ или } (H_3 \text{ и } A)] = \\ &= P(H_1) P_{H_1}(A) + P(H_2) P_{H_2}(A) + P(H_3) P_{H_3}(A) = \\ &= \frac{1}{3}\left(\frac{m_1}{m_1 + n_1} + \frac{m_2}{m_2 + n_2} + \frac{m_3}{m_3 + n_3}\right). \end{aligned}$$

Сформулируем теперь общую задачу. Пусть события H_1, H_2, ..., H_n образуют полную группу несовместных событий и при наступлении каждого из них, например H_i, событие A может наступить с некоторой условной вероятностью $P_{H_i}(A)$. Какова будет при этом вероятность наступления события A?

Воспользовавшись, как и в рассмотренном выше примере, теоремой умножения, найдем, что вероятность наступления A вместе с наступлением H_1 равна:

$$P(H_1 \text{ и } A) = P(H_1) P_{H_1}(A). \tag{1}$$

Аналогично,

$$\begin{aligned} &P(H_2 \text{ и } A) = P(H_2) P_{H_2}(A), \\ &\quad \ldots\ldots\ldots\ldots \\ &P(H_n \text{ и } A) = P(H_n) P_{H_n}(A). \end{aligned} \tag{2}$$

Теперь для нахождения вероятности события A можно воспользоваться теоремой сложения, так как события $H_1, H_2, \ldots, H_n$ несовместны. Складывая все равенства (1) и (2), приходим к формуле

$$P(A) = P(H_1) P_{H_1}(A) + P(H_2) P_{H_2}(A) + \ldots + P(H_n) P_{H_n}(A),$$

или, короче,

$$P(A) = \sum_{i=1}^{n} P(H_i) P_{H_i}(A). \tag{3}$$

Формула (3) называется *формулой полной вероятности.*

События $H_1, H_2, \ldots, H_n$ обычно называют в таких случаях гипотезами. В рассмотренном выше примере 1 имелось три гипотезы ($n = 3$), которые были равновероятны между собой:

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}.$$

Пример 2. При разрыве снаряда образуются крупные, средние и мелкие осколки, причем число крупных осколков составляет 0,1 их общего числа, а число средних и мелких — соответственно 0,3 и 0,6 общего числа осколков. При попадании в танк крупный осколок пробивает броню с вероятностью 0,9, средний — с вероятностью 0,3 и мелкий — с вероятностью 0,1. Какова вероятность того, что попавший в броню осколок пробьет ее?

В нашем примере имеется три гипотезы, вероятности которых $P(H_1) = 0{,}1$, $P(H_2) = 0{,}3$ и $P(H_3) = 0{,}6$. Пользуясь формулой полной вероятности (3), находим:

$$P(A) = 0{,}1 \cdot 0{,}9 + 0{,}3 \cdot 0{,}3 + 0{,}6 \cdot 0{,}1 = 0{,}24.$$

Используя формулу полной вероятности, можно получить еще одну важную формулу, которая называется *формулой Бейеса* или *формулой вероятностей гипотез.*

Пусть мы имеем некоторую полную группу событий — гипотез H_i ($i = 1, 2, \ldots, n$), вероятность каждой из которых $P(H_i)$ до производства опыта имеет определенное значение. Предположим, что в результате опыта наступило некоторое событие A. Появление этого нового сведения — наступление события A — может повлечь за собой изменение первоначальных вероятностей гипотез.

Поясним сказанное примером.

Пример 3. Пусть урна содержит три шара белого и черного цвета, однако распределение числа шаров по цветам неизвестно. До производства опыта о содержимом урны можно сделать четыре гипотезы:

1) 3 белых и 0 черных (H_1),
2) 2 белых и 1 черный (H_2),
3) 1 белый и 2 черных (H_3),
4) 0 белых и 3 черных (H_4),

которые можно считать равновероятными:

$$P(H_1) = P(H_2) = P(H_3) = P(H_4) = \frac{1}{4}.$$

Допустим, что в результате опыта был вынут белый шар (событие A). В таком случае вероятность гипотезы H_4 делается равной нулю. Вероятности остальных трех гипотез также изменятся, причем их уже нельзя будет считать равновероятными; вероятность гипотезы H_1, например, больше, чем вероятность гипотезы H_3.

Поставим вопрос в общем виде: выяснить, каковы будут вероятности гипотез H_i после опыта в предположении, что в результате опыта наступило событие A.

Обозначим вероятности гипотез $H_1, H_2, \ldots, H_n$ до производства опыта соответственно через $\alpha_1, \alpha_2, \ldots, \alpha_n$:

$$P(H_i) = \alpha_i \quad (i = 1, 2, \ldots, n), \sum_{i=1}^{n} \alpha_i = 1.$$

Вероятности тех же гипотез после опыта,в результате которого наступило событие A, обозначим через $\beta_1, \beta_2, \ldots, \beta_n$:

$$P_A(H_i) = \beta_i \ (i = 1, 2, \ldots, n),$$

причем снова

$$\sum_{i=1}^{n} \beta_i = 1,$$

так как события $H_1, H_2, \ldots, H_n$ по-прежнему несовместны и единственно возможны. Обозначим условную вероятность $P_{H_i}(A)$ через p_i, а полную вероятность события A через P.

Пользуясь равенством (6) предыдущего параграфа, которое следует из теоремы умножения, напишем:

$$P(H_i)\, P_{H_i}(A) = P(A)\, P_A(H_i),$$

или, с введенными обозначениями,

$$\alpha_i p_i = P\beta_i.$$

Отсюда

$$\beta_i = \frac{\alpha_i p_i}{P}.$$

Подставляя сюда выражение для полной вероятности P из формулы (3), получим:

$$\beta_i = \frac{\alpha_i p_i}{\sum\limits_{j=1}^{n} \alpha_j p_j} = \frac{\alpha_i p_i}{\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_n}. \qquad (4)$$

Нетрудно проверить, что сумма всех вероятностей β_i действительно равна единице.

Формула (4), дающая выражение вероятности гипотезы H_i после опыта, и есть нужная нам *формула Бейеса*.

Вернемся снова к примеру 3. В соответствии с принятыми обозначениями, имеем $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{4}$. Далее, находим:

$$p_1 = P_{H_1}(A) = 1, \quad p_2 = \frac{2}{3}, \quad p_3 = \frac{1}{3}, \quad p_4 = 0.$$

Окончательно получим:

$$P_A(H_1) = \frac{\frac{1}{4} \cdot 1}{\frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot 0} = \frac{1}{1 + \frac{2}{3} + \frac{1}{3} + 0} = \frac{1}{2}.$$

Аналогично,

$$P_A(H_2) = \frac{1}{3}, \quad P_A(H_3) = \frac{1}{6}, \quad P_A(H_4) = 0.$$

§ 4. ПОВТОРЕНИЕ ИСПЫТАНИЙ. СХЕМА БЕРНУЛЛИ

Правила сложения и умножения вероятностей дают возможность определять вероятности достаточно сложных комбинаций событий. Одной из наиболее простых и вместе с тем весьма распространенных схем, с которой мы сейчас познакомимся, является *схема повторения независимых испытаний*, которую называют также *схемой Бернулли*.

Пусть при некотором испытании событие A может наступить или не наступить. Обозначим вероятность наступления события A через $P(A) = p$ и вероятность его ненаступления через $P(\overline{A}) = q = 1 - p$.

Рассмотрим возможные исходы двух последовательных независимых испытаний. Они описываются схемой (табл. 1), в которой приведены также вероятности различных исходов. Теперь нетрудно подсчитать, что вероятность двукратного появления события A равна p^2, вероятность его однократного появления (безразлично, при каком испытании, т. е. вероятность того, что при двух испытаниях один раз наступит A и один раз $\overline{A}$) равна $2pq$, а вероятность того, что A не наступит ни разу равна q^2. Очевидно, что эти результаты единственно возможны, причем

$$p^2 + 2pq + q^2 = 1.$$

Таблица 1

Результаты испытаний	AA	$A\bar{A}$	$\bar{A}A$	$\bar{A}\bar{A}$
Вероятность	p^2	pq	qp	q^2

Приведенное рассуждение без труда переносится на случай большего числа испытаний. Например, при трех испытаниях вероятность наступления события A три раза подряд равна p^3 (как вероятность совмещения одинаковых событий). Чтобы найти вероятность наступления события A два раза, безразлично в каком порядке, заметим, что это возможно при следующих трех исходах: $AA\bar{A}$, $A\bar{A}A$, $\bar{A}AA$, вероятность каждого из которых p^2q, так что вероятность двукратного наступления события A при трех испытаниях $3p^2q$. Аналогично подсчитывается вероятность однократного наступления $3pq^2$ и вероятность того, что событие A не наступит ни разу, которая равна q^3. Как и выше,

$$p^3 + 3p^2q + 3pq^2 + q^3 = (p + q)^3 = 1.$$

Мы можем теперь сформулировать общую задачу. *Производится серия из n независимых испытаний, то есть вероятность наступления события A при каждом отдельном испытании постоянна. Пусть она равна p. Требуется определить вероятность $P_{m,n}$ того, что событие наступит точно m раз.*

Такая задача может встретиться, например, при подсчете вероятности m попаданий в цель при n выстрелах и во многих аналогичных случаях, которые будут рассмотрены ниже.

Заметим прежде всего, что два крайних частных случая

$$P_{n,n} = p^n$$

и

$$P_{0,n} = q^n$$

легко находятся по теореме умножения (как вероятности совмещения событий).

Подсчитаем теперь вероятность того, что при n испытаниях событие A появится ровно m раз в определенном порядке, например, как в случае:

$$\underbrace{A\,A\ldots A}_{m \text{ раз}}\ \underbrace{\bar{A}\ldots\bar{A}}_{n-m \text{ раз}}.$$

Ясно, что эта вероятность равна $p^m q^{n-m}$. Очевидно, что вероятность появления события A также m раз, но в другом порядке, будет той же самой. Число всех возможных схем из n элементов, в которых m раз встречается A в различном порядке, равно числу сочетаний C_n^m.

Поэтому, пользуясь теоремой сложения вероятностей, получаем:

$$P_{m,n} = C_n^m p^m q^{n-m} = \frac{n!}{m!\,(n-m)!} p^m q^{n-m}. \qquad (1)$$

Из этой формулы видно, что вероятности $P_{m,n}$ представляют отдельные слагаемые в разложении бинома

$$(p+q)^n = p^n + np^{n-1}q + \frac{n(n-1)}{1\cdot 2}p^{n-2}q^2 + \ldots + C_n^m p^m q^{n-m} + \\ + \ldots + q^n = 1.$$

Поэтому формулу (1) называют *биномиальной*.

Итак, сформулированная выше задача полностью решена. Проиллюстрируем теперь полученную формулу двумя примерами.

Пример 1. Бросается монета 6 раз. Какова вероятность выпадения герба 0, 1, ..., 6 раз?

В данном случае $p = q = \frac{1}{2}$. Пользуясь полученной формулой, приходим к результатам:

$$P_{0,6} = P_{6,6} = \left(\frac{1}{2}\right)^6 = \frac{1}{64},$$

$$P_{1,6} = P_{5,6} = C_6^1 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right) = \frac{3}{32},$$

$$P_{2,6} = P_{4,6} = C_6^2 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^2 = \frac{15}{64},$$

$$P_{3,6} = C_6^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3 = \frac{5}{16}.$$

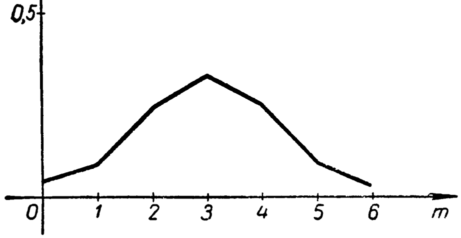


Рис. 1

Эти результаты можно изобразить графически, отложив по оси абсцисс значения m, а по оси ординат—значения $P_{m,6}$ (рис. 1). Очевидно, что наиболее вероятное число выпадений герба $m = 3$, однако вероятность эта сама по себе невелика.

Пример 2. Производится восемь выстрелов по резервуару с горючим, причем первое попадание вызывает течь, а второе — воспламенение горючего. Какова вероятность того, что резервуар будет подожжен, если вероятность попадания при отдельном выстреле равна $p = 0{,}2$?

Найдем сначала вероятность противоположного события, т. е. вероятность того, что резервуар не будет подожжен. Это произойдет лишь тогда, когда число попаданий не превзойдет единицы. Вероятность этого равна:

$$P_{0,8} + P_{1,8} = q^8 + C_8^1\, pq^7.$$

Так как здесь $p = 0{,}2$ и $q = 0{,}8$, то

$$P_{0,8} + P_{1,8} = (0{,}8)^8 + 8 \cdot 0{,}2 \cdot (0{,}8)^7 \approx 0{,}503,$$

откуда следует, что вероятность того, что резервуар будет подожжен, равна:

$$P = 1 - 0{,}503 = 0{,}497.$$

Из первого примера видно, что с ростом m вероятность $P_{m,n}$ сначала возрастает, а затем с некоторого m начинает убывать. Покажем, что это имеет место и в общем случае, и установим такое значение m, при котором $P_{m,n}$ будет наибольшим для данных n и p.

Сравним два соседних значения $P_{m,n}$ и $P_{m+1,n}$:

$$P_{m,n} = \frac{n!}{m!\,(n-m)!}\, p^m q^{n-m} = \frac{n(n-1)(n-2)\ldots(n-m+1)}{m!}\, p^m q^{n-m},$$

$$P_{m+1,n} = \frac{n!}{(m+1)!\,(n-m-1)!}\, p^{m+1} q^{n-m-1} =$$

$$= \frac{n(n-1)(n-2)\ldots(n-m+1)(n-m)}{(m+1)!}\, p^{m+1} q^{n-m-1}.$$

Отношение следующего члена к предыдущему равно:

$$\frac{P_{m+1,n}}{P_{m,n}} = \frac{n-m}{m+1}\,\frac{p}{q}. \tag{2}$$

Второй множитель этого отношения $\frac{p}{q}$ есть величина постоянная, не зависящая от m. Первый множитель $\frac{n-m}{m+1}$ с ростом m убывает, так как числитель дроби убывает, а знаменатель растет. Благодаря этому мы можем написать ряд неравенств:

$$\frac{np}{q} = \frac{P_{1,n}}{P_{0,n}} > \frac{P_{2,n}}{P_{1,n}} > \ldots > \frac{P_{n,n}}{P_{n-1,n}} = \frac{p}{nq}.$$

Если предположить, что число опытов n достаточно велико, так что $np > q$ и $nq > p$, то

$$\frac{P_{1,n}}{P_{0,n}} > 1 \quad \text{и} \quad \frac{P_{n,n}}{P_{n-1,n}} < 1.$$

Таким образом, отношение $\frac{P_{m+1,n}}{P_{m,n}}$ для малых m больше единицы, а затем становится меньше единицы. Выберем такое целое число μ, для которого $\frac{P_{\mu+1,n}}{P_{\mu,n}} \geqslant 1$, но для всех $m > \mu$ уже $\frac{P_{m+1,n}}{P_{m,n}} < 1$. Иначе говоря,

$$\begin{aligned} P_{m,n} &< P_{m+1,n} \quad \text{при} \quad m < \mu, \\ P_{m,n} &\leqslant P_{m+1,n} \quad \text{при} \quad m = \mu, \\ P_{m,n} &> P_{m+1,n} \quad \text{при} \quad m > \mu. \end{aligned}$$

Очевидно, что при $m = \mu$ вероятность $P_{\mu,n}$ будет иметь наибольшее (при данном n) значение. Если $P_{\mu+1,n} = P_{\mu,n}$, то максимальное значение вероятности достигается при двух значениях: $m = \mu$ и $m = \mu + 1$. Определим сначала величину μ в этом последнем случае. Здесь

$$\frac{P_{\mu+1,n}}{P_{\mu,n}} = \frac{n-\mu}{\mu+1} \cdot \frac{p}{q} = 1,$$

откуда находим:

$$np - \mu p = \mu q + q,$$

так что

$$\mu = np - q.$$

Следовательно, если число $np - q$ будет целым, то наиболее вероятными числами появления событий будут $\mu = np - q$ и $\mu + 1 = np - q + 1 = (n + 1)p$.

Если число $np - q$ не целое, то отношение $\frac{P_{m+1,n}}{P_{m,n}}$ переходит от значений, бо́льших единицы, сразу к значениям, меньшим единицы. Пусть μ — такое целое число, что

$$\frac{P_{\mu,n}}{P_{\mu-1,n}} > 1 \quad \text{и} \quad \frac{P_{\mu+1,n}}{P_{\mu,n}} < 1.$$

Согласно формуле (2),

$$\frac{P_{\mu,n}}{P_{\mu-1,n}} = \frac{n-\mu+1}{\mu} \cdot \frac{p}{q} > 1,$$

так что

$$np - \mu p + p > \mu q, \text{ т. е. } \mu < np + p.$$

С другой стороны,

$$\frac{P_{\mu+1,\, n}}{P_{\mu,\, n}} = \frac{n-\mu}{\mu+1} \cdot \frac{p}{q} < 1,$$

или

$$np - \mu p < q\mu + q,$$

так что

$$\mu > np - q.$$

Таким образом, в случае нецелых $np - q$ для μ получаются неравенства

$$np - q < \mu < np + p, \tag{3}$$

указывающие границы*, в которых заключено наиболее вероятное число появления события A при n испытаниях.

Рассмотрим теперь еще один пример применения выведенных формул.

Пример 3. Игральная кость бросается девять раз подряд. Каково наиболее вероятное количество выпадений числа очков, кратного трем. Определим эту вероятность, а также вероятность того, что число очков, кратное трем, выпадет не более двух раз.

Вероятность выпадения числа очков, кратного трем, равна $p = \frac{1}{3}$. В силу неравенств (3), для наиболее вероятного числа появления события находим:

$$9 \cdot \frac{1}{3} - \frac{2}{3} < \mu < 9 \cdot \frac{1}{3} + \frac{1}{3},$$

откуда следует $\mu = 3$. Вероятность того, что нужное число очков выпадет именно три раза, по формуле (1) равна:

$$P_{3,9} = C_9^3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^6 = \frac{5376}{19683} \approx 0{,}278.$$

Вероятность выпадения нужного числа очков не более двух раз может быть получена как сумма:

$$P = P_{0,9} + P_{1,9} + P_{2,9} = \left(\frac{2}{3}\right)^9 + C_9^1 \frac{1}{3} \left(\frac{2}{3}\right)^8 + C_9^2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^7 \approx 0{,}378.$$

---

* Так как $np + p - (np - q) = p + q = 1$, то интервал $(np - q,\ np + p)$ имеет длину, равную единице, и в нем обязательно лежит ровно одно целое число.

Пример 4. В некотором производстве вероятность того, что отдельная деталь окажется бракованной, равна 0,005. Какова вероятность того, что в партии из 10 000 изделий бракованных окажется:

а) ровно 40?

б) не более 70?

Ответ на первый вопрос дается непосредственно формулой $P_{m,n} = C_m^n p^m q^{n-m}$. Здесь $p = 0{,}005$, $q = 0{,}995$, $n = 10\,000$, $m = 40$, так что искомая вероятность равна:

$$P(m = 40) = P_{40,10000} = C_{10000}^{40}(0{,}005)^{40} \cdot (0{,}995)^{9960} =$$
$$= \frac{10000!}{40!\,9960!}(0{,}005)^{40}(0{,}995)^{9960}.$$

Для ответа на второй вопрос следует воспользоваться теоремой сложения вероятностей. Искомая вероятность выразится суммой:

$$P(0 \leqslant m \leqslant 70) = \sum_{m=0}^{70} P_{m,\,10000} = \sum_{m=0}^{70} C_{10000}^{m}\,(0{,}005)^m\,(0{,}995)^{10000-m}.$$

Действительно, если бракованных деталей не более 70, то это означает, что их либо 0, либо 1, либо 2, ..., либо 70. Все эти события несовместны.

Итак, ответ на оба наших вопроса мы получили. Однако практически осуществить требуемые здесь вычисления очень трудно. Это вынуждает нас искать другие формулы, позволяющие решать аналогичные вопросы хотя и приближенно, но более просто. Такие формулы будут рассмотрены нами позже (см. гл. II).

§ 5. ПРИМЕРЫ ВЫЧИСЛЕНИЯ ВЕРОЯТНОСТЕЙ

В настоящем параграфе мы разберем несколько примеров вычисления вероятностей на основе теорем, рассмотренных в предыдущих параграфах. В большинстве случаев потребуется прямой подсчет числа различных возможностей, основанный на известных формулах комбинаторики.

Пример 1. В некоторой лотерее имеется всего n билетов, из которых m являются выигрышными. Определить вероятность хотя бы одного выигрыша для лица, обладающего k билетами.

Решение. Подсчитаем сначала вероятность противоположного события, т. е. вероятность того, что ни один билет не выиграет. Представим себе, что на каждом билете заранее написано, является ли он выигрышным или нет, как например это было сделано в книжной лотерее 1965 года. Общее число равновозможных случаев выбора k билетов из имеющихся n равно числу сочетаний C_n^k. Так как невыигрышных билетов имеется $n - m$, то число элементарных событий, благоприятствующих интересующему нас сейчас событию — не выиграть ни на один билет, равно числу способов, которыми можно выбрать k билетов из $n - m$, т. е. числу сочетаний C_{n-m}^k.

Таким образом, в силу определения вероятности получаем, что вероятность не выиграть ни на один билет равна отношению $C_{n-m}^k : C_n^k$. Поскольку требуемое событие—выиграть хотя бы на один билет—является противоположным только что рассмотренному, то по следствию из теоремы сложения находим:

$$P = 1 - \frac{C_{n-m}^k}{C_n^k}.$$

Это и есть искомая вероятность.

Пример 2. Партия товара состоит из ста изделий, среди которых имеется пять бракованных. Определить вероятность того, что эта партия будет принята при испытании выбранной наудачу ее половины, если условиями приема допускается не более одного бракованного изделия на пятьдесят испытываемых.

Решение. Пусть A означает событие, состоящее в том, что среди взятых наугад 50 изделий нет ни одного бракованного. Выбор 50 изделий из общей партии в 100 штук может быть произведен C_{100}^{50} различными способами. Последнее число можно считать равным общему числу всех равновозможных случаев. Из общего числа 95 годных изделий 50 штук можно выбрать C_{95}^{50} различными способами. Очевидно, что это — число элементарных событий, благоприятствующих событию A, так как каждая из таких групп не содержит бракованных изделий. Отсюда можно найти вероятность события A, которая равна:

$$P(A) = \frac{C_{95}^{50}}{C_{100}^{50}}.$$

Пусть, далее, событие B означает выбор партии из 50 изделий, которая содержит ровно одно бракованное. Если считать все имеющиеся бракованные изделия занумерованными, то общее число групп, каждая из которых содержит одно бракованное изделие с определенным номером, равно C_{95}^{49}. Так как для нас безразлично, какое именно из бракованных изделий попало в нашу группу, то общее число случаев, благоприятствующих событию B, будет равно $C_5^1 \, C_{95}^{49}$, а вероятность его

$$P(B) = \frac{C_5^1 \cdot C_{95}^{49}}{C_{100}^{500}}.$$

Так как события A и B, очевидно, несовместны, то вероятность интересующего нас события найдется по теореме сложения:

$$p = P(A \text{ или } B) = \frac{C_{95}^{50} + C_5^1 C_{95}^{49}}{C_{100}^{50}} \approx 0{,}181.$$

Пример 3. В партии из ста деталей имеется 30 бракованных. Наугад отбирается три детали. Какова вероятность того, что среди них имеется ровно одна бракованная?

Решение. Рассуждая как и в предыдущем примере, найдем, что общее число равновозможных случаев есть C_{100}^3, а число благоприятствующих данному событию случаев равно $C_{30}^1\ C_{70}^2$. Искомая вероятность равна отношению

$$P_{100,30}(3,\ 1) = \frac{C_{30}^1 C_{70}^2}{C_{100}^3} \approx 0{,}448.$$

Рассмотренные примеры 2 и 3 являются частными случаями некоторой общей схемы, которую можно излагать в различных терминах. Удобнее всего формулировать эту схему в терминах урны с шарами. Так это и сделано в следующем примере.

Пример 4. В урне имеется N шаров, из которых M белых и $L = M - N$ черных. Вынимается наугад n шаров. Требуется определить вероятность $P_{N,\ M}(n,\ m)$ того, что среди вынутых n шаров будет ровно m белых.

Решение. Занумеруем мысленно все шары в урне таким образом, чтобы белые шары имели номера 1, 2, ..., M, а черные — номера $M+1$, ..., N. Появление любых n шаров равновероятно, и число возможных элементарных событий равно C_N^n.

Интересующему нас событию благоприятствуют те случаи, когда среди вынутых n шаров окажется ровно m белых. Последние могут быть любыми из общего числа M штук, так что общее число таких случаев равно C_M^m. Что касается черных шаров, то их должно быть $l = n - m$ штук, и число возможных выборов таких групп равно, аналогично предыдущему, C_L^l. Поскольку любая группа белых шаров может комбинироваться с любой группой черных, то общее число случаев, благоприятствующих рассматривающемуся событию, равно произведению $C_M^m\ C_L^l$.

Следовательно, искомая вероятность равна:

$$P_{N,\ M}(n,\ m) = \frac{C_M^m C_L^l}{C_N^n} = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}. \qquad (1)$$

Легко видеть, что формулы, получившиеся в примерах 3 и 4, действительно являются частными случаями формулы (1). Пользуясь известными выражениями для числа сочетаний, формулу (1) можно переписать в виде

$$P_{N,\,M}(n,\ m)=\frac{M!\,L!\,(N-n)!\,n!}{(M-m)!\,m!\,(n-m)!\,(L-l)!\,N!}. \tag{2}$$

Из смысла задачи ясно, что все разности $N-n$, $M-m$, $L-l$ положительны.

В заключение найдем вероятность $P_{N,M}(n,\ 0)$, т. е. вероятность того, что все n вынутых шаров будут черными. По формуле (1),

$$P_{N,\,M}(n,\,0)=\frac{C_M^0\,C_{N-M}^n}{C_N^n}=\frac{L!\,(N-n)!}{(L-n)!\,N!}.$$

Тот же результат легко сразу получить из формулы (2), поскольку здесь $m=0$ и $l=n$.

Пример 5. В пачке из 12 общих тетрадей имеется семь тетрадей в клетку и пять в линейку. Наугад отобраны шесть тетрадей. Какова вероятность того, что среди них будет одинаковое число тетрадей в клетку и в линейку.

Решение. В этом примере мы без труда узнаем уже разобранную схему и можем поэтому воспользоваться готовым результатом. Здесь $N=12$, $M=7$, $L=5$, $n=6$, $m=l=3$. По формуле (2) получаем:

$$P_{12,7}(6,3)=\frac{7!\,5!\,6!\,6!}{3!\,4!\,3!\,2!\,12!}=\frac{25}{66}\approx 0{,}38.$$

Пример 6. Для приема зачетов преподаватель заготовил 50 задач: 20 задач по интегральному исчислению, 20 по дифференциальным уравнениям и 10 по теории вероятностей. Для сдачи зачета студент должен решить первую же доставшуюся наугад задачу. Какова вероятность для студента сдать зачет, если он умеет решить 18 задач по интегральному исчислению, 15 задач по дифференциальным уравнениям и 5 задач по теории вероятностей?

Решение. Получение любой из 50 задач следует, очевидно, считать равновероятным. В таком случае вероятности получить задачу по интегральному исчислению (гипотеза H_1) или по дифференциальным уравнениям (гипотеза H_2) равны:

$$P(H_1)=P(H_2)=0{,}4.$$

Вероятность гипотезы H_3 получить задачу по теории вероятностей равна $P(H_3)=0{,}2$.

Если событие A означает, что задача решена, то вероятность этого события при различных гипотезах равна:

$$P_{H_1}(A)=0{,}9,$$

$$P_{H_2}(A)=0{,}75,$$

$$P_{H_3}(A)=0{,}5.$$

Окончательно по формуле полной вероятности находим:

$$P(A)=\sum_{i=1}^{3} P(H_i)\,P_{H_i}(A)=$$

$$=0{,}4\cdot 0{,}9+0{,}4\cdot 0{,}75+0{,}2\cdot 0{,}5=0{,}76.$$

Пример 7. В условиях предыдущего примера определить вероятность того, что студент решил задачу по теории вероятностей, если известно, что он сдал зачет.

Решение. Воспользуемся формулой Бейеса вероятностей гипотез. Нам требуется определить вероятность гипотезы H_3 в предположении, что наступило событие A. Пользуясь формулой (4) из § 3, находим:

$$P_A(H_3)=\frac{P(H_3)\,P_{H_3}(A)}{P(A)}=\frac{0{,}2\cdot 0{,}5}{0{,}76}=0{,}132.$$

Пример 8. Из полной карточной колоды, содержащей 52 карты, извлекается наугад одна карта. Какова вероятность извлечь картинку (короля, даму или валета) любой масти или карту пиковой масти?

Решение. Так как в колоде всего 12 картинок (по три карты каждой масти), то вероятность извлечь картинку равна $\frac{12}{52}=$ $=\frac{3}{13}$. Вероятность извлечь карту пиковой масти равна $\frac{1}{4}$. Теперь остается воспользоваться теоремой сложения. Однако необходимо учесть, что рассматриваемые события совместны. Поэтому мы должны воспользоваться расширенной теоремой сложения и формулой (9) из § 2. Для этого нужно еще определить вероятность совмещения событий, т. е. вероятность извлечь картинку пиковой масти. Таких карт в колоде три, так что соответствующая вероятность равна $\frac{3}{52}$. После этого формула (9) из § 2 дает искомую вероятность:

$$\frac{3}{13}+\frac{1}{4}-\frac{3}{52}=\frac{22}{52}=0{,}423.$$

Пример 9. Для данного баскетболиста вероятность забросить мяч в корзину равна 0,6. Произведено 8 бросков. Найти вероятность того, что при этом будет ровно два попадания. Найти наиболее вероятное число попаданий и соответствующую вероятность.

Решение. Вероятность ровно двух попаданий найдется по формуле (1) из § 4 при $n=8$, $p=0{,}6$, $q=0{,}4$, $m=2$:

$$P_{2,8}=C_8^2(0{,}6)^2\cdot(0{,}4)^6=\frac{8\cdot 7}{2}\cdot 0{,}36\cdot 0{,}0041\approx 0{,}041.$$

Для наиболее вероятного числа попаданий μ мы имели в § 4 неравенство (3), из которого следует

$$8 \cdot 0,6 - 4 < \mu < 8 \cdot 0,6 + 0,6,$$

или

$$4,4 < \mu < 5,4,$$

откуда $\mu = 5$. Соответствующая вероятность равна:

$$P_{5,8} = C_8^5 \cdot (0,6)^5 \cdot (0,4)^3 = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} 0,078 \cdot 0,064 \approx 0,28.$$

Предыдущий пример решался по схеме Бернулли повторения независимых испытаний, рассмотренной в предыдущем параграфе. В следующих примерах ситуация уже более общая — рассматривается случай более чем двух возможных исходов испытаний.

Пример 10. Игрушечный волчок разделен на 6 секторов, на которых написано по одной из цифр 1, 2, 3; каждая из них повторяется дважды. Волчок запускается три раза. Какова вероятность того, что сумма выпавших цифр будет равна 6.

Решение. Здесь мы имеем более общий случай, чем в схеме Бернулли, поскольку в результате каждого испытания может наступить одно из трех событий. Однако благодаря тому, что при каждом испытании все исходы равновозможны, решение задачи можно свести к использованию уже хорошо известных приемов.

Так как при каждом испытании возможно три исхода (может выпасть 1, 2 или 3) и результаты всех трех испытаний независимы (и, значит, могут произвольно комбинироваться друг с другом), то общее число возможных элементарных событий равно $3^3 = 27$. Остается подсчитать число элементарных событий, благоприятствующих интересующему нас событию. Для этого требуется определить количество возможных представлений числа 6 в виде суммы трех слагаемых.

Все слагаемые являются натуральными числами. Ясно, что для числа 6, без учета порядка слагаемых, возможны только два таких представления:

$$6 = 1 + 2 + 3,$$
$$6 = 2 + 2 + 2.$$

Второе представление возможно при единственном исходе — при всех трех запусках волчка выпали двойки. Что касается первого, то слагаемые в нем могут приставляться произвольно. Следовательно, возможны еще $P_3 = 6$ элементарных событий, благоприятствующих тому, что сумма очков, выпавших в результате трех запусков волчка, будет равна 6 Общее число благоприятствующих этому событию возможностей равно 7, а искомая его вероятность равна $\frac{7}{27}$.

Решение оказалось простым благодаря тому, что все исходы при каждом испытании равновероятны, и дело свелось к подсчету числа благоприятствующих возможностей. В более общем случае решение будет уже более сложным. В следующем примере рассматривается уже общая схема для некоторого конкретного числа исходов и конкретного числа испытаний.

Пример 11. Пусть в результате каждого испытания может наступить одно из трех событий A_1, A_2, A_3 с вероятностями соответственно p_1, p_2, p_3. Производится три испытания. Определить вероятность наступления k раз ($k = 0, 1, 2, 3$) каждого из возможных событий.

Решение. Для определенности проследим за вероятностями наступления различное число раз события A_1. С этой точки зрения для нас безразлично, наступит ли событие A_2 или A_3, и мы можем рассматривать только два исхода испытания — наступает или не наступает событие A_1. Вероятности этих исходов равны соответственно p_1 и $q_1 = 1 - p_1 = p_2 + p_3$, поскольку события A_1, A_2, A_3 при каждом испытании образуют полную группу.

Возможные исходы трех испытаний вместе с соответствующими вероятностями можно в этом случае записать с помощью следующей таблицы:

Таблица 2

Возможные исходы	$A_1A_1A_1$	$A_1A_1\overline{A}_1$	$A_1\overline{A}_1A_1$	$A_1A_1A_1$	$A_1\overline{A}_1\overline{A}_1$	$\overline{A}_1A_1\overline{A}_1$	$\overline{A}_1\overline{A}_1A_1$	$\overline{A}_1\overline{A}_1\overline{A}_1$
Вероятность	p_1^3	$p_1^2q_1$	$p_1^2q_1$	$p_1^2q_1$	$p_1q_1^2$	$p_1q_1^2$	$p_1q_1^2$	q_1^3

Комбинация $A_1\overline{A}_1A_1$, например, означает при этом, что событие A_1 наступает при первом и третьем испытаниях и не наступает при втором, так что при втором испытании наступает либо A_2, либо A_3. Таким образом, комбинация $A_1\overline{A}_1A_1$ равносильна осуществлению либо комбинации $A_1A_2A_1$, либо $A_1A_3A_1$. Аналогично, комбинация $\overline{A}_1A_1\overline{A}_1$ равносильна осуществлению одной из четырех следующих:

$$A_2A_1A_2,\ A_2A_1A_3,\ A_3A_1A_2,\ A_3A_1A_3.$$

Из приведенной таблицы легко находится вероятность наступления события A_1 любое число раз при трех испытаниях:

$$P_3(0) = P(\overline{A}_1\overline{A}_1\overline{A}_1) = q_1^3 = (p_2 + p_3)^3,$$

$$P_3(1) = P(A_1\overline{A}_1\overline{A}_1) + P(\overline{A}_1A_1\overline{A}_1) + P(\overline{A}_1\overline{A}_1A_1) =$$
$$= 3p_1q_1 = 3p_1(p_2 + p_3)^2, \qquad (3)$$

$$P_3(2)=P(\bar{A}_1A_1A_1)+P(A_1A_1\bar{A}_1)+P(A_1\bar{A}_1A_1)=3p_1^2q_1=3p_1^2(p_2+p_3),$$
$$P_3(3)=P(A_1A_1A_1)=p_1^3.$$

Ясно, что те же рассуждения можно повторить и для любого другого события, так что формулы, аналогичные формулам (3), справедливы и для любого числа наступлений событий A_2 и A_3. Например, для события A_2 вероятность его наступления дважды в трех испытаниях равна:

$$P_3(2)=3p_2^2q_2=3p_2^2(p_1+p_3),$$

а вероятность наступления один раз события A_3 в трех испытаниях равна:

$$P_3(1)=3p_3q_3^2=3p_3(p_1+p_2)^2.$$

Более общая задача, в которой число исходов при одном испытании и число испытаний произвольны, будет рассмотрена в следующем параграфе.

Пример 12. Найти вероятность того, что станок, работающий к моменту времени t_0, не остановится до момента t_0+t, исходя из следующих дополнительных предположений:

1) искомая вероятность $P(t)$ зависит только от величины промежутка времени t и не зависит от начального момента t_0;

2) вероятность того, что станок остановится за бесконечно малый промежуток времени продолжительности Δt, пропорциональна величине Δt с точностью до бесконечно малых более высокого порядка, чем Δt, т. е. может быть представлена в виде

$$a\Delta t+\alpha\Delta t,$$

где a — коэффициент пропорциональности, а α — бесконечно малая вместе с Δt (иначе говоря, предполагается, что $P(t)$ есть дифференцируемая функция t);

3) остановки станка в непересекающиеся промежутки времени являются независимыми событиями.

Решение. Пусть $P(t)$ означает вероятность того, что станок, работавший в момент t_0, не остановится до момента $t+t_0$. Рассмотрим промежуток времени, начиная от момента $t+t_0$, продолжительностью Δt. Вероятность того, что станок не остановится в течение этого промежутка, равна в силу предположения 1) $P(\Delta t)$, а вероятность противоположного события: $1-P(\Delta t)$. С другой стороны, вследствие предположения 2), та же вероятность равна $a\Delta t+\alpha\Delta t$, т. е.

$$1-P(\Delta t)=a\Delta t+\alpha\Delta t.$$

Определим теперь $P(t+\Delta t)$. Для этого необходимо совмещение двух событий: чтобы станок не остановился за первоначальный промежуток времени длины t и последующий промежуток длины Δt. Так как эти промежутки времени не пересекаются, то остановки за эти промежутки, в силу условия 3) независимы, и, значит, по теореме умножения вероятностей,

$$P(t+\Delta t)=P(t)\cdot P(\Delta t),$$

Учитывая предыдущее равенство, последнее можно представить в виде

$$P(t+\Delta t)=P(t)(1-a\Delta t-\alpha\Delta t).$$

Из этого равенства следует:

$$\frac{P(t+\Delta t)-P(t)}{\Delta t} = -aP(t) - \alpha P(t).$$

Переходя здесь к пределу при $\Delta t \to 0$ и заметив, что α есть бесконечно малая вместе с Δt, т. е. что $\lim \alpha = 0$, находим:

$$P'(t) = -a P(t).$$

Мы получили для неизвестной функции $P(t)$ дифференциальное уравнение первого порядка с разделяющимися переменными. Решив это уравнение обычным способом, находим его общее решение в виде

$$P(t) = Ce^{-at}.$$

Произвольное постоянное C можно определить, исходя из начального условия $P(0) = 1$, что дает $C = 1$; окончательное решение нашей задачи можно записать в виде функции

$$P(t) = e^{-at}.$$

Рассмотренная задача находит широкие применения в самых различных областях приложений теории вероятностей.

§ 6. ОБОБЩЕНИЕ СХЕМЫ БЕРНУЛЛИ. ЗАДАЧА О БЕЗВОЗВРАТНОЙ ВЫБОРКЕ

В предыдущем параграфе мы уже рассматривали пример обобщения схемы Бернулли, в котором речь шла более чем о двух исходах испытаний (см. пример 11). Однако там мы интересовались вероятностями наступления определенное число раз лишь одного из возможных событий. Сейчас мы рассмотрим такую задачу в самом общем виде.

Пусть в результате каждого испытания может наступить одно из s событий $A_1, A_2, \ldots, A_s$ с вероятностями соответственно p_1, $p_2, \ldots, p_s$, причем $\sum\limits_{i=1}^{s} p_i = 1$. Опыт повторяется n раз при неизменных условиях. Требуется найти вероятность того, что при этом событие A_1 наступит ровно m_1 раз, событие A_2 — m_2 раз, ..., событие A_s наступит m_s раз $(m_1 + m_2 + \ldots + m_s = n)$. Эту вероятность обозначим через $P_n(m_1, m_2, \ldots, m_s)$.

Запишем, как мы уже делали это и раньше, какой-либо возможный исход n испытаний в виде комбинации букв, показывающей, какое именно событие наступило при первом, втором, ... испытаниях. Например, комбинация $A_1A_3A_2A_1$ означает, что при первом испытании наступило событие A_1, при втором — A_3, при третьем — A_2, при четвертом — снова A_1 и т. д.

Интересующему нас событию благоприятствует, например, следующая комбинация:

$$\underbrace{A_1\ A_1\ \ldots\ A_1}_{m_1\ \text{раз}}\ \underbrace{A_2\ A_2\ \ldots\ A_2}_{m_2\ \text{раз}}\ \ldots\ \underbrace{A_s\ \ldots\ A_s}_{m_s\ \text{раз}},$$

которую мы назовем «главной». Вероятность ее осуществления, в силу теоремы умножения вероятностей, равна $p_1^{m_1} p_2^{m_2} \dots p_s^{m_s}$. Кроме указанной «главной» комбинации, возможны еще и «побочные», также благоприятствующие нужному событию. Это будут такие, в которых событие A_1 появляется ровно m_1 раз, хотя и не обязательно сначала, событие A_2 появляется ровно m_2 раз, хотя и не обязательно после A_1, и т. д. Важно лишь, чтобы каждое из событий A_1, A_2, ..., A_s появлялось ровно столько раз, сколько требуется. Очевидно, что вероятность каждой такой «побочной» комбинации, как и «главной», равна $p_1^{m_1} p_2^{m_2} \dots p_s^{m_s}$. Событие, состоящее в том, что в результате n испытаний A_1 наступит ровно m_1 раз, A_2 — ровно m_2 раз, ..., A_s — ровно m_s раз, распадается на частные случаи, состоящие в осуществлении «главной» или некоторой из «побочных» комбинаций, описанных выше. Все эти частные случаи попарно несовместны, поэтому для нахождения искомой вероятности можно применить теорему сложения. Как было указано, вероятность каждого из частных случаев равна $p_1^{m_1} p_2^{m_2} \dots p_s^{m_s}$. Остается определить их число.

Число всех возможных частных случаев, на которые распадается наше событие, есть число всевозможных («главной» и всех «побочных») комбинаций из n букв, в которых A_1 встречается m_1 раз, A_2 — m_2 раз, ..., A_s встречается m_s раз, т. е. число перестановок с повторениями. Как известно из курса элементарной алгебры*, число таких перестановок с повторениями равно

$$\frac{n!}{m_1!\ m_2! \dots m_s!}.$$

Таким образом, искомая вероятность равна:

$$P_n(m_1,\ m_2,\ \dots m_s) = \frac{n!}{m_1!\ m_2! \dots m_s!}\ p_1^{m_1}\ p_2^{m_2} p_s^{m_s}. \qquad (1)$$

Легко заметить, что аналогичная формула (1) из § 4 для схемы Бернулли является частным случаем полученной нами сейчас. Действительно, положим $s = 2$, т. е. примем, что возможными исходами испытания являются A_1 и A_2 с вероятностями p_1 и p_2. Тогда выведенная нами здесь формула (1) дает вероятность того, что при n испытаниях событие A_1 наступит m_1 раз, а A_2 — m_2 раз, равную

$$P_n(m_1,\ m_2) = \frac{n!}{m_1!\ m_2!}\ p_1^{m_1} p_2^{m_2}.$$

Но в случае двух возможных исходов $p_1 + p_2 = 1$, т. е. если $p_1 = p$, то $p_2 = 1 - p = q$. Далее, при $m_1 = m$ получаем $m_2 = n -$

* См., например: С. Т. Завало, Элементарная алгебра, стр. 87.

$- m_1 = n - m$. Заменяя в последней формуле m_1, m_2, p_1, p_2 указанными величинами, находим:

$$P_n(m) = \frac{n!}{m!\,(n-m)!}\,p^m q^{n-m}, \tag{2}$$

что совпадает с формулой для вероятности в схеме Бернулли.

Как было указано в § 4, вероятности $P_n(m)$ могут быть получены как отдельные слагаемые в разложении бинома $(p+q)^n$. Аналогично, вероятности $P_n(m_1, m_2, \ldots, m_s)$ могут быть получены как слагаемые в разложении полинома $(p_1 + p_2 + \ldots + p_s)^n$ в сумму по полиномиальной теореме. В самом деле, полиномиальная теорема* утверждает, что

$$(p_1 + p_2 + \ldots + p_s)^n = \sum_{m_1+m_2+\ldots+m_s=n} \frac{n!}{m_1!\,m_2!\ldots m_s!}\,p_1^{m_1} p_2^{m_2} \ldots p_s^{m_s}.$$

Поэтому формулу (1) называют *полиномиальной* формулой.

Вероятности (1) и (2) можно получить и иным путем. Составим произведение n биномов:

$$\varphi_n(\xi) = (p\xi + q)(p\xi + q)\ldots(p\xi + q) = (p\xi + q)^n. \tag{3}$$

Тогда вероятности (2) получаются как коэффициенты разложения функции (3) по степеням параметра ξ.

Функция $\varphi_n(\xi)$, разложение которой по степени параметра ξ дает в качестве коэффициентов вероятности $P_n(m)$, называется *производящей функцией* вероятностей $P_n(m)$. Из написанного выше ясно, что производящую функцию для полиномиальных вероятностей $P_n(m_1, m_2, \ldots, m_s)$ можно записать в виде

$$\psi_n(\xi_1, \xi_2, \ldots, \xi_s) = (p_1\xi_1 + p_2\xi_2 + \ldots + p_s\xi_s)^n. \tag{4}$$

Использование производящих функций позволяет рассмотреть и некоторые другие обобщения схемы Бернулли. Например, пусть при каждом испытании может наступить или не наступить событие A, но вероятность его наступления при i-м испытании равна p_i $(i = 1, 2, \ldots, n)$. Эта схема есть обобщение схемы Бернулли, в которой для всех i выполняется $p_i = p$. Нетрудно догадаться, что для такой схемы производящей функцией вероятностей будет произведение биномов:

$$\varphi_n(\xi) = (p_1\xi + q_1)(p_2\xi + q_2)\ldots(p_n\xi + q_n) = \prod_{i=1}^{n}(p_i\xi + q_i), \tag{5}$$

где $q_i = 1 - p_i$. Подробнее рассматривать эту схему мы не будем.

Рассмотрим несколько примеров применения формулы (1).

Пример 1. Мишень состоит из трех зон: круга I и колец

* См. там же, стр. 100—101.

II и III (рис. 2). Вероятности попадания в зоны I, II, III равны соответственно $p_1 = 0{,}15$, $p_2 = 0{,}22$, $p_3 = 0{,}13$. Какова вероятность того, что при десяти выстрелах получится шесть попаданий в зону I, три попадания в зону II и одно попадание в зону III?

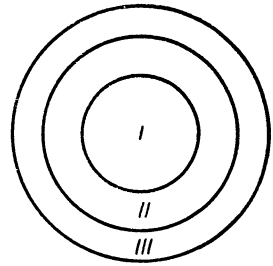


Рис. 2.

Решение. Число возможных исходов в данном случае равно $s = 4$, так как, кроме попадания в одну из зон мишени, возможен еще и промах. Вероятности попадания в различные зоны мишени указаны выше. Вероятность промаха равна $p_4 = 1 - (p_1 + p_2 + p_3) = 0{,}5$. В соответствии с условием задачи следует принять $m_1 = 6$, $m_2 = 3$, $m_3 = 1$, $m_4 = 0$. По формуле (1) находим:

$$P_{10}(6,\ 3\ 1,\ 0) = \frac{10!}{6!\ 3!\ 1!\ 0!}(0,\ 15)^6 \cdot (0{,}22)^3 \cdot (0{,}13)^1 \cdot (0{,}5)^0 =$$
$$= 840 \cdot (0{,}15)^6 \cdot (0{,}22)^3 \cdot 0{,}13 \approx 0{,}000013.$$

Пример 2. В процессе некоторого производства вероятности получить деталь с диаметром, меньшим допустимого, большим допустимого или в допустимых пределах, равны соответственно 0,05; 0,10 и 0,85. Из общей партии отобрано наудачу 100 деталей. Определить вероятность того, что среди них будет 5 деталей с диаметрами меньшими и 10 деталей с диаметрами, бóльшими допустимых.

Решение. В нашем случае имеем $p_1 = 0{,}05$, $p_2 = 0{,}10$, $p_3 = 0{,}85$, $m_1 = 5$, $m_2 = 10$ и $m_3 = 85$. Поэтому в силу формулы (1) получаем:

$$P_{100}(5,\ 10,\ 85) = \frac{100}{5!\ 10!\ 85!}(0{,}05)^5 \cdot (0{,}10)^{10} \cdot (0{,}85)^{85}.$$

Для практического выполнения вычислений удобно воспользоваться логарифмами, причем, кроме логарифмов чисел, существуют также специальные таблицы логарифмов факториалов. Логарифмирование дает:

$$\lg P_{100}(5,\ 10,\ 85) = \lg 100! - \lg 5! - \lg 10! - \lg 85! +$$
$$+ 5 \lg 0{,}05 + 10 \lg 0{,}10 + 85 \lg 0{,}5.$$

По таблице логарифмов факториалов* находим:

$$\lg 100! = 157{,}9700,$$
$$\lg 85! = 128{,}4498,$$
$$\lg 10! = 6{,}5598,$$
$$\lg 5! = 2{,}0792.$$

---

* См., например: Л. М. Милн-Томсон и Л. Дж. Комри, Четырехзначные математические таблицы, Физматгиз, 1961, стр. 36—37.

Кроме того, $5 \lg 0{,}05 + 10 \lg 0{,}10 = 5 \lg 5 - 20$. Подставив все найденные из таблиц значения, получим:

$$\lg P_{100}(5,\ 10,\ 85) = \overline{2}{,}3752,$$

откуда

$$P_{100}(5,10,85) = 0{,}0237.$$

Итак, хотя появление именно такого числа деталей естественно считать наиболее вероятным, искомая вероятность очень мала.

В заключение настоящего параграфа рассмотрим еще одну задачу, которую можно рассматривать как обобщение схемы Бернулли.

Нам уже приходилось рассматривать задачу о вынимании шара того или иного цвета из урны, извлечении карт из карточной колоды и т. п., т. е. вообще о выборе одного предмета из совокупности предметов, обладающих или не обладающих некоторым признаком. Вероятность при однократном извлечении получить предмет, обладающий признаком, равна при этом отношению числа предметов, им обладающих, к общему числу предметов.

Если производится выборка некоторого количества предметов, например нескольких шаров из урны или нескольких карт из карточной колоды, то возможны две различные постановки задачи. Одна из них состоит в том, что предметы извлекаются по одному и каждый раз возвращаются обратно перед следующим извлечением (*возвратная*, или *повторная*, выборка). В этом случае вероятность извлечь предмет, обладающий или не обладающий некоторым признаком, остается одной и той же при каждом извлечении, так что эту задачу можно рассматривать как общую задачу повторения независимых испытаний, совпадающую со схемой Бернулли.

Иначе будет обстоять дело, если все предметы извлекаются одновременно, или, что то же самое, извлекаются по одному, но после извлечения не возвращаются обратно (*безвозвратная* выборка). Такая схема не совпадает со схемой Бернулли, так как вероятность извлечения предмета, обладающего данным признаком, при каждом испытании зависит от исхода предыдущих испытаний.

Общую задачу о безвозвратной выборке можно сформулировать следующим образом: имеется совокупность N предметов, из которых M обладают свойством A. Производится выборка n предметов, причем каждый выбранный предмет обратно не возвращается. Требуется определить вероятность того, что ровно m из выбранных предметов обладают свойством A.

Такая постановка задачи уже рассматривалась нами в примере 4 предыдущего параграфа, где речь шла о выборе шаров из урны. Как там было показано, искомая вероятность выражается формулой

$$P_{N,\,M}(n,\,m)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \qquad (6)$$

или, заменяя сочетания их выражениями через факториалы,

$$P_{N,\,M}(n,\,m)=\frac{M!\,(N-M)!\,n!\,(N-n)!}{m!\,(M-n)!\,(n-m)!\,(N-M-n+n)!\,N!}\,. \qquad (7)$$

Пример 3. Из карточной колоды, содержащей 36 карт, извлекаются одна за другой четыре карты. Какова вероятность того, что среди них будет не более одного туза?

Решение. Из сказанного выше читателю должно быть ясно, что приведенная задача не поставлена достаточно четко и допускает два различных решения в зависимости от различных уточнений. Поэтому при формулировке задачи должно быть сказано, возвращается ли вынутая карта в колоду или не возвращается. Рассмотрим оба случая.

а) Предположим, что карты вынимаются по одной и каждая извлеченная карта возвращается в колоду до того, как извлекается следующая. Тогда мы можем воспользоваться биномиальной формулой. Вероятность вынуть туз при одном испытании равна $p=\frac{1}{9}$. Производится четыре испытания, а интересующее нас событие состоит в том, что среди извлеченных карт будет 0 тузов или 1 туз. По биномиальной формуле получаем искомую вероятность:

$$P_1=C_4^0\left(\frac{1}{9}\right)^0\left(\frac{8}{9}\right)^4+C_4^1\left(\frac{1}{9}\right)^1\left(\frac{8}{9}\right)^3=\frac{8^4+4\cdot 8^3}{9^4}=\frac{2048}{2187}\approx 0{,}936.$$

б) Предположим теперь, что извлеченные карты в колоду не возвращаются, или, что то же самое, извлекаются одновременно. В этом случае мы не можем воспользоваться схемой Бернулли. Ответ на поставленный вопрос дает тогда формула (6) для безвозвратной выборки. Здесь $N=36$, $M=4$, $n=4$. Что касается m, то мы должны сложить вероятности для $m=0$ и $m=1$. По формуле (6) получаем:

$$P_2=\frac{C_4^0\,C_{32}^4}{C_{36}^4}+\frac{C_4^1 C_{32}^3}{C_{36}^4}=\frac{C_{32}^4+4C_{32}^3}{C_{36}^4}\,.$$

Подставляя сюда значения сочетаний, находим:

$$\frac{\dfrac{32!}{4!\,28!}+\dfrac{4\cdot 32!}{3!\,29!}}{\dfrac{36!}{4!\,32!}}=\frac{\dfrac{32\cdot 31\cdot 30\cdot 29}{1\cdot 2\cdot 3\cdot 4}+\dfrac{4\cdot 32\cdot 31\cdot 30}{1\cdot 2\cdot 3}}{\dfrac{36\cdot 35\cdot 34\cdot 33}{1\cdot 2\cdot 3\cdot 4}}=$$

$$=\frac{32\cdot 31\cdot 30\,(29+16)}{36\cdot 35\cdot 34\cdot 33}=\frac{1240}{1309}\approx 0{,}947.$$

Рассмотренный пример показывает, что вероятности для случая возвратной и безвозвратной выборки различаются между собою, хотя различие может оказаться и незначительным. Естественно предположить, что это различие зависит от объема совокупности. Чем больше N или меньше отношение $n : N$ объема выборки к объему совокупности, тем ближе должна быть вероятность бесповторной выборки к вероятности, полученной по биномиальной формуле.

В том, что это предположение справедливо, легко убедиться с помощью следующего рассуждения. Перепишем формулу (7) таким образом:

$$P_{N,\,M}(n,\ m) = \frac{n!}{m!\,(n-m)!}\times$$

$$\times\frac{M(M-1)\ldots(M-m+1)(N-M)(N-M-1)\ldots(N-M-m+m+1)}{N(N-1)\ \ldots\ (N-n+1)}.$$

Число множителей в числителе и знаменателе второй дроби одинаково. Разделив каждый множитель числителя и знаменателя на N и обозначив отношение $\frac{M}{N}$ через p, получим:

$$P_{N,\,M}(n,\ m) = \frac{n!}{m!\,(n-m)!}\times$$

$$\times\frac{p\left(p-\frac{1}{N}\right)\ldots\left(p-\frac{m-1}{n}\right)(1-p)\left(1-p-\frac{1}{N}\right)\ldots\left(1-p-\frac{n-m-1}{N}\right)}{1\cdot\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\ldots\left(1-\frac{n-1}{N}\right)}.$$

Если теперь объем совокупности N неограниченно растет таким образом, что доля элементов, обладающих свойством A, т. е. отношение $p=\frac{M}{N}$, остается постоянным, то из формулы (8) мы видим, что в пределе выражение для $P_{N,M}(n,\ m)$ переходит в

$$P(n,\ m) = C_n^m p^m (1-p)^{n-m}.$$

Таким образом, для неограниченно большой совокупности безразлично, производится ли выборка с возвратом или без возврата.

§ 7. ЦЕПЬ МАРКОВА КАК ОБОБЩЕНИЕ СХЕМЫ БЕРНУЛЛИ

Цепь Маркова представляет собою дальнейшее по сравнению с рассмотренным в предыдущем параграфе обобщение схемы Бернулли уже на случай з а в и с и м ы х испытаний. Пусть, как и раньше, в результате каждого испытания может наступить одно из несовместных событий A_1, A_2, ..., A_k. В отличие от предыдущего будем считать, что вероятность наступления каждого из этих событий при каждом испытании з а в и с и т о т р е з у л ь т а т о в п р е д ш е с т в у ю щ и х и с п ы т а н и й. Если *условные вероятности наступления каждого события при данном испы-*

тании однозначно определяются результатом предыдущего состояния, то мы будем говорить, что такая *последовательность испытаний образует цепь Маркова*.*

Условные вероятности необходимо теперь изображать с двумя индексами. Так, p_{14} будет означать условную вероятность наступления события A_4 при условии, что при предыдущем испытании наступило событие A_1.

Цепь Маркова бывает удобно описывать в других терминах. Представим себе некоторую физическую систему, которая может находиться в одном из k различных состояний, и заданы вероятности того, что система, находящаяся в состоянии A_i, перейдет в состояние A_j, причем эти переходы могут происходить в определенные моменты времени. В таком случае нахождение системы в том или ином состоянии зависит от того, в каком из них она находилась в предшествующий момент времени, так что мы имеем дело с цепью Маркова.

Вероятность p_{ij} перехода из i-го состояния в j-е мы будем называть *вероятностью перехода.* Пользуясь предшествующими терминами испытаний, мы можем сказать, что p_{ij} есть условная вероятность наступления события A_j в данном испытании, если в предыдущем испытании наступило событие A_i.

Цепь Маркова полностью описывается заданием всех возможных вероятностей перехода, которые естественно записать в виде квадратной матрицы k-го порядка:

$$\pi_1 = \begin{Vmatrix} p_{11} & p_{12} & \dots & p_{1k} \\ p_{21} & p_{22} & \dots & p_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ p_{k1} & p_{k2} & \dots & p_{kk} \end{Vmatrix}.$$

Эту матрицу называют *матрицей перехода.* Впрочем, для полного описания цепи необходимо задать еще безусловные вероятности различных исходов при первом испытании. Однако на вероятности в процессе последовательных испытаний эти исходные вероятности не влияют, а роль играют только условные вероятности, задаваемые матрицей перехода.

Легко установить условия, которым должны удовлетворять элементы матрицы перехода. Прежде всего ясно, что все они должны удовлетворять неравенствам

$$0 \leqslant p_{ij} \leqslant 1. \tag{1}$$

Далее, из того, что события A_1, A_2, ..., A_k при каждом испытании

* Точнее говоря, здесь идет речь об *однородной* цепи Маркова, которую мы только и будем рассматривать. В более общем случае вероятность может зависеть еще от номера испытания.

образуют полную группу событий, следует, что для любой строки сумма элементов должна равняться единице:

$$\sum_{j=1}^{k} p_{ij} = 1 \quad (i = 1,\ 2,\ \ldots,\ k). \tag{2}$$

Обратно, всякая матрица, элементы которой удовлетворяют условиям (1) и (2), может служить матрицей перехода для некоторой цепи Маркова.

Рассмотрим некоторые задачи, относящиеся к цепям Маркова.

В качестве первой из них определим вероятности различных исходов в $t+2$-м испытании при известном результате t-го испытания. Для этого достаточно воспользоваться формулой полной вероятности.

Действительно, пусть при t-м испытании наступило событие A_i. Подсчитаем вероятность события A_j при $t+2$-м испытании. Это событие распадается на частные случаи в зависимости от исхода $t+1$-го испытания. Пусть, например, при $t+1$-м испытании наступило событие A_1. Вероятность этого равна p_{i1}, и вероятность после этого исхода наступления события Aj при $t+2$-м испытании будет равна p_{1j}. Таким образом, вероятность наступления A_j при $t+2$-м испытании, при условии, что при t-м наступило A_i, а при $t+1$-м A_1, равна произведению $p_{i1}\ p_{1j}$. Аналогичные произведения получаются и для других промежуточных исходов. В результате получаем:

$$P_{ij}(2) = p_{i1}p_{1j} + p_{i2}p_{2j} + \ldots + p_{ik}p_{kj} = \sum_{s=1}^{k} p_{is}p_{sj}. \tag{3}$$

Здесь $P_{ij}(2)$ означает вероятность наступления события A_j через два шага после наступления A_i. Формула (3), собственно, является формулой полной вероятности. Мы замечаем, что правая часть (3) представляет собою произведение i-й строки матрицы π_1 на ее j-й столбец. Отсюда видно, что матрица π_2 «двухшагового перехода» равна квадрату матрицы перехода π_1:

$$\pi_2 = \pi_1 \cdot \pi_1 = \pi_1^2. \tag{4}$$

Пример 1. Матрица перехода цепи Маркова имеет вид:

$$\pi_1 = \begin{Vmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{Vmatrix}.$$

Определим матрицу вероятностей перехода за два шага. Согласно формуле (4), находим:

$$\pi_2 = \pi_1^2 = \begin{Vmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{Vmatrix} = \begin{Vmatrix} \frac{7}{18} & \frac{7}{18} & \frac{2}{9} \\ \frac{13}{36} & \frac{5}{12} & \frac{2}{9} \\ \frac{5}{18} & \frac{7}{18} & \frac{1}{3} \end{Vmatrix}.$$

Аналогичную задачу можно рассмотреть и в более общем виде. Определим вероятность наступления события A_j при $t+n$-м испытании при условии, что при t-м испытании наступило событие A_i. Введя в рассмотрение некоторое промежуточное испытание $t+m$ $(1 \leqslant m < n)$ подобно (3), по формуле полной вероятности получим:

$$P_{ij}(n) = \sum_{s=1}^{k} P_{is}(m)\, P_{sj}(n-m). \tag{5}$$

Если матрицу перехода через n испытаний обозначить через π_n, то из (5) вытекает соотношение

$$\pi_n = \pi_m \cdot \pi_{n-m} \quad (0 < m < n), \tag{6}$$

частным случаем которого при $m = 1$, $n = 2$ является уже полученное равенство (4). Кроме того, из (6) следует, что вообще

$$\pi_n = \pi_1^n. \tag{7}$$

Еще одной задачей, которую мы рассмотрим, является задача о предельных вероятностях. Как ведут себя условные вероятности наступления событий A_1, A_2. ..., A_k при неограниченном возрастании числа испытаний? Ответ на этот вопрос дается следующей теоремой.

Теорема А. А. Маркова о предельных вероятностях. *Если при некотором испытании t все элементы матрицы перехода π_t положительны, то для каждого события A_j существует предельная вероятность его наступления, т. е. такое число p_j, что независимо от i имеет место равенство*

$$\lim_{n \to \infty} P_{ij}(n) = p_j. \tag{8}$$

Доказательство этой теоремы, ввиду его сложности, мы опускаем*. Разъясним только ее вероятностный смысл. Это удобно сделать в терминах перехода физической системы из одного состояния в другое.

* Его можно найти, например, в учебнике Б. В. Гнеденко «Курс теории вероятностей».

В этих терминах теорема А. А. Маркова означает, что при достаточно большом числе переходов вероятность находиться в каждом состоянии A_j практически не зависит от того, в каком состоянии система находилась в начале процесса, и мало отличается от предельной величины p_j. Эту величину p_j можно поэтому рассматривать как безусловную вероятность того, что система находится в состоянии A_j, когда число шагов достаточно велико.

Если существование предельных вероятностей p_j доказано, то легко показать, что $\sum_{j=1}^{k} p_j = 1$. Действительно,

$$\sum_{j=1}^{k} p_j = \lim_{n \to \infty} \sum_{j=1}^{k} p_{ij}(n),$$

а так как при любом n справедливо равенство $\sum_{j=1}^{k} p_{ij}(n) = 1$, то отсюда и вытекает

$$\sum_{j=1}^{k} p_j = 1, \tag{9}$$

что еще более подтверждает взгляд на предельные вероятности p_j как на безусловные вероятности наступления события A_j при достаточно большом числе испытаний.

Пример 2. Цепь Маркова задается следующей матрицей перехода:

$$\pi = \begin{Vmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{Vmatrix}.$$

Выясним, справедлива ли для данной цепи теорема о предельных вероятностях. Сама матрица π не удовлетворяет условиям теоремы Маркова, поскольку она содержит нулевые элементы. Однако уже в матрице перехода за два шага все вероятности перехода отличны от нуля. Действительно,

$$\pi_2 = \pi^2 = \begin{Vmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{Vmatrix} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{Vmatrix}.$$

Таким образом, условия теоремы о предельных вероятностях для данной цепи Маркова выполняются и, следовательно, предельные вероятности существуют. Теперь легко определить и значения этих предельных вероятностей.

Начнем с того, что вычислим несколько степеней первоначальной матрицы перехода.

$$\pi_4=(\pi_2)^2=\begin{Vmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{Vmatrix} = \begin{Vmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{3}{8} & \frac{5}{16} \\ \frac{5}{16} & \frac{5}{16} & \frac{3}{8} \end{Vmatrix};$$

$$\pi_8=(\pi_4)^2=\begin{Vmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{3}{8} & \frac{5}{16} \\ \frac{5}{16} & \frac{5}{16} & \frac{3}{8} \end{Vmatrix}^2 = \begin{Vmatrix} \frac{43}{128} & \frac{85}{256} & \frac{85}{256} \\ \frac{85}{256} & \frac{43}{128} & \frac{85}{256} \\ \frac{85}{256} & \frac{85}{256} & \frac{43}{128} \end{Vmatrix}.$$

Полученных результатов вполне достаточно, чтобы найти предельные вероятности. Действительно, мы замечаем, что все полученные матрицы перехода, как и первоначально заданная, имеют на главной диагонали одинаковые элементы. Если считать этот факт установленным, то отсюда следует, что предельные вероятности всех возможных исходов должны быть равны между собою, а так как их всего три, то предельные вероятности должны равняться $\frac{1}{3}$.

Это станет еще более ясным, если записать элементы матриц π_4 и π_8 в десятичных дробях. В самом деле,

$$\pi_4 = \begin{Vmatrix} 0{,}375 & 0{,}3125 & 0{,}3125 \\ 0{,}3125 & 0{,}375 & 0{,}3125 \\ 0{,}3125 & 0{,}3125 & 0{,}375 \end{Vmatrix},$$

$$\pi_8 = \begin{Vmatrix} 0{,}336 & 0{,}332 & 0{,}332 \\ 0{,}332 & 0{,}336 & 0{,}332 \\ 0{,}332 & 0{,}332 & 0{,}336 \end{Vmatrix}.$$

Матрица $\pi_{16} = \pi_8^2$ имеет уже вид:

$$\pi_{16} = \begin{Vmatrix} 0{,}33334 & 0{,}33332 & 0{,}33332 \\ 0{,}33332 & 0{,}33334 & 0{,}33332 \\ 0{,}33332 & 0{,}33332 & 0{,}33334 \end{Vmatrix},$$

что не оставляет никаких сомнений в значении предельных вероятностей. Однако приведенные вычисления не могут, разумеется, служить доказательством, которое мы еще должны будем провести.

Докажем прежде всего, что матрица перехода $\pi_2{}^n$ всегда имеет вид:

$$\pi_2{}^n = \begin{Vmatrix} a_n & b_n & b_n \\ b_n & a_n & b_n \\ b_n & b_n & a_n \end{Vmatrix} \tag{10}$$

где, следовательно, $b_n = \frac{1-a_n}{2}$, так как сумма элементов строки всегда должна равняться единице. Вид матриц π_2, π_4, π_8, которые мы вычислили, показывает, что для $n = 1, 2, 3$ это утверждение справедливо.

Пусть для некоторого n матрица перехода $\pi_2{}^n$ имеет вид (10). Чтобы доказать справедливость этой формулы для $n+1$, возведем матрицу (10) в квадрат:

$$\pi_{2^{n+1}} = \pi_{2^n} \cdot \pi_{2^n} = \begin{Vmatrix} a_n & b_n & b_n \\ b_n & a_n & b_n \\ b_n & b_n & a_n \end{Vmatrix} \cdot \begin{Vmatrix} a_n & b_n & b_n \\ b_n & a_n & b_n \\ b_n & b_n & a_n \end{Vmatrix}.$$

Перемножая написанные матрицы, находим, что диагональный элемент матрицы произведения всегда равен

$$a_n^2 + 2b_n^2,$$

все недиагональные элементы —

$$2a_n b_n + b_n^2.$$

Таким образом, матрица $\pi_2{}^{n+1}$ имеет вид:

$$\pi_{2^{n+1}} = \begin{Vmatrix} a_{n+1} & b_{n+1} & b_{n+1} \\ b_{n+1} & a_{n+1} & b_{n+1} \\ b_{n+1} & b_{n+1} & a_{n+1} \end{Vmatrix},$$

где $a_{n+1} = a_n^2 + 2b_n^2$ и $b_{n+1} = 2a_n b_n + b_n^2$. Следовательно, формула (10) справедлива для любого n.

Как уже было замечено, $b_n = \frac{1-a_n}{2}$. Пользуясь этим равенством, мы можем выразить a_{n+1} через a_n:

$$a_{n+1} = a_n^2 + 2\left(\frac{1-2a_n}{2}\right)^2 = \frac{3a_n^2 - 2a_n + 1}{2}.$$

Остается доказать, что $\lim\limits_{n\to\infty} a_n = \frac{1}{3}$. Положим $a_n = \frac{1}{3} + \varepsilon_n$ и рассмотрим разность $\varepsilon_{n+1} = a_{n+1} - \frac{1}{3}$:

$$\varepsilon_{n+1} = a_{n+1} - \frac{1}{3} = \frac{3a_n^2 - 2a_n + 1}{2} - \frac{1}{3} =$$

$$= \frac{9a_n^2 - 6a_n + 1}{6} = \frac{9\left(\frac{1}{3} + \varepsilon_n\right)^2 - 6\left(\frac{1}{3} + \varepsilon_n\right) + 1}{6} = \frac{3}{2}\,\varepsilon_n^2.$$

Применяя полученное равенство при $n = 1, 2, \ldots$, получим:

$$\varepsilon_n = \left(\frac{3}{2}\right)^{2n-1} \varepsilon_1^{2n}.$$

Но $a_1 = 0$; поэтому $\varepsilon_1 = \frac{1}{3}$, а, значит,

$$\varepsilon_n = \frac{1}{3 \cdot 2^{2n-1}},$$

откуда следует $\lim\limits_{n \to \infty} \varepsilon_n = 0$ и $\lim\limits_{n \to \infty} a_n = \frac{1}{3}$.

Впрочем, из равенства диагональных элементов уже следует равенство предельных вероятностей для всех исходов, то есть

$$p_1 = p_2 = p_3 = \frac{1}{3},$$

так что последнее доказательство можно было бы и не проводить.

§ 8. ДРУГИЕ ОПРЕДЕЛЕНИЯ ВЕРОЯТНОСТИ. АКСИОМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Рассмотренное выше определение вероятности, которое мы назвали классическим, оказывается неприменимым во многих случаях, важных для приложений.

Прежде всего для классического определения вероятности требуется рассмотрение к о н е ч н о г о ч и с л а единственно возможных и несовместных равновозможных событий; между тем, как раз добиться к о н е ч н о с т и числа допустимых случаев далеко не всегда возможно. Неосуществимым часто оказывается и само допущение о разложении рассматриваемого события на р а в н о в о з м о ж н ы е случаи.

Преодолеть первую из указанных трудностей иногда удается с помощью так называемого *геометрического определения вероятности*, сущность которого хорошо видна из следующего примера.

П р и м е р 1. Территория нефтебазы имеет форму прямоугольника со сторонами $a = 50$ *м* и $b = 30$ *м*. На территории имеются четыре круглых нефтебака диаметром 10 *м* каждый (рис. 3). Какова вероятность прямого поражения нефтебаков бомбой, попавшей на территорию нефтебазы, если попадание бомбы в любую точку одинаково вероятно?

Попробуем прежде всего понять, в каком точном смысле вообще можно говорить о «вероятности попадания точки в некоторую область».

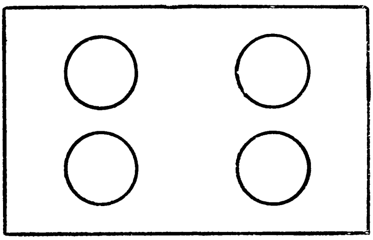


Рис. 3

Рассмотрим сначала более простые случаи. Пусть прямоугольник разделен осями симметрии на четыре равные части. Тогда можно считать, что попадание в каждую из этих частей равновозможно. Так как число частей конечно, то здесь применимо классическое определение вероятности и вероятность попадания в любую из этих частей равна 1/4. Аналогично, если разбить прямоугольник на n равных частей, то вероятность попадания в каждую из них будет равна $\frac{1}{n}$, т. е. отношению площади каждой части к площади всего прямоугольника. Если указанные части не равны, но равновелики, т. е. имеют одинаковую площадь, то все равно попадание точки в каждую из них будем считать равновероятным. Таким образом, если прямоугольник разбит на некоторое число n равновеликих частей, то вероятность попадания в каждую из них равна $\frac{1}{n}$.

Теперь ясно, что если взять область, соизмеримую с прямоугольником, т. е. если прямоугольник можно разбить на конечное число равновеликих частей, таких, что в данной области будет содержаться целое число их, то вероятность попадания в эту область будет равна отношению ее площади к площади прямоугольника.

Если же область и прямоугольник несоизмеримы (т. е. отношение их площадей иррационально), то указанное разбиение невозможно. Это означает, что при определении вероятности попадания в такую область мы не сможем воспользоваться классическим определением вероятности, так как не сможем образовать полную группу, состоящую из конечного числа равновозможных событий. Введем поэтому новое определение, которое назовем *геометрическим определением вероятности*.

Пусть D — некоторая конечная область. В эту область бросается точка, причем попадания ее в любые части области, имеющие одинаковую площадь, считаются равновозможными. Тогда *вероятностью попадания точки в любую область D_1, лежащую в D, назовем отношение площади области D_1 к площади области D.*

Обозначим эту вероятность через $P(D_1)$. По определению,

$$P(D_1)=\frac{S_1}{S},$$

где S — площадь области D, а S_1 — области D_1.

Легко заметить, что это новое определение не противоречит классическому: в тех случаях, когда удается воспользоваться двумя определениями, оба приводят к одному и тому же значению вероятности. Теоремы сложения и умножения вероятностей легко могут быть сформулированы и для геометрической вероятности.

1) Если области D_1 и D_2 принадлежат области D и не пересекаются, то вероятность попадания в D_1 или в D_2 равна сумме вероятностей попадания в D_1 или в D_2:

$$P(D_1 \text{ или } D_2) = P(D_1) + P(D_2).$$

Условие, что области D_1 и D_2 не пересекаются, соответствует условию несовместности событий.

2) Если области D_1 и D_2 имеют общую часть, то вероятность попадания в их общую часть D_3 равна произведению вероятности попадания в D_1 на условную вероятность попадания в D_2 при условии, что имеет место попадание в D_1. Запишем это так:

$$P(D_1 \text{ и } D_2) = P(D_1)\cdot P_{D_1}(D_2).$$

Обе теоремы читатель легко докажет самостоятельно.

Пользуясь первой из сформулированных теорем, мы можем легко решить вопрос, поставленный в примере 1. Здесь областью D является прямоугольник и $S = 1500 м^2$. Области D_1, D_2, D_3 и D_4 — круги, площадь каждого из которых равна $25\pi м^2 \approx 78{,}5\ м^2$. Так как эти области не пересекаются, то искомая вероятность P равна:

$$P = 4\cdot\frac{78{,}5}{1500}\approx 0{,}21.$$

При рассмотрении геометрической вероятности мы встречаемся уже с некоторыми новыми фактами. Так, при классическом определении справедливо не только утверждение, что вероятность достоверного события равна единице, но и обратное: если вероятность события равна единице, то событие достоверно. Действительно, $P(A) = 1$ означает, что $m = n$, т. е. данному событию благоприятствуют все элементарные события A_i полной группы, так что A необходимо наступит.

Для геометрических вероятностей это обратное заключение оказывается несправедливым. В самом деле, выделим в рассматриваемой области D конечное число точек или даже целую линию. Площадь оставшейся части равна площади всей области, а поэтому и вероятность попадания точки в эту оставшуюся часть равна единице. Тем не менее это событие не является достоверным, ибо возможно попадание в выделенную точку или линию. Точно так же вероятность попадания в выделенные точки равна нулю, в то время как это событие является возможным. С подобного рода явлением мы встретимся и дальше при изучении непрерывных случайных величин.

Пример 2. (Задача Бюффона.) На плоскости проведены параллельные прямые на расстоянии $2a$ друг от друга. На плоскость бросается наугад игла длиной $2l$ ($l < a$). Какова вероятность того, что игла пересечет какую-либо из проведенных прямых?

Решение. Положение иглы относительно проведенных прямых определяется расстоянием x ее центра до ближайшей прямой и углом φ, образованным иглой с этими прямыми. Эти величины удовлетворяют очевидным неравенствам $0 \leqslant x \leqslant a$, $0 \leqslant \varphi < \pi$. Поэтому положение иглы можно определить точкой (φ, x) прямоугольника $OABC$ на плоскости φOx (рис.4), имеющего указанные выше размеры. Утверждение, что игла бросается наугад, можно рассматривать как равновероятность выбора любой точки этого прямоугольника.

Если брошенная игла пересечет прямую, то, как видно из рис. 5, она образует прямоугольный треугольник LMN, стороны которого связаны соотношением $LM = MN \sin \varphi$. Так как $MN \leqslant l$, то мы получаем неравенство $x \leqslant l \sin \varphi$. Обратно, если это неравенство выполняется, то ясно, что игла пересечет одну из линий.

Таким образом, положения иглы, при которых она пересекает какую-либо прямую, характеризуются такими точками прямоугольника $OABC$, координаты которых удовлетворяют неравенству

$$x \leqslant l \sin \varphi.$$

Очевидно, что такие точки лежат под синусоидой $x = l \sin \varphi$, т. е. заполняют в прямоугольнике $OABC$ заштрихованную область (см. рис. 4).

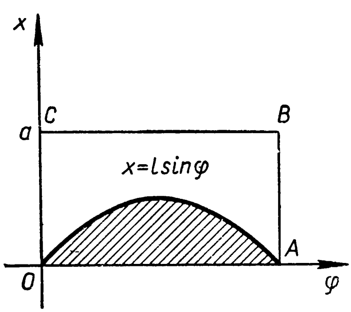


Рис. 4

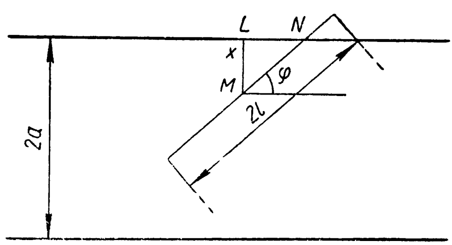


Рис. 5

Речь идет, следовательно, о том, чтобы определить вероятность попадания точки в заштрихованную область. Эта вероятность равна отношению площади этой области $S_1 = \int_0^\pi l \sin\varphi \, d\varphi$ к площади всего прямоугольника $S = \pi a$.

Итак, искомая вероятность p равна:

$$p = \frac{S_1}{S} = \frac{1}{\pi a}\int_0^\pi l \sin\varphi \, d\varphi = \frac{1}{\pi a}\Big[-l\cos\varphi\Big]_0^\pi = \frac{2l}{\pi a}.$$

В различных приложениях теории вероятностей в естественно-научных и технических вопросах часто пользуются так называемым *статистическим определением вероятности*.

Допустим, что имеется возможность неограниченного повторения испытаний, в каждом из которых при сохранении неизменных условий отмечается появление или непоявление некоторого события. (Примеры: бросание монеты или игральной кости, извлечение шара из урны (с возвратом), стрельба по цели и т. п.)

Пусть при достаточно большом числе n испытаний интересующее нас событие наступило m раз. Отношение $\mu = \frac{m}{n}$ принято называть *частостью* (иногда также *частотой* или *относительной частотой*) события.

Наблюдение за появлением некоторых событий показало, что в ряде случаев при очень большом числе испытаний их частота сохраняет почти постоянную величину, причем колебания ее становятся тем меньше, чем больше число испытаний.

Например, распределение новорожденных по полу может быть каким угодно, пока мы ограничиваемся рассмотрением нескольких семей или даже небольшого города, да еще за сравнительно короткий промежуток времени. Если же перейти к рассмотрению большой территории с большим населением, то здесь полностью обнаруживается устойчивость частоты рождения девочек и мальчиков, причем она оказывается примерно одинаковой для разных территорий.

По данным шведской статистики, частота рождения девочек в 1935 г. изменялась по месяцам следующим образом (табл. 3):

Таблица 3

Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь	За год
0,486	0,489	0,490	0,471	0,478	0,482	0,462	0,484	0,485	0,491	0,482	0,473	0,482

Подобная устойчивость частот дает основание полагать, что рассматриваемое событие (в данном случае рождение девочки) имеет определенную «вероятность», вокруг которой и происходит колебание частоты. Это допущение тем более естественно, что в тех случаях, когда применимо классическое определение вероятности, поведение частот оказывается примерно таким же*.

Для изучения частот выпадения герба ** были произведены эксперименты, давшие следующие результаты (табл. 4):

Таблица 4

Экспериментатор	Число бросаний	Число выпадений герба	Частота
Бюффон	4 040	2 048	0,5080
К. Пирсон	12 000	6 019	0,5016
К. Пирсон	24 000	12 012	0,5005

Используя указанные свойства частот, *вероятностью события называют характеризующее его число, около которого колеблется частота появления события, при сохранении неизменных условий опыта.*

Приведенное определение называют статистическим определением вероятности.

Это определение также не может охватить всех случаев применения понятия вероятности. Более того, оно и не определяет однозначно численного значения вероятности, поскольку колебания частот оставляют для этого значения довольно широкие границы.

В настоящее время при строгом построении теории вероятностей принято пользоваться *аксиоматическим определением.* Согласно этому определению, каждому событию из определенного множества событий ставится в соответствие некоторое число, причем это соответствие должно обладать определенными, заранее предположенными свойствами, то есть удовлетворять заданным аксиомам.

Аксиомы, приводимые в определении, не являются произвольно придуманными. Они заимствованы из практики и отражают именно те свойства вероятности, которые нужны для приложений и которые были установлены нами выше для классического определения.

Для большего удобства формулировок введем еще определения *суммы* и *произведения событий.*

*Суммой событий A и B называется новое событие C, состоящее в том, что наступило или событие A или событие B***.*

---

* Подробнее об этом будет сказано ниже, в § 11.

** К этому случаю применимо классическое определение вероятности.

*** Не исключается возможность наступления обоих.

Таким образом, запись $C = A + B$ заменяет использовавшееся в § 2 обозначение (A или B).

Аналогично *произведением событий A и B называется событие C, состоящее в совмещении событий A и B, так что запись $C = AB$ заменяет* (A *и* B).

Чтобы пояснить введенные термины, рассмотрим следующий пример.

Пусть в сосуде с газом выделены области α и β (рис. 6). Если событие A означает попадание некоторой молекулы газа в область α, а событие B — попадание той же молекулы в область β, то сумма событий $A + B$ означает попадание в какую-либо из этих областей, т. е. в сумму $\alpha + \beta$, а произведение AB — попадание в их общую часть, т. е. в пересечение $\alpha\beta$.

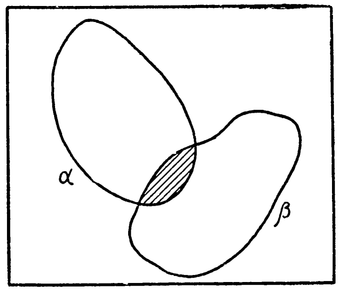


Рис. 6.

Рассмотрим теперь некоторое множество событий, обладающее тем свойством, что вместе с каждым событием множеству принадлежит и противоположное; сумма и произведение конечного числа или бесконечной последовательности событий, принадлежащих нашему множеству, также принадлежит этому множеству*. Это множество событий должно содержать также достоверное событие. События, входящие в рассматриваемое множество, будем называть *допустимыми*. Предположим, что на этом множестве допустимых событий определена числовая функция, ставящая в соответствие каждому событию A число $P(A)$ и обладающая следующими свойствами:

1°. *Для любого допустимого события A*

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

2°. *Если событие E достоверно, то*

$$P(E) = 1.$$

3°. *Если событие A влечет событие B,* т. е. из наступления события A следует также наступление события B, то

$$P(A) \leqslant P(B).$$

4°. *Если события $A, B, \ldots$ попарно несовместны, то*

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

5°. *Если события $A, B, \ldots$ в совокупности независимы, то*

$$P(AB\ldots) = P(A)\,P(B)\ldots .$$

---

* Такое множество называют *булевой алгеброй*.

Свойства 1° — 5° называют *аксиомами вероятности.*

Числовое значение функции $P(A)$, удовлетворяющей аксиомам вероятности, называют вероятностью допустимого события A.

При практическом применении этого определения обычно поступают следующим образом. Выделяется множество A_1, A_2, ... основных «элементарных событий», каждому из которых, исходя из смысла задачи, приписывается определенная вероятность $P(A_i) = p_i$. Множество допустимых событий составляется из всех событий, которые могут быть получены в виде комбинаций сумм и произведений элементарных событий. Вероятность каждого из допустимых событий находится теперь из вероятностей элементарных событий с помощью аксиом 1° — 5°.

Что касается вероятностей элементарных событий, то необходимо еще раз подчеркнуть, что они не получаются из математических соображений, а должны быть выведены исходя из смысла задачи. В условиях применимости классического определения число n элементарных событий конечно, и вероятность каждого получается равной $\frac{1}{n}$, так как события равновозможны, а их сумма достоверна.

Аксиоматический подход к определению вероятности тесно связан с теорией множеств и теорией функций действительного переменного. В самом деле, если рассматривать множество допустимых событий, то определенным выше действиям сложения и умножения событий соответствуют обычные операции сложения и пересечения множеств. Вероятности события соответствует в этом случае мера множества. Определенная как числовая функция, ставящая в соответствие каждому измеримому множеству действительное число между нулем и единицей (за единицу принимается мера всего пространства), мера удовлетворяет всем аксиомам вероятности, перечисленным выше.

Впервые такой подход к понятию вероятности и строгое аксиоматическое построение теории вероятностей на базе теории множеств были даны А. Н. Колмогоровым в его книге «Основные понятия теории вероятностей».

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ К ГЛАВЕ 1

1. Что понимается в теории вероятностей под словом «событие»?

2. Какие события называются противоположными? Что такое несовместные события? Являются ли противоположные события несовместными?

3. Можно ли говорить, что несовместные события противоположны? Верно ли такое утверждение хотя бы в каких-либо случаях; если да, то в каких именно?

4. Что такое полная группа событий? В каких случаях употребляется термин «элементарное событие»?

5. Сформулируйте классическое определение вероятности.

6. Является ли в теореме сложения вероятностей несовместность событий необходимым условием справедливости теоремы?

7. Как найти вероятность противоположного события $\bar{A}$ по известной вероятности некоторого события A?

8. Что такое условная вероятность?

9. Когда события называются независимыми? Приведите примеры зависимых и независимых событий.

10. В чем различие между теоремами умножения для зависимых и независимых событий?

11. Обоснуйте равенство

$$P(A)\, P_A(B) = P(B)\, P_B(A).$$

В чем состоит вероятностный смысл этого равенства? Объясните все используемые здесь обозначения.

12. Можно ли перенести теоремы сложения и умножения вероятностей на случай нескольких событий? Как они формулируются?

13. Справедлива ли теорема сложения вероятностей в случае совместных событий? Как она формулируется? Можно ли считать, что теорема сложения для несовместных событий является частным случаем последней? Как это доказать?

14. Что такое «полная вероятность»? В каких случаях применяется формула полной вероятности?

15. Для чего служит формула Бейеса? В каких случаях она может быть применена?

16. Что такое биномиальная формула вероятностей и почему она так называется?

17. В каких пределах заключено наиболее вероятное число наступлений события μ в схеме Бернулли?

18. В чем разница между возвратной и безвозвратной выборкой? Как ведет себя вероятность при безвозвратной выборке при неограниченном увеличении объема совокупности? Можно ли ожидать такого поведения вероятности без вычислений и предельного перехода?

19. Почему задачу о безвозвратной выборке можно рассматривать как обобщение схемы Бернулли? В чем общность и различие таких задач?

20. Какая формула называется полиномиальной формулой вероятностей? Что выражает полиномиальная формула и почему она так называется?

21. Докажите, что биномиальная формула является частным случаем полиномиальной.

22. Приведите выражение производящей функции для биномиальной формулы вероятностей.

23. Что такое цепь Маркова? Почему цепь Маркова можно рассматривать как обобщение схемы Бернулли?

24. Что называется вероятностью перехода и матрицей перехода? Каким условиям должны удовлетворять вероятности перехода p_{ij}?

25. В чем состоит теорема Маркова о предельных вероятностях? Как можно истолковать физически предельные вероятности p_j?

26. Почему классическое определение вероятности оказывается недостаточным? Приведите примеры, в которых полная группа событий оказывается бесконечной.

27. В чем состоит геометрическое определение вероятности? Приведите примеры его использования.

28. В чем состоит статистическое определение вероятностей? Приведите примеры статистических закономерностей из физики и других областей естествознания.

29. Что называется суммой и произведением событий?

30. В чем смысл аксиоматического построения теории вероятностей?

ГЛАВА II

АСИМПТОТИЧЕСКИЕ ФОРМУЛЫ

§ 9. ЛОКАЛЬНАЯ ТЕОРЕМА МУАВРА — ЛАПЛАСА

Рассмотрим снова (ср. § 4) задачу о повторении испытаний производится серия из n независимых испытаний, в каждом из которых появление события A имеет вероятность p. Найдем для вероятности появления события A ровно m раз при n испытаниях *асимптотическую формулу,* т. е. приближенную формулу (удобную для практических вычислений), которая при достаточно большом числе испытаний дает сколь угодно малую относительную погрешность. Поставленная нами задача решается теоремой, которая носит название *локальной предельной теоремы Муавра—Лапласа.*

Локальная предельная теорема. *Если вероятность наступления события A при каждом отдельном испытании постоянна и равна p, то вероятность наступления события A ровно m раз в серии из n испытаний может быть представлена в виде*

$$P_{m,n}=\frac{1}{\sqrt{2\pi npq}}e^{-\frac{x^2}{2}}(1+\varepsilon_n),\tag{1}$$

где $q=1-p$, $x=\dfrac{m-np}{\sqrt{npq}}$ и $\varepsilon_n\to 0$ *при неограниченном возрастании n, m и $n-m$.*

Для достаточно больших n формула (1) дает

$$P_{m,n}\approx\frac{1}{\sqrt{2\pi npq}}e^{-\frac{x^2}{2}}.\tag{2}$$

Относительная погрешность этой приближенной формулы стремится к нулю при неограниченном возрастании n. Применение формулы (2) требует для вычисления $P_{m,n}$ знания значений функции

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Ввиду важности этой функции* для нее составлены специальные таблицы, которые приведены в приложении. Ясно, что функция $\varphi(x)$ — четная.

Доказательство теоремы. Воспользуемся *формулой Стирлинга*, дающей асимптотическое выражение для факториалов больших чисел**. Она имеет вид:

$$n! = n^n e^{-n} \sqrt{2\pi n}\,(1 + \alpha_n), \tag{3}$$

где величина α_n стремится к нулю при неограниченном возрастании n.

Рассмотрим полученное в § 4 выражение (1) для вероятности $P_{m,n}$,

$$P_{m,n} = \frac{n!}{m!\,(n-m)!}\,p^m q^{n-m},$$

и заменим в нем факториалы по формуле Стирлинга. Тогда

$$P_{m,n} = \frac{n^m e^{-n} \sqrt{2\pi n}\,(1+\alpha_n)\,p^m q^{n-m}}{m^m e^{-m} \sqrt{2\pi m}\,(1+\alpha_m)\,(n-m)^{n-m} e^{-n+m} \sqrt{2\pi (n-m)}\,(1+\alpha_{n-m})}$$

или

$$P_{m,n} = \frac{n^n p^m q^{n-m}}{m^n (n-m)^{n-m}} \sqrt{\frac{n}{2\pi m\,(n-m)}}\left(1+\alpha_n'\right), \tag{4}$$

где мы положили

$$1+\alpha_n' = \frac{1+\alpha^n}{(1+\alpha_m)(1+\alpha_{n-m})}\,.$$

Числа m и $n-m$ неограниченно возрастают вместе с n***. Тогда α_n, α_m, α_{n-m} стремятся к нулю, так что $1+\alpha_n' \to 1$ и $\alpha_n' \to 0$.

Обозначим первый множитель в правой части равенства (4) через $\frac{1}{H}$, т. е. положим

$$\frac{1}{H} = \frac{n^n p^m q^{n-m}}{m^n (n-m)^{n-m}} = \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m}.$$

---

* Эта функция будет подробнее рассматриваться в следующем параграфе.

** Доказательство формулы Стирлинга можно найти в подробных курсах математического анализа. См., например: Г. М. Фихтенгольц, Курс дифференциального и интегрального исчисления, т. II.

*** Легко понять, что если число m не возрастает вместе с n, то вероятность $P_{m,n}$ при конечном p стремится к нулю.

Далее, введем обозначение

$$\frac{m-np}{\sqrt{npq}}=x.$$

Тогда

$$m=np+x\sqrt{npq},$$

$$n-m=n-np-x\sqrt{npq}=nq-x\sqrt{npq}.$$

Используя полученные выражения для m и $n-m$ через x, получим для величины H следующее выражение:

$$H=\left(\frac{n}{np}\right)^m\left(\frac{n-m}{nq}\right)^{n-m}=\left(1+x\sqrt{\frac{q}{np}}\right)^m\left(1-x\sqrt{\frac{p}{np}}\right)^{n-m}.$$

Логарифмируя полученное выражение для H, получаем:

$$\ln H=m\ln\left(1+x\sqrt{\frac{q}{np}}\right)+(n-m)\ln\left(1-x\sqrt{\frac{p}{nq}}\right),$$

или, еще раз воспользовавшись выражениями m и $n-m$,

$$\ln H=(np+x\sqrt{npq})\ln\left(1+x\sqrt{\frac{q}{np}}\right)+$$
$$+(nq-x\sqrt{npq})\ln\left(1-x\sqrt{\frac{p}{nq}}\right). \tag{5}$$

Выберем фиксированное значение x (что означает выбор определенной зависимости между m и n) и предположим n настолько большим, чтобы $\left|x\sqrt{\frac{q}{np}}\right|<\frac{1}{2}$ и $\left|x\sqrt{\frac{p}{nq}}\right|<\frac{1}{2}$, что возможно, так как n находится в знаменателе. Для логарифмов, входящих в формулу (5), можно тогда воспользоваться известным рядом

$$\ln(1+t)=t-\frac{1}{2}t^2+\frac{1}{3}t^3-\ldots,$$

сходящимся при $|t|<1$, полагая соответственно $t=x\sqrt{\frac{q}{np}}$ и $t=x\sqrt{\frac{p}{np}}$

Получим:

$$\left.\begin{aligned}\ln\left(1+x\sqrt{\frac{q}{np}}\right)&=x\sqrt{\frac{q}{np}}-\frac{x^2}{2}\frac{q}{np}+\rho_1,\\ \ln\left(1-x\sqrt{\frac{p}{nq}}\right)&=-x\sqrt{\frac{p}{nq}}-\frac{x^2}{2}\frac{p}{nq}+\rho_2.\end{aligned}\right\} \tag{6}$$

Оценим остаточные члены ρ_1 и ρ_2 рядов (6). Для первого из них находим:

$$\rho_1=\frac{1}{3}\left(x\sqrt{\frac{q}{np}}\right)^3-\frac{1}{4}\left(x\sqrt{\frac{q}{np}}\right)^4+\ldots,$$

так что

$$|\rho_1|<\frac{1}{3}\left|x\sqrt{\frac{q}{np}}\right|^3+\frac{1}{4}\left|x\sqrt{\frac{q}{np}}\right|^4+\ldots$$

Отбросив справа числовые множители, меньшие единицы, мы только усилим неравенство. При этом справа получится геометрическая прогрессия со знаменателем $q^* = \left| x \sqrt{\frac{q}{np}} \right| < \frac{1}{2}$. Поэтому

$$|\rho_1| < \left| x \sqrt{\frac{q}{np}} \right|^3 \cdot \frac{1}{1 - q^*} < 2 \left(\frac{q}{p} \right)^{\frac{3}{2}} \frac{x^3}{n^{3/2}}.$$

Аналогично получаем оценку для остаточного члена второго ряда:

$$|\rho_2| < 2 \left(\frac{p}{q} \right)^{\frac{3}{2}} \frac{x^3}{n^{3/2}}.$$

Если обозначить через A наибольшее из двух чисел $2\left(\frac{q}{p}\right)^{\frac{3}{2}}$ и $2\left(\frac{p}{q}\right)^{\frac{3}{2}}$, то можно написать общую оценку для обоих остаточных членов:

$$|\rho_1| < A \frac{x^3}{n^{3/2}}, \quad |\rho_2| < A \frac{x^3}{n^{3/2}}. \tag{7}$$

Подставив ряды (6) в выражение (5) для $\ln H$, получим формулу

$$\ln H = (np + x\sqrt{npq}) \left(x \sqrt{\frac{q}{np}} - \frac{x^2}{2} \frac{q}{np} + \rho_1 \right) +$$
$$+ (nq - x\sqrt{npq}) \left(-x \sqrt{\frac{p}{nq}} - \frac{x^2}{2} \frac{p}{nq} + \rho_2 \right),$$

откуда после простых алгебраических преобразований найдем:

$$\ln H = \frac{x^2}{2} + \frac{x^3}{2} \left(-q \sqrt{\frac{q}{np}} + p \sqrt{\frac{p}{nq}} \right) + \rho_1 (np +$$
$$+ x\sqrt{npq}) - \rho_2 (nq - x\sqrt{npq}) = \frac{x^2}{2} + \delta_n.$$

Мы положили здесь

$$\delta_n = \frac{x^3}{2} \left(-q \sqrt{\frac{q}{np}} + p \sqrt{\frac{p}{nq}} \right) + \rho_1 (np + x\sqrt{npq}) - \rho_2 (nq - x\sqrt{npq}).$$

Оценим величину δ_n, для чего воспользуемся известным свойством абсолютной величины суммы и оценками (7) для величин ρ_1 и ρ_2. Тогда

$$|\delta_n| < \frac{|x^3|}{2} \left| -q \sqrt{\frac{q}{np}} + p \sqrt{\frac{p}{np}} \right| + A \frac{|x^3|}{n^{3/2}} |np + x\sqrt{npq}| +$$
$$+ A \frac{|x^3|}{n^{3/2}} |nq - x\sqrt{npq}|.$$

Вынесем из всех членов за скобку множитель $\frac{|x^3|}{\sqrt{n}}$ и запишем неравенство в виде

$$|\delta_n| = \frac{|x^3|}{\sqrt{n}} \left\{ \left| \frac{1}{2} \left(-q \sqrt{\frac{q}{p}} \right) \right| + A \left| p + \frac{x\sqrt{pq}}{n} \right| + A \left| q - \frac{x\sqrt{pq}}{n} \right| \right\}.$$

Так как n неограниченно возрастает, то члены вида $\frac{x\sqrt{pq}}{n}$ могут быть сделаны меньше любого фиксированного числа. Поэтому выражение, стоящее в фигурных скобках, не превосходит фиксированного числа B, не зависящего от x и n, а только от p и q. Следовательно, мы можем написать:

$$|\delta_n| < B \frac{|x^3|}{\sqrt{n}}. \tag{8}$$

Итак,

$$\ln H = \frac{x^2}{2} + \delta_n,$$

где величина δ_n оценивается неравенством (8). Отсюда следует

$$H = e^{\frac{x^2}{2} + \delta_n} = e^{\frac{x^2}{2}} (1 + \gamma_n), \tag{9}$$

где $\gamma_n = e^{\delta_n} - 1$. Так как, в силу неравенства (8), имеем $\delta_n \to 0$ при $n \to \infty$, то и $\gamma_n \to 0$.

Мы изучили первый множитель выражения (4) для вероятности $P_{m,n}$. Остается рассмотреть второй множитель, в котором мы выпишем отдельно часть подкоренного выражения

$$\frac{n}{m(n-m)}.$$

Заменяя здесь m и $n-m$ их выражениями через x, получаем:

$$\frac{n}{m(n-m)} = \frac{n}{(np + x\sqrt{pq})(nq - x\sqrt{npq})} =$$

$$= \frac{1}{n\left(p + x\sqrt{\frac{pq}{n}}\right)\left(q - x\sqrt{\frac{pq}{n}}\right)}.$$

Раскрыв скобки, представим знаменатель этой дроби в виде

$$n\left[pq + x(q-p)\sqrt{\frac{pq}{n}} - x^2\frac{pq}{n}\right] = npq\left(1 + x\frac{q-p}{\sqrt{npq}} - \frac{x^2}{n}\right) = npq(1 + \beta_n),$$

где $\beta = x\dfrac{q-p}{\sqrt{npq}} - \dfrac{x^2}{n}$, причем для фиксированного x и $n \to \infty$ имеем $\beta_n \to 0$. Таким образом,

$$\frac{n}{m(n-m)} = \frac{1}{npq(1+\beta_n)}. \tag{10}$$

Подставляя в равенство (4) полученные равенства (9) и (10), придем к выражению для вероятности:

$$P_{m,n} = \frac{e^{-\frac{x^2}{2}}}{1+\gamma_n} \cdot \frac{1+\alpha_n'}{\sqrt{2\pi npq(1+\beta_n)}}$$

или окончательно:

$$P_{m,n} = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{x^2}{2}} (1 + \varepsilon_n),$$

где положено $1+\varepsilon_n=\frac{1+\alpha'_n}{(1+\gamma_m)\sqrt{1+\beta_n}}$. Из поведения α'_n, β_n и γ_n следует, что $\varepsilon_n \to 0$ при $n\to\infty$ для всех целых $m = np + x\sqrt{npq}$ при любом фиксированном x, взятом из фиксированного конечного интервала.

Пример 1. Мы можем воспользоваться формулой (2) для окончательного решения задачи, приведенной в примере 4 из § 4. Там требовалось определить вероятность того, что в партии из 10 000 изделий встретится ровно 40 бракованных, если вероятность быть бракованной для отдельной детали равна 0,005.

Здесь $p = 0{,}005$, $q = 0{,}995$, $n = 10\,000$, $m = 40$. Поэтому

$$\sqrt{npq}=\sqrt{10000\cdot 0{,}005\cdot 0{,}995}\approx 7{,}05,$$
$$x=\frac{m-np}{npq}=\frac{40-10000\cdot 0{,}005}{7{,}05}\approx -1{,}42.$$

В силу четности функции $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ отыскиваем ее значение при $x = 1{,}42$. По таблице получаем $\varphi(1{,}42) = 0{,}1456$. Таким образом,

$$P_{40,10000}\approx\frac{0{,}1456}{7{,}05}\approx 0{,}0206.$$

Подсчет по точной формуле дает $P_{40,10000} = 0{,}0197$. Наша приближенная формула дает в этом случае ошибку в 0,0009, что составляет около 4,5%. Такая ошибка, однако, вполне искупается серьезным облегчением вычислений.

Пример 2. Произведено 10 000 испытаний, в каждом из которых вероятность появления события A равна $p = 0{,}5$. Требуется приближенно найти вероятность наиболее вероятного числа появлений события A.

Как следует из рассуждений, приведенных в § 4, наиболее вероятным числом появления событий будет здесь $m = np = 5000$. Таким образом, получаем:

$$p = 0{,}5,\qquad q = 0{,}5,\ n = 10000,\ m = 5000,$$
$$\sqrt{npq}=\sqrt{10000\cdot 0{,}5\cdot 0{,}5}=50,$$
$$x=\frac{n-np}{\sqrt{npq}}=0,\quad \varphi(0)=\frac{1}{\sqrt{2\pi}}=0{,}3989.$$

Полученные числовые данные приводят к вероятности

$$P_{5000,10000}=\frac{0{,}3989}{50}=0{,}007978.$$

Как было указано, число $m = 5000$ является наиболее вероятным числом наступлений события A, однако вероятность появления точно этого числа весьма мала. Для сравнения приведем некоторые вероятности $P_{m,n}$ для других m в тех же условиях:

$$P_{4950,10000} = P_{5050,10000} = 0{,}004894 \quad (x = 1),$$
$$P_{4900,10000} = P_{5100,10000} = 0{,}001108 \quad (x = 2).$$

Пример 3. Игральная кость (кубик) бросается 1200 раз. Определить вероятность наиболее вероятного числа выпадений шестерки.

Вероятность наступления интересующего нас события — выпадения шестерки — при одном испытании равна $p = \frac{1}{6}$. Неравенство (9) из § 4 показывает, что наиболее вероятным числом наступлений события является $m = np = 200$. Как и в предыдущем примере, можем написать:

$$p = \frac{1}{6}, \quad q = \frac{5}{6}, \quad n = 1200, \quad m = 200.$$

Тогда

$$\sqrt{npq} = \sqrt{200 \cdot \frac{1}{6} \cdot \frac{5}{6}} = \frac{5\sqrt{10}}{3} \approx 5{,}28,$$

$$x = \frac{m - np}{\sqrt{npq}} = 0, \quad \varphi(0) = \frac{1}{\sqrt{2\pi}} = 0{,}3989.$$

Отсюда находим вероятность

$$P_{200,1200} = \frac{0{,}3989}{5{,}28} \approx 0{,}0757.$$

Для сравнения укажем, что при $m = 210$, т. е. при

$$x = \frac{210 - 200}{5{,}28} = 1{,}9,$$

получаем $\varphi(x) = 0{,}0656$, откуда вероятность 210 выпадений шестерки

$$P_{210,1200} = \frac{0{,}0656}{5{,}28} \approx 0{,}0124.$$

§ 10. НОРМАЛЬНАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

При разборе примеров в § 9 нам пришлось пользоваться функцией $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Эту функцию называют *плотностью вероятности нормального распределения.* Причины этого названия будут выяснены позже (см. § 16).

Построим график функции $y = \varphi(x)$. Так как $\varphi(x)$ — четная функция, то ее график симметричен относительно оси Oy. Кривая имеет ось Ox горизонтальной асимптотой при $x \to \pm\infty$. Функция достигает максимума при $x = 0$, причем $\varphi(0) = \frac{1}{\sqrt{2\pi}} \approx 0{,}3989$. График функции $\varphi(x)$ приведен на рис. 7, где для большего удобства приняты различные масштабы по осям.

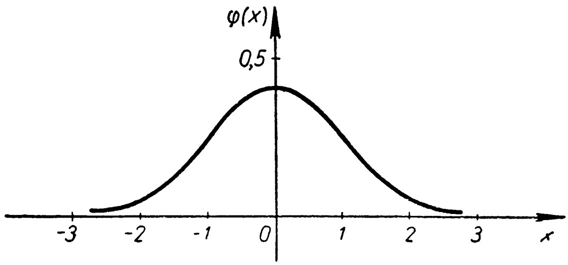


Рис. 7

Площадь, ограниченная этой кривой и осью Ox, равна единице:

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

Чтобы в этом убедиться, вычислим интеграл

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx = 2\int_{0}^{+\infty} e^{-\frac{x^2}{2}}\,dx.$$

Положим

$$I = \int_{0}^{\infty} e^{-\frac{x^2}{2}}\,dx.$$

Вычисляя несобственный двойной интеграл

$$I_1 = \iint_D e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy,$$

где область D означает первую четверть плоскости xOy, получаем:

$$I_1 = \iint_D e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy = \int_0^{\infty} dx\left(\int_0^{\infty} e^{-\frac{1}{2}(x^2+y^2)}\,dy\right) = \int_0^{\infty} e^{-\frac{x^2}{2}}\left(\int_0^{\infty} e^{-\frac{y^2}{2}}\,dy\right)dx = I^2,$$

так как внутренний интеграл не зависит от x и его величина не изменяется при изменении обозначения переменной интегрирования.

С другой стороны, вычисляя тот же двойной интеграл переходом к полярным координатам, получаем из

$$x = \rho\cos\varphi,$$
$$y = \rho\sin\varphi$$

известные соотношения:

$$x^2+y^2=\rho^2,$$
$$dxdy=\rho d\rho d\varphi;$$

следовательно,

$$I_1=\iint\limits_D e^{-\frac{1}{2}(x^2+y^2)}dxdy=\int\limits_0^{\frac{\pi}{2}}d\varphi\int\limits_0^{\infty}e^{-\frac{1}{2}\rho^2}\rho d\rho=-\frac{\pi}{2}e^{-\frac{1}{2}\rho^2}\Big|_0^{\infty}=\frac{\pi}{2}.$$

Сравнивая полученные результаты, находим:

$$I^2=\frac{\pi}{2},$$

откуда

$$I=\sqrt{\frac{\pi}{2}}.$$

Поэтому

$$\int\limits_{-\infty}\varphi(x)\,dx=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}2I=1.$$

Еще бо́льшую роль, нежели функция $\varphi(x)$, играет интеграл от нее, взятый в пределах от 0 до x, т. е. функция

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt.$$

Эту функцию называют *нормальной функцией распределения*. Иногда употребляются также названия *функция Лапласа*, или *интеграл вероятности*.

Функция $\Phi(x)$ является нечетной. Действительно, положив $t=-\tau$, получим:

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt=-\frac{1}{\sqrt{2\pi}}\int\limits_0^{-x}e^{-\frac{\tau^2}{2}}d\tau=-\Phi(-x).$$

Кривая $y=\Phi(x)$ симметрична поэтому относительно начала координат. Она имеет две горизонтальные асимптоты, ибо

$$\lim_{x\to-\infty}\Phi(x)=-\frac{1}{2},$$
$$\lim_{x\to+\infty}\Phi(x)=\frac{1}{2}.$$

При $x=0$ имеем $\Phi(0)=0$. Начало координат служит точкой перегиба кривой, причем угловой коэффициент касательной в точке перегиба равен $k=\frac{1}{\sqrt{2\pi}}\approx 0{,}3989$.

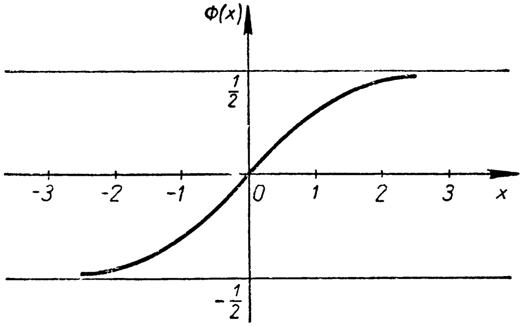


Рис. 8

График функции $y = \Phi(x)$ изображен на рис. 8.

В некоторых случаях взамен функции $\Phi(x)$ рассматривается функция $F(x)$, определяемая равенством

$$F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

Между этими функциями существует очевидное соотношение

$$F(x) = \Phi(x) + \frac{1}{2}.$$

Ввиду большой важности функций $\varphi(x)$ и $\Phi(x)$ для них составлены подробные таблицы, которые приведены в приложении. Вследствие четности $\varphi(x)$ и нечетности $\Phi(x)$ в таблице даны лишь положительные значения аргумента.

§ 11. ТЕОРЕМА ПУАССОНА

В § 9 мы получили асимптотическую формулу для вероятности $P_{m,n}$ наступления события m раз при n испытаниях. Эта формула дает при одном и том же n тем лучшие результаты, чем ближе число p (вероятность наступления события при отдельном испытании) к половине. Однако при достаточно малых p (т. е. при рассмотрении редких событий) ошибка, даваемая этой формулой, может быть уже довольно значительной.

Рассмотрим случай, когда вероятность p положительного исхода каждого испытания в серии из n испытаний равна $\frac{\lambda}{n}$, где λ — некоторая постоянная величина*, и выведем в этом случае новую

* Как мы увидим в следующей главе (см. § 16), коэффициент λ имеет определенный вероятностный смысл.

приближенную формулу для $P_{m,n}$. Справедлива следующая теорема.

Теорема Пуассона. *Пусть в серии из n испытаний вероятность появления события A в каждом испытании равна $\frac{\lambda}{n}$. Тогда вероятность появления события A в этой серии m раз при большом n выражается приближенной формулой*

$$P_{m,n} \approx \frac{\lambda^m}{m!} e^{-\lambda}. \tag{1}$$

Доказательство. Как мы уже знаем из § 4 (формула (1)),

$$P_{m,n} = \frac{n!}{m!\,(n-m)!} p^m q^{n-m} = \frac{n!}{m!\,(n-m)!} p^m (1-p)^{n-m}.$$

Положим здесь $p = \frac{\lambda}{n}$. Тогда

$$P_{m,n} = \frac{n!}{m!\,(n-m)!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m},$$

или, после алгебраических преобразований,

$$P_{m,n} = \frac{n(n-1)\dots(n-m+1)(n-m)\dots 1\cdot\lambda^m}{m!\,(n-m)(n-m-1)\dots 1\cdot n^m}\left(1-\frac{\lambda}{n}\right)^{n-m} =$$

$$= \frac{\lambda^m}{m!}\cdot 1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-m}.$$

Найдем теперь предел $P_{m;n}$ при $n \to \infty$ и постоянном m. Так как пределы множителей $\left(1-\frac{1}{n}\right), \dots, \left(1-\frac{m-1}{n}\right)$ и $\left(1-\frac{\lambda}{n}\right)^{-m}$ равны единице, а $\lim\limits_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$,

$$\text{то } \lim_{n\to\infty} P_{n,m} = \frac{\lambda^m}{m!} e^{-\lambda},$$

откуда и вытекает приближенное равенство (1).

Пример 1. Работница обслуживает 800 веретен. Вероятность обрыва пряжи на каждом из них за промежуток времени τ равна 0,005. Найти наиболее вероятное число обрывов и его вероятность.

Наиболее вероятное число обрывов будет $\mu = np = 4$. Точное значение вероятности четырех обрывов равно:

$$P_{4,800} = C_{800}^4 (0{,}005)^4 \cdot (0{,}995)^{796}.$$

Пользуясь формулой Пуассона с $\lambda = np = 4$, получаем:

$$P_{4,800} \approx \frac{4^4}{4!} e^{-4} = \frac{256}{24}\cdot 0{,}0183 = 0{,}1954.$$

Вычисление по точной формуле дает 0,1945, так что ошибка при пользовании формулой Пуассона составляет 0,009. Локальная предельная теорема Муавра—Лапласа дает для данного случая

$$P_{4,800} \approx \frac{1}{\sqrt{2\pi \cdot 800 \cdot 0{,}005 \cdot 0{,}995}}\, e^{-\frac{x^2}{2}} \approx 0{,}2000,$$

ибо здесь $x = 0$, так что ошибка составляет уже 0,0055, т. е. в шесть раз больше, чем при использовании формулы Пуассона.

Пример 2. Вероятность любому абоненту позвонить на коммутатор в течение часа равна 0,01. Телефонная станция обслуживает 300 абонентов. Какова вероятность, что в течение часа позвонит ровно 4 абонента.

Здесь можно снова воспользоваться законом Пуассона, полагая $n = 300$, $m = 4$ и $\lambda = np = 3$. Тогда искомая вероятность равна:

$$P_{3.300} \approx \frac{3^4}{4!} e^{-3} \approx 0{,}169.$$

§ 12. ИНТЕГРАЛЬНАЯ ТЕОРЕМА МУАВРА — ЛАПЛАСА. ТЕОРЕМА БЕРНУЛЛИ

Примеры, разобранные в § 9, показывают, что при больших n вероятность появления события A определенное число раз весьма мала. В практических приложениях нас обычно интересует не наступление события A какое-то определенное число раз, а вероятность того, что число наступлений события A заключено в некоторых пределах. Как было отмечено еще в примере 4 из § 4, эту вероятность можно получить путем суммирования, однако это требует громоздких вычислений.

Интегральная предельная теорема, которую мы сейчас рассмотрим, дает возможность подсчитать эту вероятность значительно проще, подобно тому как это делалось для вероятности $P_{m,n}$ с помощью локальной теоремы, разобранной в § 9.

Интегральная теорема Муавра—Лапласа. *Если вероятность наступления события A при каждом отдельном испытании равна p, то при неограниченном увеличении числа испытаний n вероятность того, что число m наступлений события A удовлетворяет неравенству*

$$t_0 < \frac{m - np}{\sqrt{npq}} < t_1, \qquad (1)$$

где $q = 1 - p$, имеет своим пределом интеграл

$$\frac{1}{\sqrt{2\pi}} \int_{t_0}^{t_1} e^{-\frac{x^2}{2}}\, dx. \qquad (2)$$

Доказательство. Как было указано выше, искомая вероятность может быть получена суммированием:

$$P\left(t_0 < \frac{m-np}{\sqrt{npq}} < t_1\right) = \sum_m P_{m,n}, \tag{3}$$

где сумма распространяется на все значения m, удовлетворяющие неравенству (1). Эти значения можно записать в виде

$$m = np + x\sqrt{npq}, \tag{4}$$

где x принимает значения, лежащие между t_0 и t_1. Из (4) следует, что при неограниченном возрастании n неограниченно возрастает также и m.

Так как формула (4) выражает m через x, то суммирование в (3) распространяется на некоторые значения x из участка (t_0, t_1). Подсчитаем приращение x при переходе от значения m к значению $m+1$ для одного и того же числа испытаний n. Пусть

$$m + 1 = np + (x + \Delta x)\sqrt{npq}. \tag{5}$$

Вычитая из равенства (5) равенство (4), получаем:

$$1 = \Delta x\sqrt{npq},$$

или

$$\Delta x = \frac{1}{\sqrt{npq}}.$$

Пользуясь этим выражением, перепишем выражение вероятности $P_{m,n}$ (см. формулу (2) из § 9) в виде

$$P_{m,n} \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Delta x. \tag{6}$$

Выражение (6) для вероятности $P_{m,n}$ допускает простую геометрическую иллюстрацию. Действительно, построим кривую $y = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (плотность вероятности нормального распределения).

При заданных n, m, p и q вычислим по формуле (4) значение x и построим на кривой $y = \varphi(x)$ точку M с абсциссой x (рис. 9). Затем выберем $\Delta x = \frac{1}{\sqrt{npq}}$ и построим прямоугольник (на рис. 9 он заштрихован). Площадь этого прямоугольника будет равна $\varphi(x)\Delta x = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Delta x$. Таким образом, согласно формуле (6) вероятность $P_{m,n}$ приближенно выражается площадью заштрихованного на рис. 9 прямоугольника.

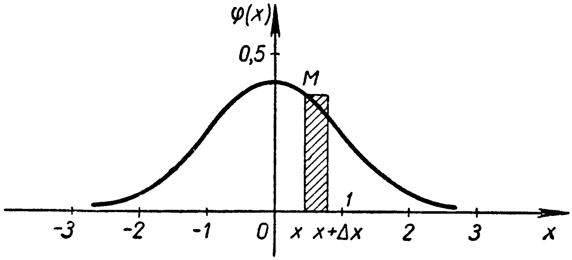


Рис. 9

Подставим выражение (6) в сумму (3). Тогда вероятность неравенства (1) приближенно выразится суммой

$$P\left(t_0 < \frac{m}{\sqrt{npq}} < t_1\right) \approx \sum_i \varphi\,(x_i)\;\Delta\,x. \tag{7}$$

Сумма в правой части равенства (7) представляет собой интегральную сумму, построенную для функции $\varphi(x)$ на участке $(t_0,\ t_1)$, при разбиении этого участка на равные части, длины которых $\Delta\,x = \frac{1}{\sqrt{npq}}$. При неограниченном возрастании n величина $\Delta\,x = \frac{1}{\sqrt{npq}}$ стремится к нулю, и поэтому предел этой суммы равен интегралу (2).

Можно показать, что при этом ошибка приближенного равенства (7) стремится к нулю* .

Таким образом, при неограниченном возрастании n вероятность неравенства (1) стремится к интегралу (2). Поэтому при больших значениях n можно считать, что

$$P\left(t_0 < \frac{m - np}{\sqrt{npq}} < t_1\right) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{t_0}^{t_1} e^{-\frac{x^2}{2}}\,dx. \tag{8}$$

Заметим, что

$$\frac{1}{\sqrt{2\pi}} \int\limits_{t_0}^{t_1} e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{t_1} e^{-\frac{x^2}{2}}\,dx - \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{t_2} e^{-\frac{x^2}{2}}\,dx =$$

$$= \Phi\,(t_1) - \Phi\,(t_0).$$

* Точное доказательство сходимости суммы в правой части равенства (7) к интегралу приводится в более подробных курсах теории вероятностей. См., например: Б. В. Гнеденко, Курс теории вероятностей.

В частности, если $t_0 = -t_1$, то

$$\frac{1}{\sqrt{2\pi}} \int_{-t_1}^{t_1} e^{-\frac{x^2}{2}} dx = \Phi(t_1) - \Phi(-t_1) = 2\Phi(t_1). \qquad (9)$$

Значения функции $\Phi(x)$ могут быть найдены в таблице. Это дает возможность легко находить приближенные значения искомой вероятности интересующего нас неравенства.

Пример 1. Вернемся к рассмотрению второй части примера 4 из § 4. В ней требовалось определить вероятность того, что в партии из 10 000 деталей окажется не более 70 бракованных, если для каждой детали вероятность оказаться бракованной равна 0,005.

Здесь $p = 0{,}005$, $q = 0{,}995$, $n = 10\,000$, $\sqrt{npq} \approx 7{,}05$, $np = 50$. Поэтому неравенство $0 \leqslant m \leqslant 70$, вероятность которого нас интересует, равносильно неравенству $-50 \leqslant m - np \leqslant 20$, или

$$-7{,}09 \leqslant \frac{m - np}{\sqrt{npq}} \leqslant 2{,}84.$$

В силу доказанной теоремы,

$$P\left(-7{,}09 \leqslant \frac{m - np}{\sqrt{npq}} \leqslant 2{,}84\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-7{,}09}^{2{,}84} e^{-\frac{x^2}{2}} dx.$$

Пользуясь равенством (8), получаем:

$$P = \Phi(2{,}84) - \Phi(-7{,}09) = \Phi(2{,}74) + \Phi(7{,}09) = \\ = 0{,}4977 + 0{,}5 = 0{,}9977.$$

Таким образом, вероятность того, что число бракованных деталей не превзойдет 70, очень мало отличается от единицы.

Пример 2. Определим вероятность P того, что при 8000 бросаний игральной кости частота выпадения шестерки будет отклоняться от вероятности $p = \frac{1}{6}$ меньше, чем на $\frac{1}{80}$.

Нас интересует вероятность неравенства

$$\left| \frac{m}{8000} - \frac{1}{6} \right| < \frac{1}{80},$$

где m — число выпадений шестерки, $p = \frac{1}{6}$, $q = \frac{5}{6}$, $n = 8000$.

Приведенное неравенство можно переписать в виде:

$$-100 < m - 8000 \cdot \frac{1}{6} < 100.$$

Последнее неравенство равносильно такому:

$$-\frac{100}{\sqrt{8000\cdot\frac{1}{6}\cdot\frac{5}{6}}}<\frac{m-800\cdot\frac{1}{6}}{\sqrt{npq}}<\frac{100}{\sqrt{8000\cdot\frac{1}{6}\cdot\frac{5}{6}}},$$

откуда следует, что $t_0 = -3$ и $t_1 = 3$. Поэтому, в силу формулы (9), искомая вероятность равна:

$$P = 2\Phi\,(3) = 2\cdot 0{,}49865 = 0{,}9973$$

(см. таблицу значений функции $\Phi\,(x)$ в приложении).

Пример 3. Вероятность попадания в цель при отдельном выстреле равна 0,63. Сколько выстрелов нужно произвести, чтобы с вероятностью 0,9 получить не менее 10 попаданий?

Требование получить не менее 10 попаданий равносильно неравенству

$$10\leqslant m< +\infty,$$

в котором верхняя граница не определена. Перепишем это неравенство в нужной форме. Тогда

$$\frac{10-0{,}63n}{\sqrt{n\cdot 0{,}63\cdot 0{,}37}}<\frac{m-0{,}63m}{\sqrt{n\cdot 0{,}63\cdot 0{,}37}}<+\infty,$$

откуда следует $t_0=\frac{10-0{,}63n}{\sqrt{0{,}23\,n}}$ и $t_1=\infty$. Вероятность $P\left(t_0<\frac{m-np}{\sqrt{npq}}<t_1\right)$ в данном случае по условию равна 0,9. С другой стороны, по формуле (8) имеем:

$$P\,(t_0<x<t_1)=\Phi\,(t_1)-\Phi\,(t_0).$$

Так как $t_1 = +\infty$ и $\Phi\,(\infty) = 0{,}5$, то $\Phi\,(t_0) = 0{,}4000$, откуда, пользуясь таблицей функции $\Phi\,(x)$, находим $t_0 = -1{,}28$, что дает для определения n уравнение

$$\frac{10-0{,}63n}{\sqrt{0{,}23n}}=-1{,}28.$$

Решив это уравнение и заметив, что с увеличением n вероятность получить не менее 10 попаданий может лишь возрасти, получаем окончательно:

$$n\geqslant 20.$$

В § 8 указывалось, что наблюдаемые на практике частоты появления некоторых событий тем меньше отличаются от вероятности, чем больше число испытаний. Мы можем сейчас выяснить этот вопрос с теоретической стороны.

Пусть производится n независимых испытаний, причем в каждом из них вероятность появления некоторого события A равна p. Если через m обозначить число испытаний, в результате которых событие A появлялось, то отношение $\frac{m}{n}$ представит частость события A. Имеет место следующая теорема.

Теорема Бернулли. *Для всякого* $\varepsilon > 0$ *вероятность неравенства*

$$\left|\frac{m}{n} - p\right| < \varepsilon$$

при неограниченном возрастании n стремится к единице.

Иными словами, при достаточно большом числе испытаний с вероятностью, как угодно близкой к единице, частота появления события мало отличается от вероятности события.

Доказательство. Интересующее нас неравенство равносильно следующему:

$$-\varepsilon < \frac{m - np}{n} < \varepsilon.$$

Умножив все части неравенства на $\sqrt{\frac{n}{pq}}$, приведем его к нужной нам форме:

$$-\varepsilon\sqrt{\frac{n}{nq}} < \frac{m - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}.$$

Таким образом,

$$P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = P\left\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{m - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\right\}.$$

В силу интегральной предельной теоремы, эта вероятность близка к интегралу $\frac{1}{\sqrt{2\pi}}\int\limits_{-\alpha}^{\alpha} e^{-\frac{x^2}{2}}\,dx$, где $\alpha = \varepsilon\sqrt{\frac{n}{pq}}$.

При неограниченном возрастании n имеем также $\alpha \to \infty$, поэтому вероятность интересующего нас неравенства стремится к интегралу

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx,$$

который, как это было показано в § 10, равен единице. Теорема доказана.

Теорема Бернулли является одним из частных случаев так называемого *закона больших чисел.* Некоторые другие его формы будут рассмотрены в главе IV (см. § 23).

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ К ГЛАВЕ II

1. Какой цели служат асимптотические формулы теории вероятностей?

2. Какая ошибка асимптотической формулы при неограниченном возрастании числа испытаний стремится к нулю — абсолютная или относительная?

3. Приведите пример применения локальной теоремы Муавра — Лапласа.

4. Чем отличаются условия теоремы Муавра — Лапласа от условий теоремы Пуассона? Приведите примеры, когда вероятность определенного числа наступлений события может быть найдена по формуле Пуассона.

5. Почему закон Пуассона называют законом редких событий?

6. Какую роль играет интегральная предельная теорема Муавра — Лапласа? В каких задачах она применяется? В чем заключается различие между задачами, в которых требуется применение локальной и интегральной предельных теорем?

7. В чем состоит теорема Бернулли?

8. Можно ли на основании теоремы Бернулли утверждать, что при неограниченном возрастании числа испытаний частота наступления события стремится к вероятности его наступления?

9. Приведите примеры применения теоремы Бернулли.

10. Согласуется ли с теоремой Бернулли оценка для наиболее вероятного числа наступлений события, полученная в § 4 предыдущей главы?

ГЛАВА III

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 13. СЛУЧАЙНАЯ ВЕЛИЧИНА И ЕЕ ЗАКОН РАСПРЕДЕЛЕНИЯ

Понятие случайной величины является одним из центральных понятий теории вероятностей. *Под случайной величиной понимается величина, принимающая то или иное числовое значение в зависимости от случая.* Более точное определение будет дано позднее.

Приведем некоторые примеры случайных величин.

1. *Число очков, выпадающих на игральной кости.* Эта случайная величина может принять одно из значений 1, 2, ..., 6.

2. *Число наступлений события A при n независимых испытаниях.* В результате n испытаний эта случайная величина может принять одно из значений 0,1,2, ..., n.

3. *Число выстрелов, производимых до первого попадания в цель.* Эта величина может принимать любое целое положительное значение.

4. *Расстояние от центра мишени до точки попадания.* Эта случайная величина, вообще говоря, может принимать любые положительные значения или нуль.

Случайная величина, принимающая конечное число или бесконечную последовательность различных значений, называется дискретной случайной величиной.

Из описанных выше случайных величин дискретными являются случайные величины примеров 1, 2, 3.

*Случайная величина, принимающая все значения из некоторого интервала, называется непрерывной случайной величиной** . К таким величинам относится случайная величина примера 4. (Существуют, разумеется, случайные величины, не являющиеся ни дискретными, ни непрерывными; рассмотрение их, однако, не входит в наши задачи.)

---

* Это определение неполно. Более точное определение непрерывной случайной величины будет приведено в § 15.

Мы ограничимся пока рассмотрением *дискретных* случайных величин. Чтобы охарактеризовать случайную величину, прежде всего необходимо указать возможные ее значения. Однако этого недостаточно: нужно еще знать, насколько часто принимаются различные значения этой величины. Эта частота характеризуется вероятностью отдельных ее значений. Иначе говоря, для случайной величины X следует указывать не только ее значения x_1, x_2, ..., но и вероятности событий

$$p_i = P(X = x_i) \qquad (i = 1, 2, \ldots),$$

состоящих в том, что случайная величина X приняла значение x_i.

Если перечислены все возможные значения X, то события $X = x_i$ не только несовместны, но и единственно возможны, так что сумма заданных вероятностей p_i должна равняться единице.

Соотношение, устанавливающее связь между значениями случайной величины и вероятностями этих значений, называют *законом распределения* случайной величины. Для дискретной случайной величины закон распределения удобно записывать в виде таблицы (табл. 1).

Таблица 1

x_1	x_2	...	x_n	...
p_1	p_2	...	p_n	...

Здесь $\sum\limits_i p_i = 1$. Число значений может быть конечным или бесконечным. В последнем случае ряд из вероятностей $\sum\limits_{i=1}^{\infty} p_i$ должен сходиться и его сумма должна быть равна единице.

Иногда удобно изображать закон распределения графически. Для этой цели откладывают по оси абсцисс возможные значения случайной величины X, а по оси ординат — соответствующие значения вероятностей. Ломаную, соединяющую полученные точки, называют *многоугольником распределения*.

Приведем примеры законов распределения случайных величин.

Пример 1. Пусть X — число очков, выпадающих на игральной кости. Возможные значения 1, 2, ..., 6 равновероятны. Поэтому закон распределения имеет вид таблицы 2.

Таблица 2

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Многоугольник распределения изображен на рис. 10.

Пример 2. Пусть X — число выстрелов по цели до первого попадания, причем вероятность попадания при отдельном выстреле равна p.

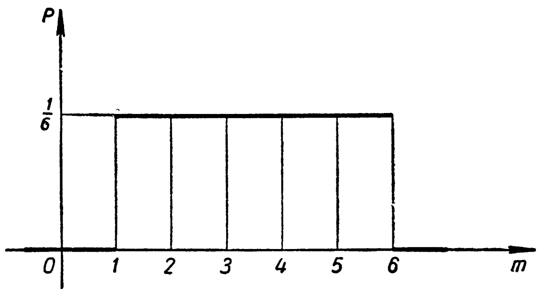


Рис. 10

Возможными значениями случайной величины будут здесь все натуральные числа 1, 2, ..., n, ... Вероятность того, что $X = 1$, равна, очевидно, p. Если $X = 2$, то это означает промах при первом выстреле и попадание при втором, так что вероятность этого равенства, как вероятность совмещения событий, равна произведению qp, где $q = 1 - p$. Аналогично найдем, что вероятность того, что $X = n$, равна $q^{n-1}\,p$, так что закон распределения задается здесь таблицей (табл. 3):

Таблица 3

x_i	1	2	3	4	...	n	...
p_i	p	pq	pq^2	pq^3	...	pq^{n-1}	...

Найдем сумму ряда $\sum\limits_{i=1}^{\infty} p_i$. Имеем:

$$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} pq^{n-1} = p(1 + q + q^2 + \ldots + q^{n-1} + \ldots).$$

Ряд в скобках представляет геометрическую прогрессию со знаменателем $0 < q < 1$; его сумма равна

$$\frac{1}{1-q} = \frac{1}{p}; \quad \text{следовательно,}$$

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

Многоугольник распределения для $p = q = \frac{1}{2}$ изображен на рис. 11.

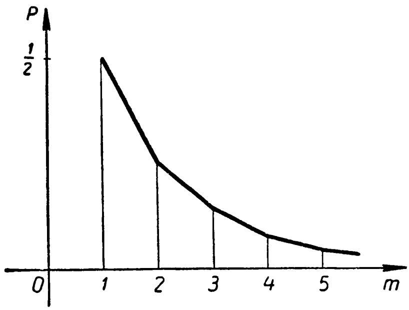


Рис. 11

§ 14. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТЬ ВЕРОЯТНОСТИ

Закон распределения случайной величины не всегда может быть задан таблицей. Например, если речь идет о непрерывной случайной величине (в том смысле, как она была определена в предыдущем параграфе), то все ее значения перечислить невозможно. Поэтому ее характеризуют не вероятностями отдельных значений, как характеризовали дискретную, а вероятностями того, что случайная величина принимает значения из определенного интервала, т. е. вероятностями неравенств вида $\alpha < X < \beta$. Впрочем, такого рода характеристики бывают полезными отнюдь не только для непрерывных случайных величин.

В дальнейшем мы будем говорить о вероятности неравенства $-\infty < X < x$, т. е. вероятности того, что случайная величина принимает значение, меньшее x. Эта вероятность $P(X < x)$ является, очевидно, функцией от x; обозначим ее через $F(x)$:

$$F(x) = P(X < x).$$

Функцию $F(x)$ называют *интегральным законом распределения* или *функцией распределения случайной величины.*

Если изображать значения случайной величины точками числовой оси и обозначить буквой N точку с абсциссой x, то $F(x)$ есть вероятность того, что точка M, являющаяся значением случайной величины, окажется левее N.

Установим некоторые свойства функции распределения.

Пусть X — случайная величина, x_1 и x_2 — две произвольные точки, причем $x_1 < x_2$. Сравним значения функции распределения $F(x)$ в этих точках. Так как событие $X < x_1$ влечет событие $X < x_2$, то ясно, что

$$P(X < x_1) \leqslant P(X < x_2),$$

или, по определению функции распределения,

$$F(x_1) \leqslant F(x_2).$$

Таким образом, функция распределения для любой случайной величины является *монотонно неубывающей.*

Далее, очевидно, что

$$F(-\infty) = 0,$$
$$F(+\infty) = 1,$$

откуда следует, что $F(+\infty) - F(-\infty) = 1$. Так как $F(x)$ монотонна и заключена между 0 и 1, то график функции $y = F(x)$ имеет две горизонтальные асимптоты: $y = 0$ при $x \to -\infty$ и $y = 1$ при $x \to +\infty$.

Перечисленные свойства функции распределения делают достаточно ясным ее поведение на всей действительной оси. К сказанному следует только добавить, что если все значения случайной величины принадлежат интервалу (a, b), то слева от точки a имеем $F(x) \equiv 0$, а справа от точки b $F(x) \equiv 1$.

Заметим, что функцию распределения можно построить для случайных величин любого типа — как непрерывных, так и дискретных.

Для дискретной случайной величины

$$F(x) = P(X < x) = P(-\infty < X < x) = \sum_{x_i < x} p_i,$$

где суммирование распространяется на значения x_i, удовлетворяющие неравенству $x_i < x$. В промежутке между двумя последовательными значениями X функция $F(x)$ постоянна. При переходе же аргумента x через возможное значение случайной величины x_i функция $F(x)$ скачком возрастает на величину $p_i = P(X = x_i)$, так что x_i будет точкой разрыва первого рода функции $F(x)$. Таким образом, функция распределения для дискретной случайной величины будет *ступенчатой функцией*. Вид ее указан на рис. 12.

Рассмотрим некоторые примеры построения функций распределения.

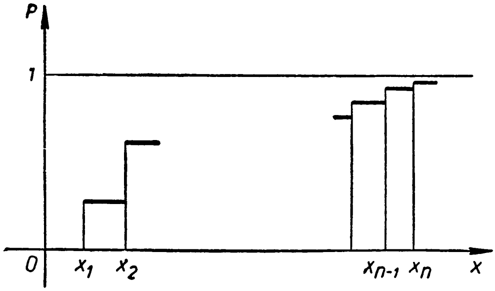


Рис. 12

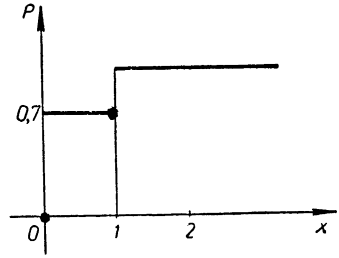


Рис. 13

Пример 1. Случайная величина X — число попаданий при одном выстреле, причем вероятность попадания равна 0,3.

Очевидно, для X возможны два значения — 0 и 1. Закон распределения может быть записан таблицей вида табл. 1. Отсюда следует, что функция распределения может быть задана равенствами:

$$F(x) = \begin{cases} 0 & \text{для} -\infty < x \leqslant 0, \\ 0{,}7 & \text{для} \quad 0 < x \leqslant 1, \\ 1 & \text{для} \quad 1 < x \leqslant \infty. \end{cases}$$

График функции распределения приведен на рис. 13.

Пример 2. Точка бросается наугад (без прицеливания) на отрезок [0,1]. Случайная величина X — абсцисса точки попадания (считается, что бросаемая точка обязательно попадает на отрезок [0, 1]).

В этом случае мы имеем дело с непрерывной случайной величиной, все значения которой принадлежат отрезку [0, 1]. Поэтому для $x \leqslant 0$ имеем $F(x) = 0$, ибо неравенство $X < x$ в этом случае невозможно. Напротив, для $x \geqslant 1$ неравенство $X < x$ достоверно, так что здесь $F(x) = 1$.

Пусть теперь число x удовлетворяет неравенству $0 < x < 1$. Так как точка бросается наугад, то следует считать все значения абсциссы одинаково вероятными. Согласно определению геометрической вероятности, данному в § 8, вероятность попадания на некоторый интервал пропорциональна длине этого интервала или, в данном случае, просто равна ей, потому что длина всего интервала возможных значений равна единице. Таким образом, вероятность $P(X<x)$ равна длине интервала $(0, x)$, т. е.

$$F(x) = P(X < x) = x.$$

Окончательное выражение для функции распределения нашей случайной величины может быть записано равенствами:

$$F(x) = \begin{cases} 0 & \text{при} -\infty < x < 0, \\ x & \text{при} \quad 0 \leqslant x \leqslant 1, \\ 1 & \text{при} \quad 1 < x < +\infty. \end{cases}$$

График функции распределения изображен на рис. 14.

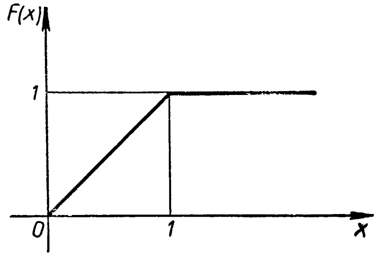


Рис. 14

Разберем теперь следующую задачу.

З а д а ч а 1. Зная функцию распределения величины X, вычислим вероятность того, что случайная величина X удовлетворяет неравенствам $\alpha < X < \beta$.

Пользуясь теоремой сложения вероятностей, напишем:

$$P(X < \beta) = P(X < \alpha) + P(\alpha \leqslant X < \beta),$$

откуда

$$P(\alpha \leqslant X < \beta) = P(X < \beta) - P(X < \alpha).$$

Исходя из определения функции распределения, можем написать:

$$P(\alpha \leqslant X < \beta) = F(\beta) - F(\alpha). \qquad (1)$$

Равенство (1) можно использовать для нахождения вероятности $P(X = \alpha)$. Для этого достаточно рассмотреть предел при $\beta \to \alpha$, так как при $\beta \to \alpha$ отрезок $[\alpha, \beta]$ стремится к точке α и X совпадает с α. Получаем:

$$P(X = \alpha) = \lim_{\beta \to \alpha} P(\alpha \leqslant X < \beta) = \lim_{\beta \to \alpha} [F(\beta) - F(\alpha)].$$

Если функция $F(x)$ непрерывна, то последний предел равен нулю. Следовательно,

$$P(X = \alpha) = 0,$$

т. е. *если функция распределения случайной величины непрерывна, то вероятность того, что случайная величина примет заранее заданное значение, равна нулю.*

Так как

$$P(\alpha \leqslant X < \beta) = P(X = \alpha) + P(\alpha < X < \beta),$$

то вместо (1) можно написать также

$$P(\alpha < X < \beta) = F(\beta) - F\alpha). \qquad (2)$$

Особо обращаем внимание читателя на следующий факт. Известно, что вероятность невозможного события равна нулю. При классическом определении вероятности, когда полная группа событий состоит из к о н е ч н о г о их числа, верно и обратное: событие, вероятность которого равна нулю, невозможно. Для непрерывной случайной величины это о б р а т н о е у т в е р ж д е н и е н е в е р н о. Хотя, как мы только что видели, вероятность того, что $X = \alpha$, где α — заранее выбранное число, р а в н а н у л ю, это событие **не является невозможным**. С подобным явлением мы сталкивались уже в § 8 при рассмотрении геометрических вероятностей.

Рассмотрим непрерывную случайную величину X, функция распределения которой предполагается **непрерывной и дифференцируемой**. Производная функции распределения

$$\varphi(x) = F'(x)$$

также играет важную роль. Эту функцию $\varphi(x)$ называют *дифференциальным законом распределения* или *плотностью вероятности случайной величины* X.

Легко установить вероятностный смысл этой функции. Из определения производной следует:

$$\varphi(x) = F'(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}. \tag{3}$$

Как было замечено выше, числитель (3) выражает вероятность того, что случайная величина X принимает значения между x и $x + \Delta x$,

$$F(x + \Delta x) - F(x) = P(x < X < x + \Delta x).$$

Таким образом, вместо (3) можно написать:

$$\varphi(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x}, \tag{4}$$

т. е. *плотность вероятности случайной величины X в точке x равна пределу отношения вероятности попадания величины X в интервал $(x, x + \Delta x)$ к Δx, когда Δx стремится к нулю.*

Для дальнейшего выяснения роли плотности вероятности рассмотрим еще одну задачу, аналогичную разобранной выше.

Задача 2. Зная плотность распределения величины X, вычислим вероятность того, что случайная величина X удовлетворяет неравенствам $\alpha < X < \beta$.

Используя (2) и известную из интегрального исчисления формулу Ньютона—Лейбница, получим:

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \int_{\alpha}^{\beta} F'(x)\,dx = \int_{\alpha}^{\beta} \varphi(x)\,dx. \tag{5}$$

Формуле (5) легко дать наглядное геометрическое истолкование. Построим график плотности распределения $y = \varphi(x)$ (рис. 15). Из (4) вытекает, что, с точностью до бесконечно малых высшего порядка по сравнению с Δx,

$$P(x < X < x + \Delta x) \approx \varphi(x)\,\Delta x.$$

Полученное произведение геометрически изображается площадью наклонно заштрихованного прямоугольника I (рис. 15). Вероятность того, что случайная величина примет значения внутри участка (α, β), выражается площадью криволинейной трапеции, расположенной над этим участком и ограниченной сверху графиком $y = \varphi(x)$. В этом и состоит геометрический смысл формулы (5).

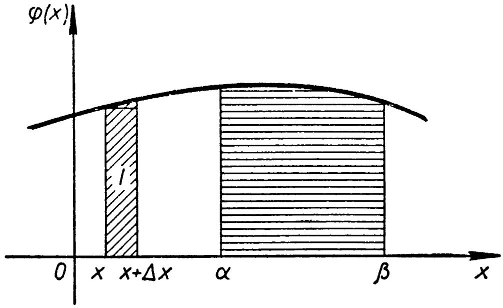


Рис. 15

Заметим, что кривая $y = \varphi(x)$ всегда лежит над осью Ox. Действительно, функция распределения $F(x)$, как было доказано выше, монотонно неубывающая, поэтому, как известно из дифференциального исчисления, ее производная должна быть неотрицательна, т. е. $\varphi(x) \geqslant 0$.

Из равенства (5) и свойств функции распределения вытекает также, что

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1. \tag{6}$$

Легко установить выражение интегрального закона распределения через дифференциальный. Интегральный закон распределения есть первообразная от дифференциального, причем такая, которая обращается в нуль при $x = -\infty$. Поэтому

$$F(x) = \int_{-\infty}^{x} \varphi(t)\,dt. \tag{7}$$

Пример 3. Найдем плотность распределения вероятности случайной величины X, рассмотренной в примере 2.

Из определения легко получаем, что

$$\varphi(x) = \begin{cases} 0 \text{ при } -\infty < x < 0, \\ 1 \text{ при } \quad 0 < x < 1 \\ 0 \text{ при } \quad 1 < x < +\infty. \end{cases}$$

В точках $x = 0$ и $x = 1$ функция $\varphi(x)$ не существует. В этих точках график функции $F(x)$ — интегрального закона распределения — имеет излом (см. рис. 14).

Пример 4. Показательный дифференциальный закон распределения задается формулой

$$\varphi(x) = \begin{cases} 0 & \text{при } x < 0 \\ Ae^{-\lambda x} & \text{при } x \geqslant 0, \end{cases}$$

где A и λ — постоянные величины. Считая λ заданным, найдем A и построим функцию распределения.

Согласно формуле (6), имеем:

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx = \int_{0}^{\infty} Ae^{-\lambda x}\,dx = -\frac{A}{\lambda}e^{-\lambda x}\Big|_0^{\infty} = \frac{A}{\lambda} = 1,$$

т. е. $A = \lambda$. Далее по формуле (7) при $x > 0$

$$F(x) = \int_{-\infty}^{x} \varphi(t)\,dt = \lambda\int_{-\infty}^{x} e^{-\lambda t}\,dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}.$$

При $x \leqslant 0$, очевидно, $F(x) = 0$. Рекомендуем читателю самостоятельно построить графики интегрального и дифференциального законов. Здесь, как и в примере 3, график функции $F(x)$ имеет угловую точку в начале координат.

П р и м е р 5. Пусть время работы электронной лампы есть случайная величина X. Найдем ее функцию распределения и плотность вероятности, если экспериментально установлено, что вероятность выхода работающей лампы из строя за небольшой промежуток времени наблюдения приблизительно пропорциональна времени ее дальнейшей работы и не зависит от того, сколько времени работала лампа до этого.

Пусть событие A состоит в том, что лампа проработает не менее x дней, а событие B — в том, что лампа выйдет из строя между x и $x + \Delta x$ днями. Мы имеем:

$$P(A) = P(X \geqslant x) = 1 - P(X < x) = 1 - F(x),$$
$$P(B) = P(x < X < x + \Delta x) = F(x + \Delta x) - F(x).$$

Так как мы начинаем вести наблюдение за р а б о т а ю щ и м и л а м п а м и, то экспериментально устанавливается не сама вероятность $P(B)$, а условная вероятность $P_A(B)$. Воспользуемся следствием из теоремы умножения

$$P(A)\,P_A(B) = P(B)\,P_B(A)$$

и заметим, что если событие B произошло, то наверняка произошло и событие A, т. е. $P_B(A) = 1$. Поэтому

$$P_A(B) = \frac{P(B)}{P(A)} = \frac{F(x+\Delta x) - F(x)}{1 - F(x)}. \qquad (8)$$

Если Δx достаточно мало, то можно записать, что

$$P_A(B) \approx \frac{\varphi(x)\,\Delta x}{1 - F(x)},$$

где $\varphi(x) = F'(x)$ — плотность распределения. С другой стороны, по условию вероятность $P_A(B)$ приближенно равна:

$$P_A(B) \approx k\,\Delta x, \tag{9}$$

где k — коэффициент пропорциональности. Сравнивая приближенные равенства (8) и (9) и учитывая, что ошибка их стремится к нулю при $\Delta x \to 0$, получаем уравнение:

$$\frac{\varphi(x)}{1 - F(x)} = k \text{ или } \varphi(x) = k - kF(x).$$

Дифференцируя последнее равенство, приходим к дифференциальному уравнению для искомой плотности вероятности:

$$\varphi'(x) = -k\varphi(x).$$

Интегрирование этого уравнения дает:

$$\varphi(x) = Ce^{-kx},$$

где C — произвольная постоянная. Для ее нахождения заметим, что в нашей задаче $X \geqslant 0$, поэтому для $x < 0$ имеем $\varphi(x) \equiv 0$. Согласно решению примера 3, получим, что $C = k$. Поэтому окончательно:

$$\varphi(x) = \begin{cases} 0 & \text{при } x < 0, \\ ke^{-kx} & \text{при } x \geqslant 0; \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 - e^{-kx} & \text{при } x \geqslant 0. \end{cases}$$

Коэффициент k должен быть установлен экспериментально. Примем его равным 0,01 (при измерении времени в днях) и вычислим: 1) вероятность того, что электронная лампа проработает не менее 30 дней; 2) вероятность того, что электронная лампа выйдет из строя между 30 и 40 днями работы.

Ответ на первый вопрос дается формулой

$$P_1 = 1 - F(30) = 1 - (1 - e^{-0,3}) = e^{-0,3} \approx 0,74.$$

Для второй вероятности получаем:

$$P_2 = \int_{30}^{40} \varphi(x)\,dx = F(40) - F(30) = e^{-0,3} - e^{-0,4} \approx$$

$$\approx 0,74 - 0,67 = 0,07.$$

§ 15. ОСНОВНЫЕ ПРИМЕРЫ ДИСКРЕТНЫХ И НЕПРЕРЫВНЫХ РАСПРЕДЕЛЕНИЙ

Пользуясь введенными в предыдущем параграфе понятиями функции распределения и плотности вероятности, можно теперь дать точное определение непрерывной случайной величины.

Случайная величина называется непрерывной, если ее функция распределения $F(x)$ *непрерывна на всей числовой оси, а плотность распределения* $\varphi(x) = F'(x)$ *существует и непрерывна всюду, кроме, быть может, дискретного множества точек.* При этом плотность $\varphi(x)$ может иметь как точки разрыва первого рода, так и точки бесконечного разрыва.

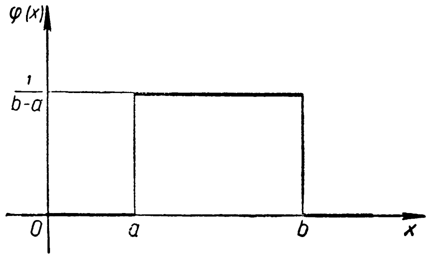


Рис. 16

В настоящем параграфе приводятся наиболее часто встречающиеся типы распределений непрерывных и дискретных случайных величин и примеры их применения.

I. Равномерное распределение вероятностей. Пусть плотность вероятности равна нулю всюду, кроме интервала (a, b), на котором она постоянна. Если обозначить эту постоянную через A, то, в силу формулы (6) из § 14, получим:

$$\int_a^b A dx = 1,$$

откуда $A = \frac{1}{b-a}$. Поэтому плотность равномерного распределения задается формулой

$$\varphi(x) = \begin{cases} 0 & \text{при } -\infty < x < a, \\ \frac{1}{b-a} & \text{при } \quad a < x < b, \\ 0 & \text{при } \; b < x < +\infty \end{cases} \tag{1}$$

(рис. 16). В точках $x = a$ и $x = b$ функция $\varphi(x)$ разрывна. Для нахождения интегрального закона распределения воспользуемся формулой (7) предыдущего параграфа.

Если $x \leqslant a$, то

$$F(x) = \int_{-\infty}^{x} \varphi(t)\, dt = 0.$$

Для $a < x < b$ получаем:

$$F(x) = \int_{-\infty}^{x} \varphi(t)\, dt = \int_{-\infty}^{a} \varphi(t)\, dt + \int_{a}^{x} \varphi(t)\, dt =$$

$$= 0 + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a},$$

ибо для $a < t < b$ имеем $\varphi(t) = \frac{1}{b-a}$. Наконец, при $x \geqslant b$ получим:

$$F(x) = \int_{-\infty}^{x} \varphi(t)\,dt = \int_{-\infty}^{a} \varphi(t)\,dt + \int_{a}^{b} \varphi(t)\,dt + \int_{b}^{x} \varphi(t)\,dt =$$

$$= 0 + \int_{a}^{b} \frac{1}{b-a}\,dt + 0 = 1.$$

Таким образом, интегральный закон равномерного распределения задается формулой

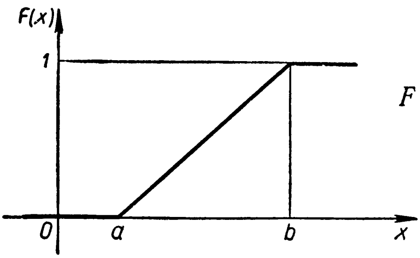


Рис. 17

$$F(x) = \begin{cases} 0 & \text{при } -\infty < x \leqslant a, \\ \dfrac{x-a}{b-a} & \text{при } \quad a < x < b, \\ 1 & \text{при } \quad b \leqslant x < +\infty \end{cases}$$

(рис. 17). Эта функция непрерывна всюду. Частный случай равномерного распределения встречался нам в § 14 (см. пример 2).

II. Закон Коши. Пусть плотность распределения случайной величины X задана формулой

$$\varphi(x) = \frac{A}{1+x^2}.$$

Как и выше, величина A определяется из соотношения

$$\int_{-\infty}^{+\infty} \varphi(x)dx = \int_{-\infty}^{+\infty} \frac{A}{1+x^2}\,dx = A \operatorname{arctg} x \Big|_{-\infty}^{+\infty} = \pi A = 1,$$

откуда $A = \frac{1}{\pi}$. Таким образом, закон Коши принимает вид:

$$\varphi(x) = \frac{1}{\pi(1+x^2)} \tag{2}$$

(рис. 18). Функция распределения получается с помощью формулы (7) из § 14:

$$F(x) = \int_{-\infty}^{x} \varphi(t)\,dt = \int_{-\infty}^{x} \frac{dt}{\pi(1+t^2)} = \frac{1}{\pi} \operatorname{arctg} t \Big|_{-\infty}^{x} = \frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}$$

(рис. 19).

В качестве примера найдем вероятность того, что случайная величина, распределенная по закону Коши, примет значение из интервала (—1, 1). Формула (2) предыдущего параграфа дает:

$$P(-1 < x < 1) = F(1) - F(-1) = \frac{1}{\pi}\left[\frac{\pi}{4} - \left(-\frac{\pi}{4}\right)\right] = \frac{1}{2}.$$

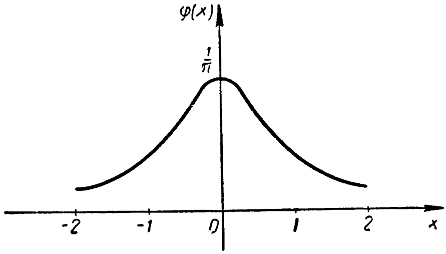


Рис. 18

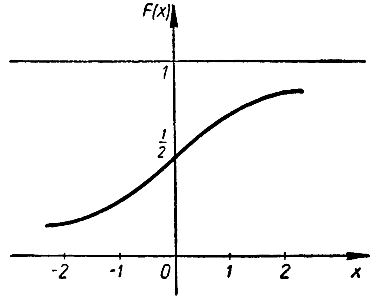


Рис. 19

III. Закон арксинуса. Пусть плотность распределения задана следующим образом:

$$\varphi(x) = \begin{cases} 0 & \text{при } -\infty < x < -1, \\ \dfrac{A}{\sqrt{1-x^2}} & \text{при } -1 < x < 1, \\ 0 & \text{при } 1 < x < +\infty. \end{cases}$$

Прежде всего находим величину A:

$$\int_{-1}^{1} \frac{A}{\sqrt{1-x^2}}\, dx = A \arcsin x \Big|_{-1}^{1} = \pi A = 1,$$

т. е. $A = \dfrac{1}{\pi}$. Составим выражение для функции распределения:

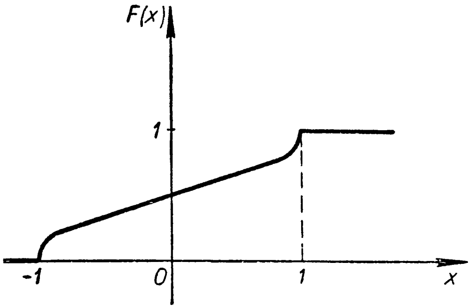
$$F(x) = \begin{cases} 0 & \text{при } -\infty < x < -1, \\ \dfrac{1}{\pi} \displaystyle\int_{-1}^{x} \frac{dt}{\sqrt{1-t^2}} = \frac{\arcsin x}{\pi} + \frac{1}{2} & \text{при } -1 < x < 1, \\ 1 & \text{при } 1 < x < +\infty. \end{cases}$$

Графики функций $\varphi(x)$ и $F(x)$ изображены на рис. 20; здесь функция $\varphi(x)$ имеет бесконечные разрывы в точках $x = -1$ и $x = 1$.

IV. Биномиальное распределение. Рассмотрим серию из n независимых испытаний, в каждом из которых вероятность наступления события A равна p. Случайная величина X означает число наступления событий. Она дискретна, и ее возможными значениями являются неотрицательные целые числа $0, 1, 2, \ldots, n$.

Закон распределения случайной величины X задается уже известной нам формулой (см. § 4):

$$P_{m,n} = C_n^m p^m q^{n-m}, \tag{3}$$



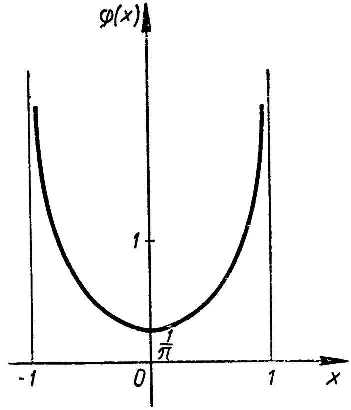


Рис. 20

определяющей вероятность равенства $X = m$. Как было указано еще в § 4, это выражение представляет член разложения бинома $(p + q)^n$. Поэтому говорят, что *случайная величина X подчиняется биномиальному закону распределения.*

Примеры приложений биномиального распределения уже встречались нам в предыдущих параграфах.

V. Нормальный закон распределения (*закон Гаусса*). Среди законов распределения, которым подчиняются встречающиеся на практике случайные величины, чаще всего приходится иметь дело с нормальным законом распределения. Для этого закона плотность вероятности $\varphi(x)$ задается формулой

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \tag{4}$$

Функция $\varphi(x)$ положительна, и остается только проверить выполнение условия, выраженного формулой (6) из § 14. Для этого произведем в интеграле $\frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$ замену переменного, положив $\frac{x-a}{\sigma} = t$, $dx = \sigma dt$. Получим:

$$\frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt.$$

Последний же интеграл, как указано в § 10, равен единице.

С частным случаем нормального закона распределения (при $a = 0$, $\sigma = 1$) мы уже встречались в § 9 при рассмотрении асимптотической формулы в задаче о повторении испытаний. Функция $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ была подробно изучена в § 10. Там же приведен и ее график (см. рис. 7).

Для построения графика функции (4) выясним прежде всего геометрический смысл параметров a и σ, отложив выяснение их вероятностного смысла на более поздний срок.

Из формулы (4) видно, что кривая $y = \varphi(x)$ достигает максимума при $x = a$, причем максимальное значение $y_{max} = \frac{1}{\sigma\sqrt{2\pi}}$.

С ростом σ величина максимального значения уменьшается, а так как площадь, ограниченная всей кривой и осью абсцисс, равна единице, то с ростом σ кривая как бы растягивается вдоль оси Ox. Наоборот, при уменьшении σ кривая вытягивается вверх возле максимумов, но сжимается в горизонтальном направлении.

На рис. 21 изображены графики нормальных законов $y = \varphi(x)$ при различных a, но одном и том же σ. На рис. 22, наоборот, даны графики функций $y = \varphi(x)$ при $a = 0$, но различных σ.

Причина того, что нормальный закон распределения встречается так часто, кроется в следующем. Оказывается, что если значения, которые принимает случайная величина, зависят от большого числа различных факторов, каждый из которых, взятый в отдельности влияет на эту величину сравнительно мало, то можно приближенно считать, что рассматриваемая случайная величина подчиняется нормальному закону распределения. Точная формулировка всех условий, при которых это действительно будет так, рассматривается в § 23.

В качестве примера можно указать на задачу о рассеивании снарядов. Производится стрельба из орудия при постоянном прицеле и сохранении неизменными всех прочих условий. Предполагается,

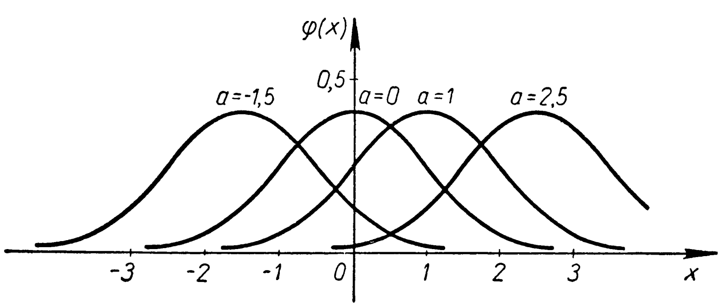


Рис. 21

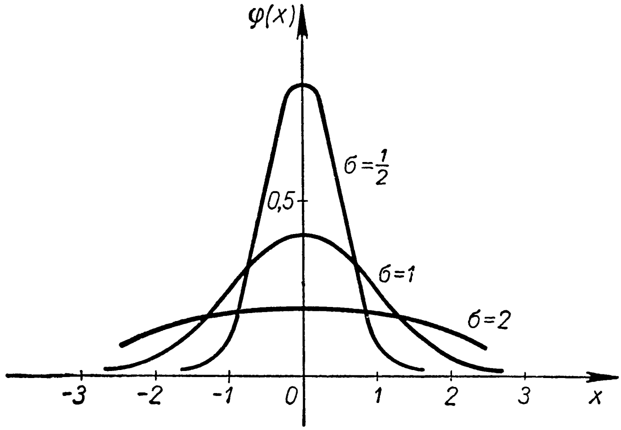


Рис. 22

что все снаряды ложатся на прямой и что рассеяние снарядов симметрично, т. е. одинаковые отклонения вперед и назад от точки прицеливания равновероятны. Случайной величиной X является координата точки падения снаряда. Ясно, что отклонение точки падения снаряда от точки прицеливания зависит от очень большого числа факторов: изменившихся атмосферных условий, изменения формы снаряда, различного веса заряда (пороха) и т. д., т. е., согласно сказанному выше, мы вправе ожидать, что случайная величина X должна приближенно подчиняться нормальному закону. И действительно, опытные данные показывают, что закон распределения обычно хорошо согласуется с нормальным законом. Тот факт, что нормальный закон распределения хорошо приближает биномиальный (см. § 9 и 12), также теперь может быть объяснен. Дело в том, что общее число положительных исходов испытаний слагается из большого числа отдельных положительных исходов, каждый из которых сам по себе мало влияет на их общее число.

Рассмотрение нормального закона распределения мы закончим разбором следующей задачи, аналогичной разобранным в § 14 задачам 1 и 2.

Задача 3. Случайная величина X распределена по нормальному закону с плотностью вероятности

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Вычислить вероятность того, что случайная величина X удовлетворяет неравенствам $\alpha < X < \beta$.

Пользуясь формулой (5) из § 14, получаем:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \varphi(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx.$$

Последний интеграл легко сводится к интегралу Лапласа. В самом деле, положим

$$\frac{x-a}{\sigma} = t;\ dx = \sigma\,dt.$$

Тогда

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} \sigma\,dt = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \quad (5)$$

Значения $\Phi\left(\frac{\beta-a}{\sigma}\right)$ и $\Phi\left(\frac{\alpha-a}{\sigma}\right)$ могут быть найдены из таблицы, помещенной в приложении. Для $a = 0$, $\sigma = 1$ мы получаем простой результат:

$$P(\alpha < X < \beta) = \Phi(\beta) - \Phi(\alpha).$$

Для интервала, симметричного относительно точки a, формула (5) упрощается. Пусть $\beta = a + \gamma$ и $\alpha = a - \gamma$; тогда

$$P(a-\gamma < X < a+\gamma) = \Phi\left(\frac{\gamma}{\sigma}\right) - \Phi\left(-\frac{\gamma}{\sigma}\right) = 2\Phi\left(\frac{\gamma}{\sigma}\right). \quad (6)$$

Проделаем некоторые числовые расчеты. Положим сперва $\alpha = a$ и $\beta = a + \sigma$. Тогда формула (5) дает:

$$P(a < X < a + \sigma) = \Phi(1) - \Phi(0) = 0{,}3413.$$

Таким же точно образом найдем:

$$P(a + \sigma < X < a + 2\sigma) = \Phi(2) - \Phi(1) = 0{,}4772 - 0{,}3413 = 0{,}1359;$$

$$P(a + 2\sigma < X < a + 3\sigma) = \Phi(3) - \Phi(2) = 0{,}49865 - 0{,}4772 = 0{,}02145.$$

Наконец, если положить $\alpha = a - 3\sigma$, $\beta = a + 3\sigma$, то

$$P(a - 3\sigma < X < a + 3\sigma) = 2\Phi(3) = 0{,}9973.$$

Последний результат означает, что *с вероятностью, близкой к единице* ($= 0{,}9973$), *случайная величина, подчиняющаяся нормальному закону распределения, не выходит за пределы интервала* $(a - 3\sigma, a + 3\sigma)$. Это утверждение носит название *правила трех сигм.*

VI. Закон распределения Пуассона. Как и нормальный закон распределения, закон Пуассона весьма часто встречается в различных приложениях теории вероятностей. Как мы уже видели в § 11, закон Пуассона может быть получен как асимптотический для биномиального закона распределения при малых вероятностях p. По этой причине закон Пуассона часто называют законом редких событий.

Пусть X — дискретная случайная величина, которая может принимать целые неотрицательные значения. Если вероятность равенства $X = m$ определяется формулой

$$P_m = \frac{\lambda^m}{m!} e^{-m}, \tag{7}$$

то мы говорим, что величина X *распределена по закону Пуассона*.

Легко проверить, что

$$\sum_{m=0}^{\infty} P_m = e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} e^{\lambda} = 1.$$

К случайным величинам, подчиненным закону Пуассона, приводит большое число задач, относящихся к вопросам массового обслуживания.

В качестве примера укажем работу телефонной станции. Можно доказать, что при выполнении некоторых условий вероятность m вызовов за промежуток времени t определяется формулой

$$P_m(t) = \frac{(at)^m}{m!} e^{-at}. \tag{8}$$

Если положить $at = \lambda$, то формула (8) означает, что случайная величина X *распределена по закону Пуассона*.

§ 16. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ

Выше мы познакомились с тем, каким образом изучается случайная величина, если известен ее закон распределения. В ряде случаев для решения задач, поставленных практикой, можно ограничиться рассмотрением лишь некоторых числовых характеристик случайной величины. Определение этих характеристик дается с помощью закона распределения. Бывает даже, что закон распределения случайной величины неизвестен, и поэтому ее полное изучение невозможно, но знание этих числовых характеристик все же позволяет решить некоторые вопросы, относящиеся к случайной величине.

Основными числовыми характеристиками случайной величины, с которыми мы сейчас познакомимся, являются *математическое ожидание* (или *среднее значение*) и *дисперсия*.

Пусть X — дискретная случайная величина, возможные значения которой $x_1, x_2, \ldots, x_n$ принимаются соответственно с вероятностями $p_1, p_2, \ldots, p_n$, так что $\sum_{i=1}^{n} p_i = 1$. Математическим ожиданием случайной величины X называется число $M(X)$ (иногда обозначается также $E(X)$), определяемое равенством

$$M(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = \sum_{i=1}^{n} x_i p_i. \quad (1)$$

Число возможных значений дискретной случайной величины может оказаться и бесконечным. В этом случае сумма вероятностей представляет ряд (сходящийся к единице). Для определения математического ожидания нужно тогда воспользоваться рядом

$$M(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n + \ldots = \sum_{i=1}^{\infty} x_i p_i, \quad (2)$$

причем для существования математического ожидания следует предполагать, что ряд (2) сходится абсолютно*.

Таким образом, *математическим ожиданием* (или *средним значением*) *дискретной случайной величины называется сумма произведений всех ее возможных значений на соответствующие вероятности.*

Выражение математического ожидания имеет следующий механический смысл. Учитывая, что $\sum_{i} p_i = 1$, запишем (1) в виде

$$M(X) = \frac{\Sigma x_i p_i}{\Sigma p_i}. \quad (3)$$

Из (3) следует, что $M(X)$ есть абсцисса центра тяжести системы точек, абсциссы которых равны возможным значениям X, а массы, помещенные в эти точки, равны соответствующим вероятностям.

Рассмотрим некоторые примеры.

Пример 1. Найдем математическое ожидание случайной величины X — числа очков, выпадающих на игральной кости (см. пример 1 из § 13).

Здесь X принимает значения от 1 до 6 с вероятностями, равными $\frac{1}{6}$; следовательно,

$$M(X) = \sum_{n=1}^{6} n \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{1+6}{2} \cdot 6 = 3{,}5.$$

* Если ряд (2) не сходился бы абсолютно, то математическое ожидание случайной величины зависело бы от порядка расположения ее значений. Эту зависимость мы хотим исключить, так как математическое ожидание должно быть вполне определенным числом.

Итак, математическое ожидание равно 3,5; как видим в данном примере, это число не совпадает ни с одним из значений случайной величины.

Пример 2. Пусть X — число выстрелов по цели до первого попадания, причем вероятность попадания при отдельном выстреле равна p (см. пример 2, § 13). Найдем $M(X)$.

При решении примера 2 из § 13 указывалось, что случайная величина X может принимать значения 1, 2, ..., n, ... с вероятностями p, pq, ..., pq^{n-1},..., где $q = 1 - p$. Имеем:

$$M(X) = \sum_{n=1}^{\infty} npq^{n-1} = p\sum_{n=1}^{\infty} nq^{n-1}.$$

Для вычисления последней суммы воспользуемся формулой для суммы членов геометрической прогрессии ($|q| < 1$)

$$1 + q + q^2 + \ldots + q^n + \ldots = \frac{1}{1-q}$$

и продифференцируем обе части этого равенства по q:

$$\sum_{n=1}^{\infty} nq^{n-1} = 1 + 2q + 3q^2 + \ldots + nq^{n-1} + \ldots = \frac{1}{(1-q)^2}.$$

Окончательно:

$$M(X) = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

Итак, среднее число требующихся для поражения цели выстрелов, равно $\frac{1}{p}$. Это число показывает, какое число выстрелов потребуется в среднем при большом числе серий стрельб. Оно может служить исходным при расчете числа необходимых снарядов.

Пример 3. Беспроигрышная лотерея содержит N билетов, из которых m_1 билетов с выигрышем в x_1 рублей, m_2 билетов с выигрышем в x_2 рублей, ..., m_n билетов с выигрышем в x_n рублей. Какова должна быть цена билета a, чтобы сумма денег, вырученных от продажи билетов, равнялась сумме всех выигрышей?

Это условие означает, что

$$Na = m_1x_1 + m_2x_2 + \ldots + m_nx_n,$$

где

$$N = m_1 + m_2 + \ldots + m_n.$$

Отсюда

$$a = \frac{m_1}{N}x_1 + \frac{m_2}{N}x_2 + \ldots + \frac{m_n}{N}x_n.$$

Рассмотрим случайную величину X, означающую выигрыш, выпавший на один билет. Возможными значениями этой случайной величины являются $x_1, x_2, \ldots, x_n$. Нетрудно подсчитать, что вероятность каждого такого значения $p_i = \frac{m_i}{N}$. Сравнив выражение для математического ожидания случайной величины

$$M(X) = \sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} \frac{m_i}{N} x_i$$

с выражением для цены билета a, замечаем, что искомая величина цены билета равна математическому ожиданию суммы выигрыша $a = M(X)$.

Для непрерывной случайной величины определения (1) или (2) для математического ожидания непригодны, потому что, как было сказано выше, вероятность каждого отдельного значения для непрерывной случайной величины равна нулю.

Математическим ожиданием непрерывной случайной величины называется интеграл

$$M(X) = \int_{-\infty}^{+\infty} x\varphi(x)\,dx, \tag{4}$$

где $\varphi(x)$ — *плотность распределения случайной величины* X.

Формула (4) является интегральным аналогом формулы (1). Действительно, все значения в интервале $(x, x + dx)$ можно считать примерно равными x, а вероятность таких значений, как было выяснено в § 14, равна $\varphi(x)\,dx$. Поэтому значения x_i заменяются на x, вероятности p_i — на $\varphi(x)\,dx$, а сумма заменяется интегралом.

Как и для дискретной случайной величины, понятию математического ожидания непрерывной случайной величины можно придать механическое истолкование. Для этого достаточно, воспользовавшись установленным в § 14 равенством $\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1$, записать формулу (4) в виде

$$M(X) = \frac{\int_{-\infty}^{+\infty} x\varphi(x)\,dx}{\int_{-\infty}^{+\infty} \varphi(x)\,dx}. \tag{5}$$

Из формулы (5) видно, что математическое ожидание равно абсциссе центра тяжести однородной плоской фигуры, ограниченной графиком функции $y = \varphi(x)$ (плотности вероятности) и осью абсцисс.

Определим теперь математические ожидания дискретных и непрерывных случайных величин, законы распределения которых рассматривались нами в § 15.

I. Биномиальное распределение. Определим математическое ожидание числа наступлений события A в серии из n независимых испытаний, если вероятность наступления A при каждом отдельном испытании равна p.

Как известно из §§ 4 и 15, случайная величина X, означающая число наступлений события A, подчиняется биномиальному закону распределения: она может принимать значения $m = 0,1,2, \ldots$ с вероятностями $P_{m,n} = C_n^m p^m q^{n-m}$. Иначе говоря, закон распределения случайной величины X задается таблицей (табл. 4):

Таблица 4

x_i	0	1	2	...	m	...	n
p_i	q^n	npq^{n-1}	$\frac{n(n-1)}{1 \cdot 2} p^2 q^{n-2}$	...	$C_n^m p^m q^{n-m}$	...	p^n

По определению математического ожидания находим:

$$M(X) = 0 \cdot q^n + 1 \cdot npq^{n-1} + 2\,\frac{n(n-1)}{1 \cdot 2} p^2 q^{n-2} + \ldots +$$
$$+ m\,\frac{n(n-1)\ldots(n-m+1)}{1 \cdot 2 \cdot \ldots \cdot m} p^m q^{n-m} + \ldots + np^n.$$

Выносяза скобку np и производя сокращения, получаем:

$$M(X) = np\,[q^{n-1} + (n-1)\,pq^{n-2} + \ldots +$$
$$+ \frac{(n-1)\ldots(n-m+1)}{1 \cdot 2 \cdot \ldots \cdot (m-1)} p^{m-1} q^{n-m} + \ldots + p^{n-1}\Big].$$

Выражение, стоящее в квадратных скобках, представляет разложение бинома $(p+q)^{n-1}$ и равно единице, ибо $p + q = 1$. Поэтому мы получаем:

$$M(X) = np. \tag{6}$$

Заметим, что в § 4 для наиболее вероятного числа наступлений события A мы получили оценку

$$np - q < \mu < np + p. \tag{7}$$

Сравнивая (6) и (7), мы видим, что μ есть такое целое число, которое отличается от математического ожидания числа наступлений события A меньше чем на единицу.

II. Закон Пуассона. Рассмотрим дискретную случайную величину X, распределенную по закону Пуассона, т. е. принимающую значения $0,1,2, \ldots, m, \ldots$ с вероятностями

$$P_m = \frac{\lambda^m}{m!} e^{-\lambda}.$$

Запишем закон распределения в виде таблицы (табл. 5):

Таблица 5

x_i	0	1	2	...	m	...
p_i	$e^{-\lambda}$	$\frac{x}{1}e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$	...	$\frac{\lambda^m}{m!}e^{-\lambda}$	...

Для математического ожидания имеем:

$$M(X) = 0 \cdot e^{-\lambda} + 1 \cdot \frac{\lambda}{1} e^{-\lambda} + 2 \cdot \frac{\lambda^2}{2!} e^{-\lambda} + \ldots + \\ + m \frac{\lambda^m}{m!} e^{-\lambda} + \ldots = \\ = \lambda e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^{m-1}}{(m-1)!} + \ldots\right).$$

Но, как известно, ряд в скобках представляет разложение функции e^{λ} в ряд Маклорена. Поэтому математическое ожидание равно $\lambda e^{-\lambda} \cdot e^{\lambda}$ или

$$M(X) = \lambda. \tag{8}$$

Мы выяснили, таким образом, вероятностный смысл параметра λ, входящего в закон распределения Пуассона: *параметр λ равен математическому ожиданию случайной величины.*

III. Равномерное распределение. Пусть непрерывная случайная величина X подчиняется закону равномерного распределения вероятностей. Как было найдено в § 15, плотность распределения в этом случае имеет вид:

$$\varphi(x) = \begin{cases} 0 & \text{при } -\infty < x < a, \\ \frac{1}{b-a} & \text{при } a < x < b, \\ 0 & \text{при } b < x < +\infty. \end{cases}$$

Пользуясь формулой (4) для математического ожидания, получим:

$$M(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{a+b}{2}, \tag{9}$$

так что *математическое ожидание случайной величины, равномерно распределенной на интервале (a, b), находится в центре этого интервала.*

IV. Нормальное распределение. Непрерывная случайная величина X подчиняется закону распределения Гаусса, т. е. ее плотность распределения имеет вид:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Формула (4) для математического ожидания дает:

$$M(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} xe^{-\frac{(x-a)^2}{2\sigma^2}}\,dx.$$

Для вычисления интеграла применим подстановку $\frac{x-a}{\sigma}=t$. Тогда

$$M(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma t+a)\,e^{-\frac{t^2}{2}}\,\sigma\,dt=$$

$$=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} te^{-\frac{t^2}{2}}\,dt+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt. \tag{10}$$

Первый из полученных двух интегралов равен нулю, так как подынтегральная функция нечетна. Второе слагаемое, как было показано в § 10, равно

$$a\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt=a.$$

Итак,

$$M(X)=a. \tag{11}$$

Мы выяснили, таким образом, вероятностный смысл параметра a, входящего в выражение нормального закона распределения: *параметр a равен математическому ожиданию случайной величины*.

Нам остается только рассмотреть некоторые свойства математического ожидания.

Теорема 1. *Математическое ожидание постоянной величины есть сама эта величина*:

$$M(C)=C.$$

Доказательство. Постоянную величину C можно рассматривать как дискретную случайную величину, принимающую лишь одно значение с вероятностью единица. Поэтому

$$M(C)=C\cdot 1=C.$$

Теорема 2. *Постоянный множитель можно выносить за знак математического ожидания.*

Доказательство придется проводить отдельно для дискретных и непрерывных случайных величин. Для дискретной случайной величины, пользуясь определением (1), получаем:

$$M(CX)=\sum_{i=1}^{n} C\,x_i\,p_i=C\sum_{i=1}^{n} x_i\,p_i=CM(x).$$

Для непрерывных случайных величин нужно воспользоваться формулой (4), которая дает:

$$M(CX) = \int_{-\infty}^{+\infty} Cx\,\varphi(x)\,dx = C\int_{-\infty}^{+\infty} x\,\varphi(x)\,dx = CM(X).$$

Итак, мы познакомились с одной из основных числовых характеристик случайной величины — математическим ожиданием, которое характеризует среднее значение случайной величины.

Однако, зная только среднее значение случайной величины, нельзя представить себе расположение значений случайной величины. Например, для случайной величины, принимающей значение $+1$ и -1 с вероятностью 0,5 каждое, как и для другой случайной величины, принимающей значения $+100$ и -100 с теми же вероятностями, математическое ожидание одинаково и равно нулю. Между тем разброс этих величин относительно их общего математического ожидания совершенно различен.

Чтобы охарактеризовать отклонение случайной величины от ее среднего значения, т. е. охарактеризовать разброс значений этой величины, вводят другую ее числовую характеристику — *дисперсию*, или *рассеяние*.

Для характеристики разброса не удается использовать разность между случайной величиной и ее средним значением, хотя на первый взгляд это и кажется наиболее естественным. Дело в том, что сама эта разность есть также случайная величина. Если же взять ее математическое ожидание, то, в силу свойств математического ожидания, для любой случайной величины X имеем:

$$M[X - M(X)] = M(X) - M[M(X)] = 0,$$

так что такая характеристика оказывается бесполезной.

Среднее значение отклонения получилось равным нулю, потому что положительные и отрицательные отклонения, т. е. отклонения в ту и другую сторону от среднего, взаимно уравновешиваются. Чтобы этого избежать, рассматривают не сами отклонения от среднего, а их квадраты, которые все неотрицательны, и в качестве характеристики рассеяния принимают среднее значение квадрата отклонения.

Дисперсией, или рассеянием, случайной величины X называется математическое ожидание квадрата разности случайной величины и ее математического ожидания

$$D(X) = M[X - M(X)]^2. \qquad (12)$$

Обозначив для краткости $M(X) = \bar{x}$, можем вместо (12) написать:

$$D(X) = M(X - \bar{x})^2,$$

т. е. дисперсия случайной величины X равна математическому ожиданию случайной величины $(X-\bar{x})^2$.

Пусть случайная величина X дискретна и принимает значения $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$. Так как события $X = x_i$ и $X-\bar{x}=x_i-\bar{x}$ равносильны, то их вероятности равны. Поэтому случайная величина $(X-\bar{x})^2$ принимает значения $(x_i-\bar{x})^2$ с вероятностью p_i $(i = 1, 2, \ldots)$* . Поэтому для дискретной случайной величины формула для вычисления дисперсии имеет вид:

$$D(X) = \sum_i (x_i - \bar{x})^2 p\ . \qquad (13)$$

Аналогично, для непрерывной случайной величины получаем:

$$D(X) = \int_{-\infty}^{+\infty} (x-\bar{x})^2 \varphi(x)\, dx. \qquad (14)$$

Часто вместо обозначения $D(X)$ применяется также обозначение $\sigma^2(X)$. Величину $\sigma = \sqrt{D(X)}$ называют *средним квадратичным отклонением* или *стандартом*.

Пример 3. Число очков, выбиваемых при одном выстреле каждым из двух стрелков, подчиняется законам распределения, приведенным в табл. 6 и 7.

Таблица 6

x_i	1	2	3
p_i	0,3	0,2	0,5

Таблица 7

x_i	1	2	3
p_i	0,1	0,6	0,3

Найдем математическое ожидание числа очков при отдельном выстреле для каждого стрелка. Для первого стрелка имеем:

$$M(X_1) = 1 \cdot 0{,}3 + 2 \cdot 0{,}2 + 3 \cdot 0{,}5 = 2{,}2.$$

Для второго стрелка:

$$M(X_2) = 1 \cdot 0{,}1 + 2 \cdot 0{,}6 + 3 \cdot 0{,}3 = 2{,}2.$$

Таким образом, математическое ожидание числа очков для обоих стрелков одинаково. Определим теперь дисперсию случайных величин X_1 и X_2. Для первого стрелка:

$$D(X_1) = (1-2{,}2)^2 \cdot 0{,}3+(2-2{,}2)^2 \cdot 0{,}2 +(3-2{,}2)^2 \cdot 0{,}5 = 0{,}76.$$

* Это верно, если все значения $(x_i-\bar{x})^2$ различны. Если же $(x_i-\bar{x})^2$+ и $(x_j-\bar{x})^2$ совпадают, то это значение принимается с вероятностью p_i+p_j, так что формула (13) остается в силе.

Для второго стрелка:

$$D(X_2)=(1-2{,}2)^2\cdot 0{,}1+(2-2{,}2)^2\cdot 0{,}6+(3-2{,}2)^2\cdot 0{,}3=0{,}36.$$

Следовательно, при одинаковом среднем для числа очков, выбиваемых обоими стрелками, рассеяние результатов у первого превышает рассеяние у второго. Таким образом, у второго стрелка бо́льшая кучность, т. е. результаты его стрельбы более устойчивы.

Вообще можно заметить, что, чем меньше дисперсия, тем лучше значения случайной величины характеризуются ее математическим ожиданием.

Пользуясь тем, что математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых*, можно получить другое выражение для дисперсии, более удобное, чем формула (12). Для этого преобразуем формулу (12) следующим образом:

$$D(X)=M[X-M(X)]^2=M[X^2-2XM(X)+(M(X))^2].$$

В силу указанного свойства, последнее выражение можно представить в виде суммы математических ожиданий. Заметим еще, что $M(X)$ есть постоянная величина и ее математическое ожидание, по теореме 1, равно ей самой. Поэтому мы получаем:

$$D(X)=M(X^2)-2M(X)\cdot M(X)+(M(X))^2,$$

или окончательно:

$$D(X)=M(X^2)-(M(X))^2. \tag{15}$$

Таким образом, *дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания.*

В качестве примера применения формулы (15) подсчитаем дисперсию для результатов выстрелов первого стрелка из примера 3. Как было подсчитано, $M(X_1)=2{,}2$. Далее, математическое ожидание квадрата равно:

$$M(X_1^2)=1^2\cdot 0{,}3+2^2\cdot 0{,}2+3^2\cdot 0{,}5=5{,}6.$$

Отсюда

$$D(X_1)=5{,}6-2{,}2^2=0{,}76,$$

как и раньше.

Рассмотрим теперь некоторые свойства дисперсии.

Теорема 3. *Дисперсия постоянной величины равна нулю.*

$$D(C)=0$$

---

* Точную формулировку и доказательство этого свойства математического ожидания см. ниже, § 21.

Действительно,

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

Этого следовало ожидать, ибо математическое ожидание постоянной равно ей самой и никакого рассеяния значений в этом случае не может быть.

Теорема 4. *Постоянный множитель можно выносить из-под знака дисперсии, возводя его в квадрат*:

$$D(CX) = C^2 D(X). \tag{16}$$

Доказательство. Из формулы (15) следует:

$$D(CX) = M(C^2X^2) - (M(CX))^2 =$$
$$= C^2 [M(X^2) - (M(X))^2] = C^2 D(X).$$

Таким образом,

$$D(CX) = C^2 D(X),$$

что и утверждалось.

Вычислим теперь дисперсию для случайных величин, распределенных по рассмотренным нами ранее законам.

I. Биномиальное распределение. Чтобы вычислить дисперсию случайной величины, распределенной по биномиальному закону, воспользуемся теоремой о дисперсии суммы случайных величин*. Пусть случайная величина X означает количество наступлений события A в серии из n испытаний, причем в каждом испытании вероятность наступления события равна p. Положим

$$X = X_1 + X_2 + \dots + X_n,$$

где X_i — число наступлений события A в i-м испытании. Это случайная величина, принимающая значение 1, если в i-м испытании событие A произошло, и 0, если оно не произошло. Закон распределения каждой из величин X_i одинаков и задается таблицей

x_i	0	1
p_i	$q = 1 - p$	p

Математическое ожидание X_i равно:

$$M(X_i) = 0 \cdot q + 1 \cdot p = p.$$

* Теорема состоит в том, что дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых. Точная формулировка и доказательство этой теоремы приведены в § 21.

Отсюда, кстати, пользуясь сформулированным выше свойством математического ожидания суммы, сразу видим, что

$$M(X) = np$$

(выше этот же результат был получен более сложным путем).

Дисперсия X_i равна (см. формулу (13)):

$$D(X_i) = (0 - p)^2 q + (1 - p)^2 p = pq(p + q) = pq.$$

Отсюда, по теореме о дисперсии суммы,

$$D(X) = npq. \tag{17}$$

II. Закон Пуассона. Дискретная случайная величина, распределенная по закону Пуассона, может принимать значения $0, 1, 2, \ldots, m, \ldots$ с вероятностями $P_m = \frac{\lambda^m}{m!} e^{-\lambda}$. Как было вычислено выше, $M(X) = \lambda$. Поэтому остается определить $M(X^2)$. По формуле (2)

$$M(X^2) = \sum_{m=0}^{\infty} m^2 P_m = \sum_{m-1}^{\infty} m^2 \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} m \frac{\lambda^m}{(m-1)!}.$$

Для вычисления суммы последнего ряда воспользуемся разложением

$$\lambda e^{\lambda} = \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!}.$$

Дифференцируя это равенство по λ, находим:

$$\lambda e^{\lambda} + e^{\lambda} = \sum_{m=1}^{\infty} m \frac{\lambda^{m-1}}{(m-1)!},$$

а после умножения на λ:

$$\sum_{m=1}^{\infty} m \frac{\lambda^m}{(m-1)!} = \lambda e^{\lambda} (\lambda + 1),$$

так что математическое ожидание квадрата

$$M(X^2) = \lambda^2 + \lambda.$$

Таким образом, дисперсия случайной величины X, распределенной по закону Пуассона, вследствие (15) равна:

$$D(X) = M(X^2) - (M(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda, \tag{18}$$

т. е. в этом случае *дисперсия равна математическому ожиданию.*

III. Равномерное распределение. Непрерывная случайная величина X, равномерно распределенная в интервале (a, b), имеет плотность вероятности

$$\varphi(x)=\begin{cases} 0 & \text{при} \quad -\infty < x < a, \\ \dfrac{1}{b-a} & \text{при} \quad a < x < b, \\ 0 & \text{при} \quad b < x < +\infty \end{cases}$$

и математическое ожидание $M(X)=\frac{a+b}{2}$. Для вычисления дисперсии найдем $M(X^2)$, пользуясь формулой (4):

$$M(X^2)=\int\limits_a^b \frac{x^2}{b-a}\,dx=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+ab+b^2}{3}.$$

Поэтому дисперсия равномерно распределенной случайной величины равна:

$$D(X)=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}. \tag{19}$$

Отметим, между прочим, что среднее квадратичное отклонение (стандарт) равно в этом случае

$$\sigma=\frac{b-a}{2\sqrt{3}}. \tag{20}$$

Таким образом, *для случайной величины, равномерно распределенной в интервале* (a, b), *среднее квадратичное отклонение равно* $\frac{1}{\sqrt{12}}=0{,}288675$ *длины интервала.*

IV. Нормальное распределение. Непрерывная случайная величина X, распределенная по закону Гаусса, имеет плотность вероятности

$$\varphi(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}},$$

причем $M(X)=a$.

Пользуясь формулой (14) для дисперсии непрерывной случайной величины, находим:

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}(x-a)^2\,e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx.$$

Для вычисления интеграла положим $\frac{x-a}{\sigma}=z$.

Тогда

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} z^2\sigma^2 e^{-\frac{z^2}{2}}\,dx=\frac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} z^2 e^{-\frac{z^2}{2}}\,dz.$$

Последний интеграл легко вычислить по частям, полагая

$$z = u, \quad ze^{-\frac{z^2}{2}} dz = dv,$$

откуда

$$du = dz, \quad v = -e^{-\frac{z^2}{2}},$$

так что

$$D(X) = \frac{\sigma^2}{\sqrt{2\pi}} \left[-ze^{-\frac{z^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz \right].$$

Как легко видеть, проинтегрированный член обращается в нуль, а

$$\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}.$$

Поэтому мы получаем:

$$D(X) = \sigma^2. \tag{21}$$

Мы выяснили теперь вероятностный смысл второго параметра, входящего в выражение закона Гаусса: *величина* σ^2 *равна дисперсии случайной величины* X.

§ 17. СТЕПЕНЬ НЕОПРЕДЕЛЕННОСТИ ДИСКРЕТНОГО РАСПРЕДЕЛЕНИЯ. ПОНЯТИЕ ОБ ЭНТРОПИИ

Рассмотрим дискретную случайную величину и будем понимать под «испытанием» получение ее очередного значения. Попробуем прогнозировать результат испытания, т. е. предсказать значение, которое примет случайная величина. Конечно, в силу случайности рассматриваемой величины исход испытания несет в себе некоторую неопределенность, но, зная закон распределения, мы можем пытаться в какой-то мере оценить надежность прогноза.

Естественно пытаться использовать для этой цели уже известные числовые характеристики — математическое ожидание и дисперсию. Однако оказывается, что для прогнозирования значения случайной величины эти характеристики бесполезны. В этом можно убедиться на следующем примере.

Пусть заданы две случайные величины, обладающие следующими законами распределения:

x	0	1
p	0,5	0,5

y	$\frac{1}{3}$	2
p	0,9	0,1

Найдем их математические ожидания и дисперсии. По формулам предыдущего параграфа получаем:

$$M(X) = 0 \cdot 0{,}5 + 1 \cdot 0\ 5 = 0{,}5,$$

$$M(Y) = \frac{1}{3} \cdot 0{,}9 + 2 \cdot 0{,}1 = 0{,}5.$$

Таким образом, математические ожидания этих случайных величин совпадают. То же относится и к дисперсиям. Действительно,

$$D(X) = M(X^2) - [M(X)]^2 = 0^2 \cdot 0{,}5 + 1^2 \cdot 0{,}5 - 0{,}5^2 = 0{,}25,$$

$$D(Y) = M(Y^2) - [M(Y)]^2 = \left(\frac{1}{3}\right)^2 \cdot 0{,}9 + 2^2 \cdot 0{,}1 - 0{,}5^2 = 0{,}25.$$

Однако интуитивно ясно, что закон распределения Y несет в себе значительно меньшую неопределенность и позволяет с довольно большой надежностью предсказать, что результатом испытания будет значение $Y = \frac{1}{3}$. Следовательно, ни математическое ожидание, ни дисперсия в качестве меры неопределенности распределения использованы быть не могут.

Числовой характеристикой дискретного распределения, которая может служить его мерой неопределенности, является *энтропия закона распределения*. Понятие энтропии было введено К. Шенноном в 1948—1950 годах, а название возникло вследствие существования глубоких аналогий между введенной характеристикой и используемым ранее в статистической физике понятием энтропии как функции состояния системы.

Энтропия дискретного распределения определяется формулой

$$H(p_1, p_2, \ldots, p_n) = -(p_1 \log p_1 + p_2 \log p_2 + \ldots + p_n \log p_n) = -\sum_{i=1}^{n} p_i \log p_i$$

в предположении, что $\sum_{i=1}^{n} p_i = 1$. (1)

Как видно из формулы (1), энтропия не зависит от значений, принимаемых случайной величиной, а только от их вероятностей. Рассматривая лишь те значения дискретной случайной величины, которые действительно возможны, можно считать, что все $p_i \neq 0$, так что функция H всегда определена. В тех случаях, когда приходится по тем или иным специальным причинам включать в рассмотрение значения $p_i = 0$, принимается, что соответствующее произведение равно нулю, $p_i \log p_i = 0$.

Основание логарифмов в формуле (1) может быть взято совершенно произвольно; нужно только позаботиться, чтобы для сравнения энтропий различных распределений они вычислялись при одном и том же основании. В теории информации, для которой понятие энтропии является центральным, принято в качестве основания логарифмов брать число 2.

Прежде чем пытаться объяснять, почему выбрана именно такая структура формулы (1), подсчитаем энтропию приведенных выше распределений. Для удобства подсчета логарифмы будем считать натуральными.

Для случайной величины X получаем:

$$H\ (0{,}5,\ 0{,}5) = -2 \cdot 0{,}5 \cdot \ln\ 0{,}5 = 0{,}6931.$$

Энтропия распределения Y равна:

$$H\ (0{,}9,\ 0{,}1) = -0{,}9\ \ln\ 0{,}9 - 0{,}1 \ln 0{,}1 = \\ = 0{,}1054 \cdot 0{,}9 + 2{,}3026 \cdot 0{,}1 = 0{,}3251,$$

т. е. энтропия распределения X более чем вдвое превосходит энтропию распределения Y. Это соотношение вполне согласуется с интуитивным представлением о неопределенности имеющихся распределений.

Рассмотрим пример, который будет полезен нам в дальнейшем.

Пример 1. Вероятность наступления события при одном испытании равна p. При каком p результат испытания обладает наибольшей неопределенностью?

Если случайная величина есть число наступлений события при данном испытании, то она описывается схемой

x_i	0	1
p_i	$1-p$	p

Согласно формуле (1) энтропия этой схемы равна:

$$H = -(1-p) \log\ (1-p) - p \log p.$$

Для удобства вычислений примем, что логарифмы берутся натуральными, т. е. примем

$$H = -(1-p) \ln (1-p) - p \ln p,$$

и найдем максимум этой функции Производная равна:

$$\frac{dH}{dp} = \ln\ (1-p) + 1 - \ln\ p - 1 = \ln\ (1-p) - \ln p.$$

Приравнивая $\frac{dH}{dp}$ нулю, находим $\ln(1-p) = \ln p$, откуда $p = 1 - p = \frac{1}{2}$. Так как $\frac{d^2H}{dp^2} = -\frac{1}{1-p} - \frac{1}{p} < 0$, то ясно, что при $p = \frac{1}{2}$ величина H имеет максимум.

Попробуем теперь разобраться в том, почему в качестве меры неопределенности распределения выбрана функция (1).

Степень неопределенности зависит прежде всего от числа возможных исходов испытания, т. е. от числа n различных возможных значений случайной величины. При выбранном n естественно считать, что наибольшая степень неопределенности соответствует такому распределению, при котором все исходы испытания равновероятны, т. е. вероятности всех возможных значений случайной величины равны между собою,

$$p_i = \frac{1}{n} \ (i = 1,\ 2,\ \ldots,\ n) \text{ (см. пример 1).}$$

Обозначим через $L(n)$ неопределенность такого распределения. Пусть теперь мы имеем две различные случайные величины, из которых первая имеет k, а вторая m равновероятных значений. Это означает, что в первом случае возможны k различных равновероятных исхода, а во втором случае — m равновероятных исходов. Если случайные величины независимы, то, рассматривая всевозможные пары их значений, мы найдем, что для них существует km различных равновероятных исходов.

Естественно считать, что неопределенность новой схемы, образованной из всевозможных пар, равна сумме неопределенностей отдельных схем. В терминах функции L это означает требование

$$L(km) = L(k) + L(m), \tag{2}$$

которое сразу же приводит нас к мысли использовать в качестве функции $L(n)$ логарифмическую функцию. Поэтому можно принять $L(n) = \log n = -\log \frac{1}{n}$. Последняя замена предпринята для того, чтобы перейти к вероятностям различных исходов вместо числа исходов.

Итак, мы положили, что общая неопределенность опыта с n равновероятными исходами равна $\log n$. Но тогда естественно считать, что каждый из исходов вносит неопределенность, равную

$$H_n = \frac{\log n}{n} = -\frac{1}{n} \log \frac{1}{n} = -p \log p,$$

поскольку $\frac{1}{n} = p$ и есть вероятность отдельного исхода в нашем случае. Остается принять условие, что для случайной величины, значения которой не равновероятны, отдельное значение, которое принимается с вероятностью p, вносит неопределенность, численная величина которой равна $-p \log p$ (знак минус нужно, очевидно, сохранить, чтобы получить положительную величину, поскольку $p \leqslant 1$ и $p \log p < 0$). Суммируя полученное выражение по всем возможным значениям случайной величины, мы и приходим к формуле (1).

Приведенные выше рассуждения, разумеется, могут лишь объяснить, почему в качестве меры неопределенности распределения выбрано выражение (1), но ничего не доказывают. Но на самом деле оказывается, что при некоторых естественных предположениях функция (1) является единственной. Это вытекает из следующей теоремы.

Теорема единственности. *Пусть функция* $H(p_1, p_2, \ldots, p_n)$ *определена при любом* n *на множестве систем* $\{p_k\}$, *для которых* $p_k \geqslant 0$ *и* $\sum_{k=1}^{n} p_k = 1$, *непрерывна по совокупности своих аргументов и удовлетворяет условиям:*

1. *при данном* n *функция* $H(p_1, p_2, \ldots, p_n)$ *достигает наибольшего значения при* $p_k = \frac{1}{n}$ $(k = 1, 2, \ldots, n)$;

2. *если* A *и* B — *две независимые* дискретные случайные величины, характеризуемые наборами вероятностей* $\{p_k\}$ *и* $\{p_k'\}$, *то*

$$H(AB) = H(A) + H(B); \tag{3}$$

3. *присоединение к значениям случайной величины одного или любого конечного числа невозможных значений не изменяет значения функции* H, *т. е.*

$$H(p_1, p_2, \ldots, p_n, 0) = H(p_1, p_2, \ldots, p_n). \tag{4}$$

Тогда

$$H(p_1, p_2, \ldots, p_n) = -\lambda \sum_{k=1}^{n} p_k \log p_k, \tag{5}$$

где λ — *постоянное положительное число.* Очевидно, что значения λ можно изменять произвольно за счет основания системы логарифмов или, наоборот, можно выбирать систему логарифмов с любым основанием, изменяя значение λ.

Доказательство этой теоремы, принадлежащей А. Я. Хинчину, мы приводить не будем.

Приведем еще один пример вычисления энтропии дискретного распределения.

Пример 2. Производится стрельба по двум мишеням. По первой мишени сделано два выстрела с вероятностью попадания при одном выстреле, равной $\frac{1}{2}$. По второй мишени сделано три выстрела с вероятностью попадания $\frac{1}{3}$. Определим для каждой мишени, какой исход стрельбы является более определенным.

Составим законы распределения числа попаданий при стрельбе по первой и второй мишеням, пользуясь биномиальным законом распределения:

$$P_{m,n} = \frac{n!}{m!\,(n-m)!}\, p^m q^{n-m}.$$

---

* Подробнее о независимости случайных величин будет сказано в § 18 главы IV (см. стр. 118 и 120).

Для первой мишени при $p = q = \frac{1}{2}$ находим:

$$P_{0,2} = \frac{2!}{0!\,2!}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^2 = \frac{1}{4},$$

$$P_{1,2} = \frac{2!}{1!\,1!}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^1 = \frac{1}{2},$$

$$P_{2,2} = \frac{2!}{2!\,0!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^0 = \frac{1}{4}.$$

Для второй мишени имеем $p = \frac{1}{3}$, $q = \frac{2}{3}$. Поэтому

$$P_{0,3} = \frac{3!}{0!3!}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^3 = \frac{8}{27},$$

$$P_{1,3} = \frac{3!}{1!2!}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^2 = \frac{4}{9},$$

$$P_{2,3} = \frac{3!}{2!1!}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^1 = \frac{2}{9},$$

$$P_{3,3} = \frac{3!}{3!0!}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^0 = \frac{1}{27}.$$

Таким образом, мы получаем законы распределения:

x	0	1	2
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

x	0	1	2	3
p	$\frac{8}{27}$	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{1}{27}$

Энтропия первого распределения (мы снова рассматриваем натуральные логарифмы) будет равна:

$$H_1 = -\left(\frac{1}{4}\ln\frac{1}{4} + \frac{1}{2}\ln\frac{1}{2} + \frac{1}{4}\ln\frac{1}{4}\right) = \frac{3}{2}\ln 2 \approx 1{,}04.$$

Та же величина для второго распределения:

$$H_2 = -\left(\frac{1}{27}\ln\frac{1}{27} + \frac{2}{9}\ln\frac{2}{9} + \frac{4}{9}\ln\frac{4}{9} + \frac{8}{27}\ln\frac{8}{27}\right) = \frac{7}{3}\ln 3 - 2\ln 2 \approx 1{,}18.$$

Так как $H_2 > H_1$, то следует считать, что исход стрельбы по первой мишени обладает большей определенностью, чем по второй.

Наряду с энтропией дискретного распределения в некоторых случаях рассматривают также и энтропию непрерывной случайной величины. По аналогии с формулой (1), энтропией непрерывной случайной величины X называют величину, определяемую формулой

$$H(X) = -\int_{-\infty}^{+\infty} \varphi(x) \log \varphi(x)\, dx, \tag{6}$$

где $\varphi(x)$ — плотность вероятности величины X, а логарифм, как и раньше, берется по произвольному основанию. Рассмотрим два примера.

Пример 3. Непрерывная случайная величина равномерно распределена в интервале (a, b). Определим энтропию этой случайной величины.

Как известно из § 15, плотность вероятности равномерно распределенной случайной величины имеет вид:

$$\varphi(x) = \begin{cases} 0 & \text{при} \quad -\infty < x < a, \\ \dfrac{1}{b-a} & \text{при} \quad a < x < b, \\ 0 & \text{при} \quad b < x < +\infty. \end{cases}$$

Поэтому, в соответствии с формулой (6), энтропия равномерного распределения будет равна:

$$H(X) = -\int_a^b \frac{1}{b-a} \ln \frac{1}{b-a}\, dx = \frac{\ln(b-a)}{b-a} x \Big|_a^b = \ln(b-a).$$

Можно доказать, что, по аналогии со случаем дискретного распределения, среди всех законов распределения непрерывных случайных величин, принимающих значения из интервала (a, b), максимальную энтропию имеет равномерное распределение. Как мы уже видели, эта максимальная энтропия равна:

$$H_{\max} = \ln(b-a). \tag{7}$$

Для случайных величин, которые могут принимать всевозможные действительные значения, такого заключения сделать нельзя.

Пример 4. Дана случайная величина X, распределенная по нормальному закону с математическим ожиданием $\bar{x} = 0$ и дисперсией σ^2. Вычислим энтропию этого распределения и сравним ее с энтропией равномерного распределения, обладающего той же дисперсией.

Плотность вероятности нормального распределения с $\bar{x} = 0$

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

Отсюда энтропия нормального распределения будет равна:

$$H_{\text{норм}} = -\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \left[\ln \frac{1}{\sigma\sqrt{2\pi}} - \frac{x^2}{2\sigma^2}\right] dx.$$

Разбивая последний интеграл на два слагаемых, имеем:

$$H_{\text{норм}} = \frac{\ln\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx + \frac{1}{2\sigma^2\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx.$$

Как уже известно (см. § 15 и § 16),

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx = 1; \qquad \int\limits_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx = \sigma^2.$$

Поэтому

$$H_{\text{норм}} = \ln\sqrt{2\pi\sigma^2} + \frac{1}{2},$$

или, иначе,

$$H_{\text{норм}} = \ln\left(\sigma\sqrt{2\pi e}\right). \tag{8}$$

Сравним теперь энтропии (7) и (8), для чего выразим энтропию (7) через дисперсию. В § 16 было вычислено, что дисперсия равномерного распределения равна:

$$\sigma_{\text{равн}} = \frac{b-a}{\sqrt{12}},$$

откуда $b - a = \sigma\sqrt{12}$. Подставляя это значение в формулу (7), находим для энтропии равномерного распределения с дисперсией σ:

$$H_{\text{равн}} = \ln\left(\sigma\sqrt{12}\right).$$

Так как $2\pi e \approx 17 > 12$, то имеет место неравенство

$$H_{\text{норм}} > H_{\text{равн}},$$

т. е. *нормальный закон распределения является более неопределенным, чем равномерный закон с той же дисперсией.*

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ К ГЛАВЕ III

1. Что такое дискретная случайная величина?

2. Дайте определение закона распределения дискретной случайной величины. Приведите примеры.

3. Будут ли дискретными следующие случайные величины:

а) число выбитых очков при стрельбе по мишени;

б) расстояние от точки попадания пули до центра мишени;

в) продолжительность работы данной электронной лампы в данном приборе;

г) количество электронных ламп в данном приборе, вышедших из строя за данный промежуток времени;

д) момент выхода из строя данной электронной лампы данного прибора?

4. Что называется функцией распределения случайной величины?

5. Можно ли определить функцию распределения и для дискретной, и для непрерывной случайных величин, или только для непрерывной, или только для дискретной?

6. Приведите точное определение непрерывной случайной величины.

7. Что называется плотностью вероятности случайной величины? Каков вероятностный смысл этой функции?

8. Можно ли определить плотность вероятности для дискретной случайной величины?

9. Как связаны между собою плотность вероятности и функция распределения непрерывной случайной величины?

10. Как выражается вероятность попадания случайной X величины в конечный интервал (α, β) через функцию распределения $F(x)$?

11. Как выражается вероятность попадания непрерывной случайной величины в конечный интервал (α, β) через плотность вероятности $\varphi(x)$?

12. Какой вид имеет функция распределения и плотность вероятности случайной величины, равномерно распределенной в интервале (a, b)?

13. Является ли случайная величина, распределенная по закону Пуассона, непрерывной или дискретной?

14. Изобразите графически плотность нормального закона распределения случайной величины. Выясните геометрический смысл параметров нормального закона.

15. Случайная величина распределена по нормальному закону. Найдите вероятности неравенств $a - k\sigma < X < a + k\sigma$ для $k = 1, 2, 3$.

16. Что называется математическим ожиданием дискретной случайной величины? Что характеризует математическое ожидание?

17. Что называется математическим ожиданием непрерывной случайной величины?

18. Приведите примеры нахождения математического ожидания случайных величин. Выясните физический смысл математического ожидания в каждом из приведенных примеров.

19. Дайте определение дисперсии случайной величины. Какие свойства случайной величины характеризует дисперсия?

20. Чему равны математическое ожидание и дисперсия случайной величины, распределенной по биномиальному закону?

21. Вычислите математическое ожидание и дисперсию случайной величины, распределенной по закону Пуассона.

22. Определите вероятностный смысл параметров a и σ нормального закона распределения.

23. Как определяется энтропия дискретного распределения?

ГЛАВА IV

МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 18. ДВУМЕРНАЯ СЛУЧАЙНАЯ ВЕЛИЧИНА. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТЬ ВЕРОЯТНОСТИ

Рассматривавшиеся в предыдущей главе случайные величины принимали различные числовые значения и геометрически изображались точками прямой. Часто приходится встречаться со случайными величинами более сложной природы, которые геометрически изображаются точками плоскости. Их называют *двумерными случайными величинами.*

Примером такой случайной величины может служить точка падения снаряда, если обстреливаемый участок рассматривать как плоскую область. Аналогично, при изучении плоского броуновского движения положение частицы, наблюдаемой под микроскопом, в данный момент времени есть двумерная случайная величина.

Вводя на рассматриваемой плоскости систему координат, мы можем каждую точку, являющуюся значением двумерной случайной величины, охарактеризовать парой чисел — ее координат. Каждая из координат, в свою очередь, является обычной (одномерной) случайной величиной. Поэтому двумерную случайную величину можно рассматривать как систему двух одномерных.

При дальнейшем изучении двумерных случайных величин мы будем использовать обе возможности, в зависимости от того, какая из них в каждом случае окажется более удобной для употребления. В одних случаях мы будем рассматривать точку M плоскости, в других — пару (X, Y) ее координат.

Двумерную случайную величину называют *дискретной,* если она может принимать конечное число или бесконечную последовательность различных значений. Для полной характеристики дискретной двумерной случайной величины достаточно указать множество возможных значений (точек плоскости) и вероятность каждого значения — *закон распределения.*

Закон распределения может быть представлен в форме таблицы с двумя входами (см. табл. 1). Здесь по горизонтали записаны значения, которые может принимать абсцисса двумерной случайной величины, а по вертикали — значения ординаты. В клетках таблицы приведены соответствующие вероятности того, что двумерная случайная величина попадет в данную точку плоскости.

Так, например, $P(x_i, y_j)$ означает вероятность того, что значением двумерной случайной величины окажется точка с координатами (x_i, y_j). Если, как было указано выше, рассматривать двумерную случайную величину как совокупность двух одномерных (X, Y), то $P(x_i, y_j)$ есть вероятность совмещения двух событий $X = x_i$, $Y = y_j$

Таблица 1

$Y \backslash X$	x_1	x_2	...	x_i	...	x_n	Σ
y_1	$P(x_1,y_1)$	$P(x_2,y_2)$	...	$P(x_i,y_1)$	...	$P(x_n,y_n)$	$P(y_1)$
y_2	$P(x_1,y_2)$	$P(x_2,y_2)$	...	$P(x_i,y_2)$	...	$P(x_n,y_2)$	$P(y_2)$
...	...	...	...	...	...	...	...
y_j	$P(x_1,y_j)$	$P(x_1,y_j)$	...	$P(x_i,y_j)$	...	$P(x_n,y_j)$	$P(y_j)$
...	...	...	...	...	...	...	...
y_m	$P(x_1,y_m)$	$P(x_2,y_m)$	...	$P(x_i,y_m)$	...	$P(x_n,y_m)$	$P(y_m)$
Σ	$P(x_1)$	$P(x_2)$	...	$P(x_i)$	...	$P(x_n)$	1

Предполагая, что все приведенные в таблице комбинации $X = x_i$, $Y = y_j$ являются единственно возможными, приходим к очевидному равенству:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i,y_j) = 1. \quad (1)$$

Имея таблицу вероятностей $P(x_i, y_j)$, можно найти вероятность того, что случайная величина X примет значение x_i, независимо от того, какое значение принимает Y.

По теореме сложения находим:

$$P(x_i) = P(X = x_i) = P(x_i, y_1) + P(x_i, y_2) + \ldots P(x_i, y_m) = \\ = \sum_{j=1}^{m} P(x_i, y_j). \qquad (2)$$

Таким образом, для нахождения вероятностей $P(X = x_i)$ нужно произвести суммирование вероятностей $P(x_i, y_j)$ в таблице по i-му столбцу. Эти вероятности и записаны в последней строке таблицы.

Аналогично, вероятность того, что случайная величина Y примет значение $Y = y_j$ независимо от значения X, получается суммированием вероятностей в j-й строке:

$$P(y_j) = P(Y = y_j) = \sum_{i=1}^{n} P(x_i, y_j). \qquad (3)$$

Эти вероятности образуют последний столбец таблицы. Как следует из равенства (1), сумма вероятностей в последнем столбце, как и в последней строке, должна равняться единице, которая и поставлена в правом нижнем углу.

Найдем еще условную вероятность того, что случайная величина Y примет значение $Y = y_j$ при условии, что $X = x_i$. Эту условную вероятность будем обозначать $P_{X=x_i}(Y = y_j)$, или, короче, $P(y_j/x_i)$. Согласно теореме умножения,

$$P(x_i, y_j) = P(x_i) P(y_j/x_i),$$

откуда получаем значение искомой условной вероятности:

$$P(y_j/x_i) = \frac{P(x_i, y_j)}{P(x_i)}. \qquad (4)$$

Если менять значения y_j ($j = 1, 2, \ldots, m$), оставляя значение $X = x_i$ неизменным, то $P(y_1/x_i)$, $P(y_2/x_i)$, ..., $P(y_m/x_i)$ выражают вероятности того, что случайная величина Y принимает соответственно значения $y_1, y_2, \ldots, y_m$ при однсм и том же значении $X = x_i$. Совокупность значений $\{y_j\}$ и соответствующие им вероятности $\{P(y_j/x_i)\}$ естественно назвать *условным законом распределения* Y при постоянном $X = x_i$. При изменении x_i ($i = 1, 2, \ldots, n$) будет соответственно изменяться и условный закон распределения.

Легко показать, что сумма условных вероятностей $P(y_j/x_i)$ при данном x_i равна единице. В самом деле, пользуясь формулой (4), получаем:

$$\sum_{j=1}^{m} P(y_j / x_i) = \sum_{j=1}^{m} \frac{P(x_i, y_j)}{P(x_i)} = \frac{\sum_{j=1}^{m} P(x_i, y_j)}{P(x_i)},$$

так как знаменатель не зависит от индекса суммирования j. Но, в силу формулы (2), числитель полученной дроби равен ее знаменателю. Таким образом,

$$\sum_{j=1}^{m} P(y_j/x_i) = 1. \tag{5}$$

Все сказанное относительно условного закона распределения Y при данном значении X целиком переносится и на случайную величину X при данном Y. Условная вероятность $P(x_i/y_j)$ того, что случайная величина примет значение $X = x_i$ при условии, что $Y = y_j$, находится по формуле

$$P(x_i / y_j) = \frac{P(x_i, y_j)}{P(y_j)}. \tag{6}$$

Значения $\{x_i\}$ вместе с вероятностями $\{P(x_i/y_j)\}$ образуют *условный закон распределения* X при постоянном $Y = y_j$. Сумма условных вероятностей, вследствие формул (6) и (3), равна единице:

$$\sum_{i=1}^{n} P(x_i / y_i) = 1. \tag{7}$$

Отметим в заключение, что если для любой пары возможных значений $X = x$, $Y = y$ справедливо равенство

$$P(x, y) = P(x) P(y), \tag{8}$$

то *случайные величины X и Y называют независимыми*.

Двумерная случайная величина, принимающая все значения из некоторой области G плоскости, называется *непрерывной* двумерной случайной величиной. Впрочем, это определение нуждается в уточнении, аналогичном сделанному в § 15. Для того чтобы двумерная случайная величина была непрерывной, необходимо еще дополнительно предположить, что она обладает непрерывной плотностью вероятности.

Плотность вероятности одномерной случайной величины X в точке x определялась нами в § 14 как функция $\varphi(x)$, удовлетворяющая условию

$$P(x < X < x + dx) \approx \varphi(x)\, dx. \tag{9}$$

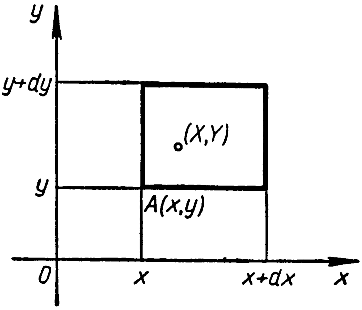


Рис. 23

Аналогично этому, для двумерной случайной величины (X, Y) *плотность вероятности определяется как функция $\varphi(x, y)$, удовлетворяющая в любой точке (x, y) условию*

$$P(x < X < x + dx,\ y < Y < y + dy) \approx \varphi(x, y)\, dx\, dy, \tag{10}$$

причем равенство (10) выполняется с точностью до бесконечно малых более высокого порядка, чем $dxdy$.

Геометрически произведение $\varphi(x, y)\, dxdy$ выражает вероятность попадания случайной величины в прямоугольник (рис. 23) с точностью до бесконечно малых более высокого порядка по сравнению с произведением $dxdy$.

Из определения плотности вероятности следует, что вероятность попадания двумерной случайной величины в некоторую область G равна:

$$P(M \in G) = \iint\limits_{G} \varphi(x, y)\, dx\, dy. \tag{11}$$

Действительно, разбив область G на прямоугольники и применив к каждому из них (10), представим по теореме сложения, искомую вероятность приближенно в форме двойной интегральной суммы. Переход к пределу при стремлении площадей прямоугольников к нулю приводит к формуле (11.) Если в с е возможные значения случайной величины принадлежат G, то, очевидно,

$$\iint\limits_{G} \varphi(x, y)\, dx\, dy = 1. \tag{12}$$

Это равенство во всяком случае справедливо, если G совпадает со всей плоскостью, так что для любой непрерывной двумерной случайной величины можно написать:

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} \varphi(x,y)\, dx\, dy. \tag{13}$$

Плотность вероятности двумерной случайной величины иначе называют *дифференциальным законом распределения.* Для двумерной случайной величины можно определить также *интегральный закон распределения,* или *двумерную функцию распределения,* как это делалось в предыдущей главе для одномерной случайной величины.

Значением функции распределения двумерной случайной величины в точке (x, y) *называют вероятность того, что случайная величина попадает левее и ниже этой точки, т. е. что ее координаты удовлетворяют неравенствам* $X < x$, $Y < y$,

$$F(x, y) = P(X < x, Y < y), \tag{14}$$

как показано на рис. 24. Из определения и формулы (11) следует, что

$$F(x,y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} \varphi(u,v)\, du\, dv. \tag{15}$$

Как было сказано выше, координаты X и Y точки, выражающей двумерную случайную величину, можно рассматривать как две

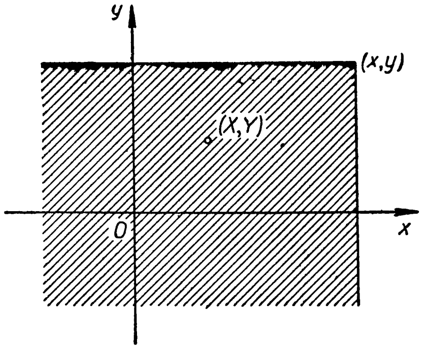


Рис. 24

случайные величины. Вообще говоря, случайные величины X и Y не являются независимыми.

При этом случайные величины X и Y называются *независимыми** , если, каковы бы ни были значения x и y, *события, заключающиеся в выполнении неравенств* $X < x$ *и* $Y < y$, *независимы* между собою.

Если предположить, что случайные величины X и Y независимы, то изучение двумерной случайной величины (X, Y) значительно упрощается. В самом деле, пусть случайная величина X имеет плотность вероятности $\varphi_1(x)$ и функция распределения $F_1(x) = \int\limits_{-\infty}^{x} \varphi_1(u)\,du$. Пусть, далее, случайная величина Y не зависит от X и имеет дифференциальный и интегральный законы распределения соответственно

$\varphi_2(y)$ и $F_2(y) = \int\limits_{-\infty}^{y} \varphi_2(v)\,dv$. Тогда вероятности события

$$x < X < x + dx, \qquad y < Y < y + dy$$

можно получить с помощью теоремы умножения как произведение вероятностей:

$$P(x < X < x + dx,\ y < Y < y + dy) =$$
$$= P(x < X < x + dx) \cdot P(y < Y < y + dy).$$

Сравнивая полученное выражение с (10), найдем, что

$$\varphi(x, y) = \varphi_1(x)\,\varphi_2(y). \qquad (16)$$

Аналогично этому для функции распределения получаем:

$$F(x,y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} \varphi_1(u)\,\varphi_2(v)\,du\,dv = \int\limits_{-\infty}^{x} \varphi_1(u)\,du \int\limits_{-\infty}^{y} \varphi_2(v)\,dv,$$

или

$$F(x, y) = F_1(x) \cdot F_2(y). \qquad (17)$$

Подчеркнем еще раз, что формулы (16) и (17) имеют место лишь в том случае, когда отдельные координаты двумерной случайной

* Заметим, что для дискретных случайных величин это определение уже было дано на стр. 118.

величины будут независимыми. Легко понять, что и обратно, если плотность вероятности и интегральный закон распределения двумерной случайной величины могут быть представлены соответственно в виде произведений (16) и (17), то отсюда следует, что координаты являются независимыми.

Пример 1. Двумерная случайная величина равномерно распределена в области G, площадь которой равна Q. Определим ее плотность вероятности.

В этом случае, как ясно из определения, плотность вероятности следует считать постоянной. Пусть $\varphi(x, y) = A$. Тогда, по формуле (12),

$$\iint_G A\,dx\,dy = 1,$$

а так как по известному свойству двойных интегралов

$$\iint_G dx\,dy = Q, \quad \text{то} \quad A = \frac{1}{Q}.$$

Пример 2. Двумерная случайная величина распределена по всей плоскости с плотностью

$$\varphi(x,y) = \frac{A}{(1+x^2)(1+y^2)} \tag{18}$$

(аналог закона Коши). Найдем двумерную функцию распределения $F(x, y)$ и вероятность попадания точки M в прямоугольник R, изображенный на рис. 25.

Определим сначала величину A. Согласно формуле (13),

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(x,y)\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{A\,dx\,dy}{(1+x^2)(1+y^2)} =$$

$$= A \operatorname{arctg} x \Big|_{-\infty}^{+\infty} \operatorname{arctg} y \Big|_{-\infty}^{+\infty} = 1,$$

т. е. $A\pi^2 = 1$, откуда $A = \frac{1}{\pi^2}$. Итак, плотность вероятности (18) должна иметь вид:

$$\varphi(x,y) = \frac{1}{\pi^2(1+x^2)(1+y^2)}.$$

Для функции распределения по формуле (15) находим:

$$F(x,y) = \frac{1}{\pi^2}\int_{-\infty}^{x}\int_{-\infty}^{y} \frac{du\,dv}{(1+u^2)(1+v^2)} = \frac{1}{\pi^2}\int_{-\infty}^{x}\frac{du}{1+u^2}\int_{-\infty}^{y}\frac{dv}{1+v^2} =$$

$$= \frac{1}{\pi^2}\left(\operatorname{arctg} x + \frac{\pi}{2}\right)\left(\operatorname{arctg} y + \frac{\pi}{2}\right) = \left(\frac{\operatorname{arctg} x}{\pi} + \frac{1}{2}\right)\left(\frac{\operatorname{arctg} y}{\pi} + \frac{1}{2}\right).$$

Вероятность попадания точки M в прямоугольник R находится с помощью (11):

$$P(M \in R) = \frac{1}{\pi^2}\int_0^a\int_0^b \frac{du dv}{(1+u^2)(1+v^2)} =$$

$$= \frac{1}{\pi^2}\int_0^a \frac{dy}{1+u^2}\int_0^b \frac{dv}{1+v^2} = \frac{\operatorname{arctg} a \operatorname{arctg} b}{\pi^2}.$$

Легко заметить, что случайные величины X и Y, являющиеся координатами точки M в данном примере, независимы. Действительно, плотность вероятности представляется в виде произведения:

$$\varphi(x,y) = \frac{1}{\pi(1+x^2)} \cdot \frac{1}{\pi(1+y^2)} = \varphi_1(x)\varphi_2(y).$$

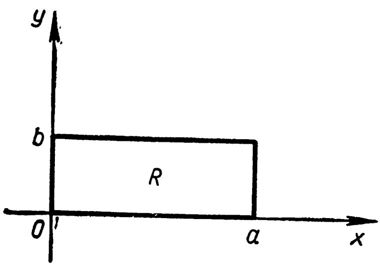


Рис. 25

Пример 3. Двумерная случайная величина распределена в первой четверти координатной плоскости ($0 \leqslant x < +\infty$, $0 \leqslant y < +\infty$) с плотностью вероятности

$$\varphi(x, y) = Ae^{-x-y}.$$

Найдем двумерную функцию распределения и вероятность попадания точки M в тот же прямоугольник R, изображенный на рис. 25.

Так как $\varphi(x, y)$, по определению, тождественно равна нулю, если $x < 0$ или $y < 0$, то формула (13) приобретает вид:

$$\int_0^\infty\int_0^\infty Ae^{-x-y}\,dx\,dy = 1,$$

откуда находим:

$$A\int_0^\infty e^{-x}dx\int_0^\infty e^{-y}dy = Ae^{-x}\Big|_0^\infty e^{-y}\Big|_0^\infty = A = 1.$$

Далее, формула (15) дает:

$$F(x,y) = \int_0^x\int_0^y e^{-u-v}du dv = \left[-e^{-u}\right]_0^x \cdot \left[-e^{-v}\right]_0^y =$$

$$= (1-e^{-x})(1-e^{-y}).$$

Вероятность попадания в прямоугольник R будет равна:

$$P(M \in R) = \int_0^a \int_0^b e^{-u-v} du\, dv = \left[-e^{-u}\right]_0^a \cdot \left[-e^{-v}\right]_0^b =$$

$$= (1 - e^{-a})(1 - e^{-b}).$$

§ 19. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ ДВУМЕРНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Роль нормального закона распределения для двумерных случайных величин так же велика, как и в разобранном ранее одномерном случае. Поэтому мы выделили рассмотрение нормального закона в отдельный параграф.

Начнем с того случая, когда координаты X и Y двумерной случайной величины независимы и распределены по нормальному закону, каждая со своим средним значением соответственно a и b и дисперсиями соответственно σ_x и σ_y. В силу независимости X и Y, плотность распределения точки (X, Y) может быть выражена в виде произведения:

$$\varphi(x,y) = \varphi_1(x)\varphi_2(y) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma_x^2}} \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-b)^2}{2\sigma_y^2}},$$

или

$$\varphi(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-a)^2}{2\sigma_x^2} - \frac{(y-b)^2}{2\sigma_y^2}}. \qquad (1)$$

Формула (1) и выражает двумерный нормальный закон распределения в простейшей форме. Точку плоскости с координатами (a, b) обычно называют *центром нормального распределения* или *центром рассеяния.*

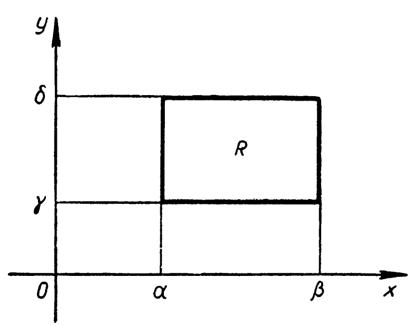


Рис. 26

Пример 1. Двумерная случайная величина распределена по нормальному закону (1) с $a = b = 0$. Определить вероятность ее попадания в прямоугольник R $(\alpha < x < \beta,\ \gamma < y < \delta)$, изображенный на рис. 26.

По формуле (11) предыдущего параграфа получаем:

$$P(M \in R) = P(\alpha < X < \beta, \gamma < Y < \delta) =$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \int_\alpha^\beta \int_\gamma^\delta e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}} dx\, dy =$$

$$= \frac{1}{\sigma_x\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{x^2}{2\sigma_x^2}} dx \cdot \frac{1}{\sigma_y\sqrt{2\pi}} \int_\gamma^\delta e^{-\frac{y^2}{2\sigma_y^2}} dy.$$

Каждый из полученных интегралов может быть выражен через функцию Лапласа. Положив $\frac{x}{\sigma_x} = t$, получим:

$$\frac{1}{\sigma_x\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{x^2}{2\sigma_x^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha}{\sigma_x}}^{\frac{\beta}{\sigma_x}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{\beta}{\sigma_x}\right) - \Phi\left(\frac{\alpha}{\sigma_x}\right).$$

Преобразовав аналогичным образом и второй интеграл, получим:

$$P(M \in R) = \left[\Phi\left(\frac{\beta}{\sigma_x}\right) - \Phi\left(\frac{\alpha}{\sigma_x}\right)\right]\left[\Phi\left(\frac{\delta}{\sigma_y}\right) - \Phi\left(\frac{\gamma}{\sigma_y}\right)\right]. \qquad (2)$$

Равенство (2) можно было бы получить непосредственно, пользуясь формулами § 14 и теоремой умножения.

Для двумерных случайных величин плотность нормального закона распределения вместо (1) чаще записывают в другой форме, вводя вместо дисперсии другую числовую характеристику.

Пусть одномерная случайная величина X распределена по нормальному закону с математическим ожиданием a и дисперсией σ_x^2. Найдем такую величину E_x, чтобы вероятность отклонения случайной величины X от a не более чем на E_x, т. е. вероятность по-

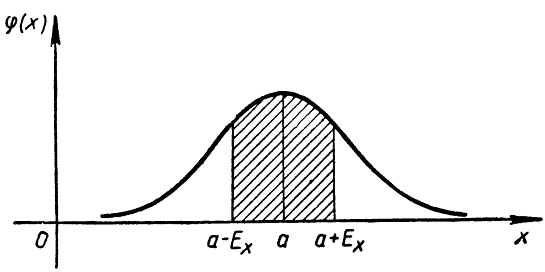


Рис. 27

падания X в интервал $(a - E_x,\ a + E_x)$, равнялась половине (см. рис. 27). Иначе говоря, с равной вероятностью случайная величина окажется внутри интервала $(a - E_x,\ a + E_x)$ или вне его. Эту величину E_x называют *вероятным* или *срединным отклонением.*

Ясно, что величина E_x характеризует рассеяние X, а потому должна быть связана с дисперсией σ_x^2. Найдем эту связь. Из определения E_x следует:

$$P(-E_x < X - a < E_x) = \frac{1}{2},$$

или

$$P(a - E_x < X < E_x + a) = \frac{1}{2}\cdot$$

Пользуясь функцией Лапласа, получаем:

$$P(a - E_x < X < a + E_x) = 2\Phi\left(\frac{E_x}{\sigma_x}\right) = \frac{1}{2},$$

откуда $\Phi\left(\frac{E_x}{\sigma_x}\right) = 0{,}25$. Из таблицы значений функции Лапласа находим:

$$E_x = 0{,}6745\sigma_x. \tag{3}$$

Часто вместо (3) пишут также

$$E_x = \rho\sqrt{2}\,\sigma_x, \tag{4}$$

где

$$\rho = 0{,}4969.$$

Введя в выражение (1) срединные отклонения E_x и E_y вместо дисперсий, получим выражение для плотности вероятности нормального распределения:

$$\varphi(x,y) = \frac{\rho^2}{\pi E_x E_y} e^{-\rho^2\left[\frac{(x-a)^2}{E_x^2} + \frac{(y-b)^2}{E_y^2}\right]}. \tag{5}$$

Выражение (5) часто называют *канонической формой* закона распределения Гаусса.

Обозначив для краткости $\frac{\rho^2}{\pi E_x E_y} = A$ и приняв $a = b = 0$, изучим форму поверхности

$$z = Ae^{-\rho^2\left[\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2}\right]}. \tag{6}$$

При любых x и y имеем $z > 0$, так что поверхность расположена всегда над плоскостью xOy. В точке $x = y = 0$ функция $z = \varphi\ (x, y)$ достигает максимума, равного A. Для дальнейшего исследования поверхности рассмотрим ее сечения плоскостями yOz ($x = 0$) и xOz ($y = 0$). Тогда получим соответственно:

$$z = Ae^{-\rho^2 \frac{y^2}{E_y^2}} \quad (\text{при } x = 0),$$

$$z = Ae^{-\rho^2 \frac{x^2}{E_x^2}} \quad (\text{при } y = 0).$$

Эти сечения представляют собою нормальные кривые распределения (рис. 28). Сечение, параллельное плоскости xOy, получим: положив $z = H$. Тогда

$$H = Ae^{-\rho^2\left(\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2}\right)}.$$

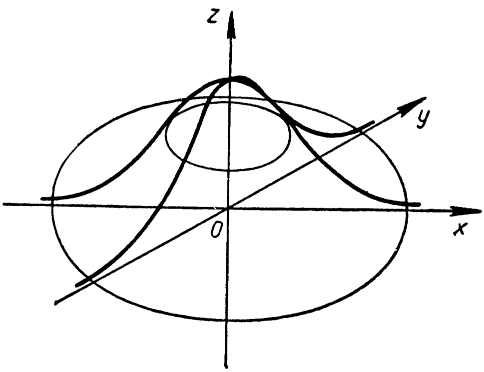


Рис. 28

Логарифмируя последнее равенство и обозначив

$\rho^2 \ln \frac{A}{H} = K^2$*, получим:

$$\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2} = K^2,$$

или

$$\frac{x^2}{(KE_x)^2} + \frac{y^2}{(KE_y)^2} = 1. \quad (7)$$

Уравнения (7) определяют семейство подобных эллипсов, полуоси которых возрастают при убывании H пропорционально вероятным отклонениям E_x, E_y. Эти эллипсы называют *эллипсами рассеяния*. При $K = 0$, то есть при $H = A$, эллипс вырождается в точку. Эллипс, получающийся при $K = 1$, называют *единичным*, а при $K = 4$ — *полным эллипсом рассеяния*. Их уравнения имеют вид:

$$\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2} = 1,$$

$$\frac{x^2}{16E_x^2} + \frac{y^2}{16E_y^2} = 1.$$

Если, в частности, срединные ошибки одинаковы ($E_x = E_y$), то эллипсы рассеяния обращаются в окружности. В этом случае рассеяние называют *круговым*.

---

* Заметим, что $A > H$, так что $\frac{A}{H} > 1$ и $\ln \frac{A}{H} > 0$.

Важно подчеркнуть, что *главные оси эллипсов рассеяния параллельны координатным осям,* как это видно из уравнения (7).

Пример 2. Двумерная случайная величина нормально распределена на плоскости в соответствии с канонической формой закона Гаусса (5). Определим вероятность ее попадания в эллипс рассеяния α_K, заданный уравнением (7).

Как известно, эта вероятность равна:

$$P(M \in \alpha_K) = \frac{\rho^2}{\pi E_x E_y} \iint\limits_{\alpha_K} e^{-\rho^2\left(\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2}\right)} dxdy. \tag{8}$$

Для вычисления интеграла (8) введем новые переменные:

$$u = \frac{\rho x}{E_x}, \quad v = \frac{\rho y}{E_y},$$

или, иначе,

$$x = \frac{uE_x}{\rho}, \quad y = \frac{vE_y}{\rho}.$$

Подставляя полученные значения x, y в уравнение (7), находим:

$$u^2 + v^2 = K^2\rho^2;$$

это означает, что эллипс рассеяния α_K перешел в круг C_K радиуса $K\rho$. Далее, $dxdy = \frac{E_xE_y}{\rho^2}$; следовательно, интеграл (8) будет иметь вид:

$$P(M \in \alpha_K) = \frac{1}{\pi} \iint\limits_{C_K} e^{-(u^2+v^2)} dudv.$$

Полученный интеграл преобразуем к полярным координатам. Положив $u = r\cos\vartheta$, $v = r\sin\vartheta$, найдем $u^2 + v^2 = r^2$ и $dudv = rdrd\vartheta$. Поэтому

$$P(M \in \alpha_K) = \frac{1}{\pi}\int_0^{2\pi} d\vartheta \int_0^{K\rho} e^{-r^2} rdr = -2\int_0^{K\rho} e^{-r^2} \frac{1}{2} d(-r^2).$$

Вычисляя последний интеграл, получаем формулу для попадания в эллипс рассеяния:

$$P(M \in \alpha_K) = 1 - e^{-K^2\rho^2}. \tag{9}$$

Формулу (9) можно, в частности, использовать для нахождения вероятности попадания в **единичный** эллипс и **полный** эллипс рассеяния. Как было сказано выше, они характеризуются соответственно значениями $K = 1$ и $K = 4$. Следовательно, по формуле (9) получаем:

$$P(M \in \alpha_1) = 1 - e^{-\rho^2} = 0{,}2034,$$

$$P(M \in \alpha_2) = 1 - e^{-16\rho^2} = 0{,}9737,$$

так как $\rho = 0{,}4769$.

Пример 3. Рассмотрим двумерную случайную величину, распределенную по нормальному закону с круговым рассеянием ($E_x = E_y = E$). Расстояние случайной точки плоскости от центра рассеяния, который можно предположить лежащим в начале координат, есть одномерная случайная величина:

$$R = \sqrt{X^2 + Y^2} \geqslant 0.$$

Определим функцию распределения и плотность вероятности случайной величины R, а также ее математическое ожидание.

Обозначив через

$$F(r) = P(R < r)$$

функцию распределения случайной величины R, с помощью формулы (9) находим:

$$P(R < r) = 1 - e^{-K^2\rho^2}.$$

Из уравнения (7) следует, что $EK = r$, т. е. $K = \frac{r}{E}$. Таким образом,

$$F(r) = 1 - e^{-\left(\rho\frac{r}{E}\right)^2}. \tag{10}$$

Для получения плотности вероятности достаточно продифференцировать формулу (10) по r. Мы получим:

$$\varphi(r) = 2r\frac{\rho^2}{E^2}e^{-\rho^2\frac{r^2}{E^2}}. \tag{11}$$

Теперь можно найти математическое ожидание случайной величины R, которое равно:

$$M(R) = \int_{-\infty}^{+\infty} r\varphi(r)\,dr.$$

Подставив сюда выражение (11) для плотности и учитывая, что при $r < 0$ плотность вероятности равна нулю, получаем:

$$M(R) = \int_0^\infty 2\frac{\rho^2}{E^2}r^2 e^{-\rho\frac{r^2}{E^2}}dr.$$

Последний интеграл вычисляем по частям, полагая

$$r = u,\ 2r\frac{\rho^2}{E^2}e^{-r^2\frac{\rho^2}{E^2}}dr = dv. \text{ Тогда } du = dr,\ v = -e^{-r^2\frac{\rho^2}{E^2}}$$

и

$$M(R) = -re^{-r^2\frac{\rho^2}{E^2}}\Big|_0^\infty + \int_0^\infty e^{-r^2\frac{\rho^2}{E^2}}dr.$$

Первое слагаемое обращается в нуль. В оставшемся интеграле полагаем $r\frac{\rho}{E} = \frac{t}{\sqrt{2}}$, после чего находим:

$$M(R) = \frac{E}{\rho\sqrt{2}}\int_0^\infty e^{-\frac{t^2}{2}}dt.$$

Но этот последний интеграл был уже вычислен нами в § 10. Используя полученный там результат, а также формулу (4), находим окончательно:

$$M(R) = \sigma\sqrt{\frac{\pi}{2}} \approx 1{,}253\sigma.$$

Для построения графика функции распределения найдем значение r, соответствующее максимальному значению плотности вероятности*. Дифференцируя формулу (11), находим:

$$\varphi'(r) = 2\frac{\rho^2}{E^2} e^{-\rho^2\frac{r^2}{E^2}}\left(1 - 2r^2\frac{\rho^2}{E^2}\right).$$

Уравнение $\varphi'(r) = 0$ имеет корень $r = \frac{E}{\rho\sqrt{2}} = \sigma$ (см. формулу (4)). Исследование достаточных условий показывает, что плотность вероятности имеет максимум при $r = \sigma$. График функции $\varphi(r)$ показан на рис. 29.

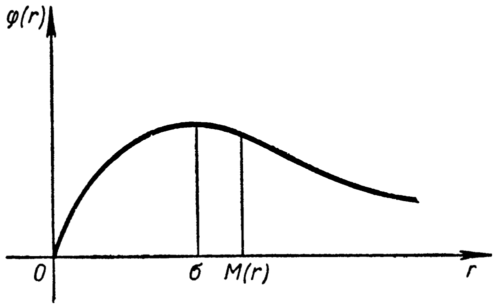


Рис. 29

До сих пор мы ограничивались рассмотрением частного случая двумерного нормального распределения, в котором эллипс рассеяния имеет оси, параллельные координатным осям.

Вообще, двумерная случайная величина называется распределенной по нормальному закону, если плотность вероятности распределения задается формулой

$$\varphi(x, y) = Re^{-Q(x, y)}, \tag{12}$$

где $Q(x, y)$ есть *положительно определенная квадратичная форма* относительно переменных $u = x - a$, $v = y - b$; a, b означают координаты центра рассеяния.

Напомним, что *квадратичной формой* в алгебре называется однородный многочлен второй степени. Квадратичная форма называется *положительно определенной*, если она при всех значениях аргумента принимает только положительные значения. В курсе высшей алгебры доказывается, что положительно определенная квадратичная форма может быть представлена в виде:

$$Q(x, y) = \frac{(x-a)^2}{2A^2} - r\frac{(x-a)(y-b)}{AB} + \frac{(y-b)^2}{2B^2},$$

где A, B — некоторые положительные числа, а число r удовлетворяет условию $-1 \leqslant r \leqslant +1$.

Коэффициент R в формуле (12) находится из условия, что интеграл от плотности вероятности равен единице. Предположив, что $|r| \neq 1$, и вводя вместо A и B новые величины σ_1, σ_2 по формулам

$$A = \sigma_1\sqrt{1-r^2}, \quad B = \sigma_2\sqrt{1-r^2},$$

а затем вычислив нормировочный коэффициент R, запишем плотность вероятности двумерного нормального распределения в виде

$$\varphi(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2} - 2r\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2}\right]} \tag{13}$$

О теоретико-вероятностном смысле параметров σ_1, σ_2 и r речь будет идти в следующем параграфе.

---

* Значение случайной величины, соответствующее максимуму ее плотности вероятности, называется *модой*.

Приравнивая квадратную скобку в показателе степени положительной постоянной, получим, как и выше, уравнение *эллипса рассеяния*:

$$\frac{(x-a)^2}{\sigma_1^2} - 2r\,\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2} = K,$$

по-прежнему имеющего центром точку (a, b), но здесь его оси уже не будут параллельны координатным осям, если только $r \neq 0$.

§ 20. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СИСТЕМЫ ДВУХ СЛУЧАЙНЫХ ВЕЛИЧИН

Числовые характеристики двумерной случайной величины удобно вводить, рассматривая двумерную случайную величину как систему двух одномерных.

Покажем сначала, как, зная двумерную плотность распределения $\varphi(x, y)$, найти плотность распределения случайной величины X. Аналогичный вопрос нахождения закона распределения случайной величины X по известному двумерному закону применительно к дискретной случайной величине (X, Y) был уже рассмотрен нами ранее, в § 18. Здесь мы займемся непрерывными случайными величинами.

Если проинтегрировать плотность $\varphi(x, y)$ по всем возможным значениям величины Y (т. е., вообще говоря, от $-\infty$ до $+\infty$), то, вследствие теоремы сложения, получим плотность вероятности того, что случайная величина X примет значения из интервала $(x, x + dx)$. Обозначив эту плотность через $\varphi_1(x)$, можем написать:

$$\varphi_1(x) = \int_{-\infty}^{+\infty} \varphi(x, y)\, dy. \tag{1}$$

Точно так же плотность вероятности случайной величины Y равна

$$\varphi_2(y) = \int_{-\infty}^{+\infty} \varphi(x, y)\, dx. \tag{2}$$

Формулы (1) и (2) соответствуют формулам (1) и (2) из § 18 для дискретной двумерной случайной величины, с той разницей, что входящие туда суммы заменены здесь интегралами.

Имея плотности вероятностей X и Y, нетрудно найти соответствующие интегральные законы распределения. Их можно записать в виде

$$F_1(x) = \int_{-\infty}^{x} \varphi_1(u)\, du = \int_{-\infty}^{x} du \int_{-\infty}^{+\infty} \varphi(u, v)\, dv, \tag{3}$$

$$F_2(y) = \int_{-\infty}^{y} \varphi_2(v)\,dv = \int_{-\infty}^{y} dv \int_{-\infty}^{+\infty} \varphi(u, v)\,du. \tag{4}$$

Можно ввести также условные законы распределения, подобно тому как это было сделано в § 18 для дискретных случайных величин. Обозначим через $\varphi(y/x)$ плотность вероятности того, что Y находится в интервале $(y, y+dy)$ при условии, что X находится в интервале $(x, x+dx)$. Аналогичный смысл имеет также и обозначение $\varphi(x/y)$. Тогда по теореме умножения найдем:

$$\varphi(x, y)\,dxdy = \varphi_1(x)\,dx\varphi(y/x)\,dy = \varphi_2(y)\,dy\,\varphi(x/y)\,dx.$$

Воспользовавшись формулами (1) и (2) для плотностей вероятности, получим:

$$\varphi(y/x) = \frac{\varphi(x, y)}{\varphi_1(x)} = \frac{\varphi(x, y)}{\int\limits_{-\infty}^{+\infty} \varphi(x, y)\,dy}, \tag{5}$$

$$\varphi(x/y) = \frac{\varphi(x, y)}{\varphi_2(y)} = \frac{\varphi(x, y)}{\int\limits_{-\infty}^{+\infty} \varphi(x, y)\,dx}. \tag{6}$$

Величины $\varphi(y/x)$ и $\varphi(x/y)$, выраженные формулами (5) и (6), носят название *условных плотностей вероятности*.

Пример 1. Случайная точка $M(X, Y)$ попадает с одинаковой вероятностью в любую точку эллипса с полуосями a, b, совпадающими с осями координат. Требуется определить плотность вероятности случайных величин X и Y, а также условные плотности вероятности $\varphi(y/x)$ и $\varphi(x/y)$.

Для непрерывной двумерной случайной величины, равномерно распределенной в области G, как мы уже вывели в примере 2 из § 18, плотность вероятности постоянна и равна $\varphi(x, y) = \frac{1}{Q}$ для $M \in G$, где Q означает площадь области G. Как известно из интегрального исчисления, площадь эллипса с полуосями a и b равна $Q = \pi ab$. Поэтому в нашем случае для двумерной плотности вероятности получаем:

$$\varphi(x, y) = \begin{cases} \dfrac{1}{\pi ab} \text{ при } \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} < 1, \\ 0 \text{ при } \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} > 1. \end{cases}$$

Если при этом x меняется на отрезке $-a \leqslant x \leqslant a$, то соответствующей областью изменения y при каждом x будет отрезок

$$-b\sqrt{1 - \frac{x^2}{a^2}} \leqslant y \leqslant b\sqrt{1 - \frac{x^2}{a^2}}.$$

Плотность вероятности случайной величины X можно получить по формуле (1) при дополнительном условии $|x| \leqslant a$:

$$\varphi_1(x) = \int_{-\infty}^{+\infty} \varphi(x, y)\, dy = \int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} \frac{1}{\pi ab}\, dy = \frac{2}{\pi a^2}\sqrt{a^2-x^2}.$$

При $|x| > a$ имеем $\varphi_1(x) \equiv 0$.

Аналогично находится плотность вероятности Y при условии $|y| \leqslant b$:

$$\varphi_2(y) = \int_{-\infty}^{+\infty} \varphi(x, y)\, dx = \int_{-\frac{a}{b}\sqrt{b^2-y^2}}^{\frac{a}{b}\sqrt{b^2-y^2}} \frac{1}{\pi ab}\, dx = \frac{2}{\pi b^2}\sqrt{b^2-y^2}.$$

Для $|y| > b$, как и выше, имеем $\varphi_2(y) \equiv 0$.

Найдем далее условную плотность вероятности $\varphi(y/x)$. Применяя формулу (5), имеем:

$$\varphi(y/x) = \frac{\varphi(x, y)}{\int_{-\infty}^{+\infty} \varphi(x, y)\, dx} = \frac{\frac{1}{\pi ab}}{\frac{2}{\pi a^2}\sqrt{a^2-x^2}} = \frac{a}{2b\sqrt{a^2-x^2}},$$

при условии $|x| < a$ и $|y| \leqslant \frac{b}{a}\sqrt{a^2-x^2}$. Если же $|x| \geqslant a$ или же $|y| > \frac{b}{a}\sqrt{a^2-x^2}$, то $\varphi(y/x) \equiv 0$. Аналогичные вычисления дают:

$$\varphi(x/y) = \frac{b}{2a\sqrt{b^2-y^2}}$$

при условии $|y| < b$ и $|x| < \frac{a}{b}\sqrt{b^2-y^2}$,
и

$$\varphi(x/y) \equiv 0,$$

если хотя бы одно из этих двух неравенств не выполняется.

Пусть теперь $M(X, Y)$ — двумерная случайная величина с плотностью вероятности $\varphi(x, y)$. Рассмотрим отдельно математические ожидания и дисперсии случайных величин X и Y.

Математические ожидания случайных величин X и Y, в соответствии с определениями из § 16, находятся по формулам:

$$M(X) = \bar{x} = \int_{-\infty}^{+\infty} x\varphi_1(x)\, dx, \quad M(Y) = \bar{y} = \int_{-\infty}^{+\infty} y\varphi_2(y)\, dy.$$

Пользуясь формулами (1) и (2), можно выразить эти математические ожидания через двумерную плотность вероятности $\varphi(x, y)$:

$$\left.\begin{aligned} M(X) &= \bar{x} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\varphi(x, y)\,dx\,dy, \\ M(Y) &= \bar{y} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\varphi(x, y)\,dx\,dy. \end{aligned}\right\} \tag{7}$$

Точка M с координатами $(\bar{x}, \bar{y})$, определенными по формуле (7), характеризует *центр рассеяния* двумерной случайной величины, около которой рассеяны случайные точки $M(X, Y)$. Меры рассеяния по осям задаются дисперсиями одномерных случайных величин:

$$\left.\begin{aligned} D(X) &= \int_{-\infty}^{+\infty} (x - \bar{x})^2 \varphi_1(x)\,dx = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - \bar{x})^2 \varphi(x, y)\,dx\,dy, \\ D(Y) &= \int_{-\infty}^{+\infty} (y - \bar{y})^2 \varphi_2(y)\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (y - \bar{y})^2 \varphi(x, y)\,dx\,dy. \end{aligned}\right\} \tag{8}$$

Как и для одномерного случая, разлагая подынтегральную функцию на слагаемые, можно получить несколько иные выражения для дисперсий, более удобные для вычислений:

$$\left.\begin{aligned} D(X) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 \varphi(x, y)\,dx\,dy - \bar{x}^2, \\ D(Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2 \varphi(x, y)\,dx\,dy - \bar{y}^2. \end{aligned}\right\} \tag{9}$$

Вычисления математических ожиданий и дисперсий, в соответствии с формулами (7) и (9), можно объединить в одну вычислительную схему. Для этой цели введем понятие центрального момента системы случайных величин.

Центральным моментом порядка k, s системы случайных величин (X, Y) будем называть математическое ожидание произведения $X^k Y^s$. Для непрерывных случайных величин центральный момент порядка k, s находится по формуле

$$\mu_{k,s} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - \bar{x})^k (y - \bar{y})^s \varphi(x, y)\,dx\,dy \tag{10}$$

(где k и s—неотрицательные целые числа).

Сумму $k + s$ называют *порядком центрального момента*, так что все моменты порядка k, s, для которых, например, $k + s = 1$, называют моментами первого порядка, а все моменты, для которых $k + s = 2$, — моментами второго порядка. Рассмотрением только таких моментов мы и ограничимся.

Из формул (7) вытекает, что *центральные моменты первого порядка равны нулю*. Действительно, так как k и s — целые неотрицательные, то существуют только два различных центральных момента первого порядка — порядка 1, 0 и порядка 0, 1:

$$\mu_{1,0}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\bar{x})\,\varphi(x,y)\,dx\,dy$$

и

$$\mu_{0,1}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\bar{y})\,\varphi(x,y)\,dx\,dy.$$

Так как $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\varphi(x,y)\,dx\,dy=1$ (см. формулу (4) из § 18), то

$$\mu_{1,0}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x\,\varphi(x,y)\,dx\,dy-\bar{x},$$

что равно нулю по определению x (см. формулу (7)). По той же причине $\mu_{0,1}=0$.

Определение (8) дисперсий показывает, что дисперсии $D(X)$ и $D(Y)$ представляют собою центральные моменты второго порядка, именно, порядка соответственно 2, 0 и 0, 2. Особую роль играет смешанный центральный момент второго порядка, т. е. момент порядка 1, 1. Его называют *корреляционным моментом* или *моментом связи* случайных величин X и Y и обозначают K_{xy}. В соответствии с определением (10), имеем:

$$K_{xy}=\mu_{1,1}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\bar{x})(y-\bar{y})\,\varphi(x,y)\,dx\,dy. \qquad (11)$$

Разбивая интеграл (11) на отдельные слагаемые и учитывая равенство нулю центральных моментов первого порядка, можно получить другое выражение для корреляционного момента. Именно,

$$K_{xy}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x(y-\bar{y})\,\varphi(x,y)\,dx\,dy-\bar{x}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\bar{y})\varphi(x,y)dx\,dy=$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\,\varphi(x,y)\,dx\,dy-\bar{y}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x\,\varphi(x,y)\,dx\,dy-\bar{x}\,\mu_{0,1}.$$

Окончательно получаем:

$$K_{xy}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\,\varphi(x,y)\,dx\,dy-\bar{x}\,\bar{y}. \qquad (12)$$

Пример 2. Точка $M(X, Y)$ распределена в квадрате $0\leqslant x\leqslant\frac{\pi}{2}$, $0\leqslant y\leqslant\frac{\pi}{2}$ с плотностью вероятности, равной $\varphi(x, y)=0{,}5\sin(x+y)$ внутри и $\varphi(x, y)=0$ вне этого квадрата. Найдем:

а) двумерную функцию распределения $F(x, y)$;

б) математические ожидания $M(X)$, $M(Y)$ случайных величин X и Y, т. е. координаты центра рассеяния;

в) дисперсии $D(X)$ и $D(Y)$;

г) корреляционный момент K_{xy}.

Решение. а) Согласно формуле (5) из § 18,

$$F(x, y) = P(X < x, Y < y) = \frac{1}{2}\int_0^x\int_0^y \sin(u+v)\,du\,dv =$$

$$= \frac{1}{2}\int_0^x \Big[-\cos(u+v)\Big]_{v=0}^{v=y} du = \frac{1}{2}\int_0^x [\cos u - \cos(u+y)]\,du =$$

$$= \frac{1}{2}\Big[\sin u - \sin(u+y)\Big]_{u=0}^{u=x} = \frac{1}{2}[\sin x - \sin(x+y) + \sin y].$$

б) Математическое ожидание $M(X)$ находится по формуле

$$M(X) = \frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} x\sin(x+y)\,dx\,dy = \frac{1}{2}\int_0^{\frac{\pi}{2}} x\,dx\int_0^{\frac{\pi}{2}}\sin(x+y)\,dy =$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{2}} x\left[\cos x - \cos\left(x + \frac{\pi}{2}\right)\right]dx.$$

Заменяя $\cos\left(x + \frac{\pi}{2}\right) = -\sin x$ и интегрируя по частям, находим окончательно $M(X) = \frac{\pi}{4}$. Аналогично для $M(Y)$;

$$M(Y) = \frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} y\sin(x+y)\,dx\,dy = \frac{\pi}{4},$$

так что центром рассеяния является в данном случае точка $\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$, т. е. центр квадрата.

в) Для вычисления дисперсий воспользуемся формулой (9):

$$D(X) = \frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} x^2\sin(x+y)\,dx\,dy - \left(\frac{\pi}{4}\right)^2 =$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{2}} x^2\left[\cos x - \cos\left(x + \frac{\pi}{2}\right)\right]dx - \frac{\pi^2}{16}.$$

Заменяя $\cos\left(x+\frac{\pi}{2}\right)=-\sin x$, как и в пункте б), и дважды интегрируя по частям, получим:

$$D(X)=\frac{\pi}{2}+\frac{\pi^2}{16}-2\approx 0{,}187.$$

Дисперсия случайной величины Y имеет то же самое значение.

г) Вычислим, наконец, корреляционный момент K_{xy}. Согласно формуле (12),

$$K_{xy}=\frac{1}{2}\int\limits_0^{\frac{\pi}{2}}\int\limits_0^{\frac{\pi}{2}} xy\sin(x+y)\,dx\,dy-\frac{\pi^2}{16}.$$

Те же приемы, что применялись и выше, дают:

$$K_{xy}\approx -0{,}045.$$

Рассмотрим теперь более подробно роль корреляционного момента в характеристике связи между случайными величинами. Эта роль устанавливается прежде всего следующей теоремой.

Теорема 1. *Если случайные величины X и Y независимы, то их корреляционный момент равен нулю:*

$$K_{xy}=0.$$

Доказательство этой теоремы мы проведем для случая непрерывных случайных величин, хотя она справедлива и в общем случае. Как было установлено в § 18, если случайные величины X и Y независимы, то плотность вероятности двумерной случайной величины (X, Y) представляется в виде произведения:

$$\varphi(x,\ y)=\varphi_1(x)\,\varphi_2(y).$$

Воспользовавшись этим равенством, запишем выражение для корреляционного момента в виде:

$$K_{xy}=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(x-\bar{x})(y-\bar{y})\,\varphi(x,y)\,dx\,dy=$$

$$=\int\limits_{-\infty}^{+\infty}(x-\bar{x})\,\varphi_1(x)\,dx\cdot\int\limits_{-\infty}^{+\infty}(y-\bar{y})\,\varphi_2(y)\,dy.$$

Но каждый из интегралов в правой части равенства представляет собою центральный момент первого порядка соответствующей случайной величины и потому равен нулю. Действительно,

$$\int\limits_{-\infty}^{+\infty}(x-\bar{x})\,\varphi_1(x)\,dx=\int\limits_{-\infty}^{+\infty}x\,\varphi_1(x)\,dx-\bar{x}\int\limits_{-\infty}^{+\infty}\varphi_1(x)\,dx=\bar{x}-\bar{x}=0,$$

поскольку первый интеграл выражает математическое ожидание X, а интеграл от плотности вероятности равен единице.

Итак, для независимости случайных величин необходимо, чтобы их корреляционный момент равнялся нулю. Обратная теорема, однако, не верна, и равенство нулю корреляционного момента не является достаточным для независимости случайных величин. В этом можно убедиться на отдельных примерах.

Часто взамен корреляционного момента для характеристики связи между случайными величинами X и Y пользуются безразмерной величиной, равной отношению корреляционного момента к квадратному корню из произведения дисперсий:

$$r_{xy} = \frac{K_{xy}}{\sqrt{D_x D_y}} = \frac{K_{xy}}{\sigma_x \sigma_y}. \tag{13}$$

Это отношение называют *коэффициентом корреляции*. Как будет показано в следующем параграфе, коэффициент корреляции r_{xy} по абсолютной величине не превосходит единицы, т. е. всегда удовлетворяет неравенствам

$$-1 \leqslant r_{xy} \leqslant +1. \tag{14}$$

Из определения коэффициента корреляции (13) и теоремы 1 сразу вытекает следующая теорема.

Теорема 2. *Если случайные величины X и Y независимы, то их коэффициент корреляции равен нулю.*

Кроме нее, имеет место еще одна теорема, более подробно выясняющая роль коэффициента корреляции для характеристики связи между случайными величинами.

Теорема 3. *Если случайные величины X и Y линейно зависимы, т. е. между ними существует соотношение $Y = aX + b$, то абсолютная величина коэффициента корреляции равна единице. Более подробно, $r_{xy} = +1$, если $a > 0$ и $r_{xy} = -1$ при $a < 0$.*

Доказательство. Воспользуемся еще раз свойством математического ожидания, которое будет доказано нами в следующем параграфе: математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых.

Заметим, что корреляционный момент равен:

$$K_{xy} = M\left[\left(X - \bar{x}\right)\left(Y - \bar{y}\right)\right] = M\left[\left(X - \bar{x}\right)\left(aX + b - \bar{y}\right)\right].$$

Далее, в силу указанного свойства математического ожидания,

$$\bar{y} = M(Y) = M(aX + b) = aM(X) + b = a\bar{x} + b;$$

поэтому для корреляционного момента получаем:

$$K_{xy} = M\left[\left(X - \bar{x}\right)\left(aX + b - a\bar{x} - b\right)\right] = M\left[\left(X - \bar{x}\right) a \left(X - \bar{x}\right)\right] =$$

$$= aM\left(X - \bar{x}\right)^2 = aD(X) = a\sigma_x^2.$$

Найдем дисперсию Y:

$$D(Y) = \sigma_y^2 = M\left(Y - \bar{y}\right)^2 = M\left(aX + b - a\bar{x} - b\right)^2 =$$
$$= a^2 M\left(X - \bar{x}\right)^2 = a^2 D(X) = a^2 \sigma_x^2,$$

откуда $\sigma_y = |a|\,\sigma_x$. Таким образом,

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{a\sigma_x^2}{\sigma_x |a| \sigma_x} = \frac{a}{|a|} = \begin{cases} +1 & \text{при } a > 0, \\ -1 & \text{при } a < 0, \end{cases}$$

чем наша теорема и доказана.

Сделаем несколько общих замечаний относительно зависимости между случайными величинами. Понятно, что крайними «полюсами зависимости» является полная независимость случайных величин или, наоборот, функциональная зависимость между ними. Если случайные величины независимы, то знание значений одной из них никак не отражается на сведениях о другой случайной величине. Наоборот, если случайные величины связаны с функциональной зависимостью, то значение, принятое одной из них, однозначно определяет значение, принимаемое другой.

Существует другой вид зависимости между случайными величинами, когда знание значения одной из них не определяет однозначно (т. е. с вероятностью единица) значение другой, а определяет лишь закон распределения другой случайной величины. Такую зависимость принято называть *корреляционной* или *стохастической*. Ее легко проиллюстрировать на примере дискретной двумерной случайной величины. Именно, в § 18 мы задавали закон распределения двумерной дискретной случайной величины в виде таблицы с двумя входами и рассматривали условные законы распределения, определяемые вероятностями $P(y_j/x_i)$ или соответственно $P(x_j/y_i)$. Выбор определенного значения случайной величины X задает определенный закон распределения для случайной величины Y.

Коэффициент корреляции численно характеризует корреляционную зависимость между случайными величинами. Если они независимы, то коэффициент корреляции равен нулю. При значениях r в пределах $0 < |r| < 1$ между случайными величинами имеет место корреляционная зависимость, тем более «тесная», чем ближе величина $|r|$ к единице. Значение $|r| = 1$ соответствует функциональной зависимости между случайными величинами вида $Y = aX + b$.

Нужно только иметь в виду, что значение r показывает отклонение зависимости между случайными величинами от линейной. При другом виде функциональной зависимости между X и Y может случиться, что коэффициент корреляции мал или даже равен нулю.

Возвратимся теперь к общему случаю двумерного нормального распределения, о котором шла речь в конце предыдущего параграфа. Вычислив моменты второго порядка для случайной величины, плотность распределения которой задается формулой (12) из § 19, найдем, что

$$D(X) = \sigma_1^2;\quad D(Y) = \sigma_2^2;\quad K_{xy} = r\sigma_1\sigma_2.$$

Тем самым выясняется теоретико-вероятностный смысл коэффициентов, входящих в общее выражение плотности вероятности нормального распределения: σ_1 и σ_2 суть дисперсии случайных величин X и Y, а r — коэффициент корреляции между ними.

Попутно отметим и роль коэффициента корреляции r: если X и Y независимы, то $r = 0$ и оси эллипса рассеяния будут параллельны координатным осям. Исключение из рассмотрения случая, когда Y есть линейная функция от X, соответствует исключению значения $|r| = 1$.

§ 21. ТЕОРЕМЫ О МАТЕМАТИЧЕСКОМ ОЖИДАНИИ И ДИСПЕРСИИ СЛУЧАЙНЫХ ВЕЛИЧИН

В предыдущих параграфах нам приходилось пользоваться некоторыми свойствами математического ожидания и дисперсии, которые не были еще доказаны. Сейчас мы можем уже приступить к доказательству этих теорем.

Теорема 1. *Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий*:

$$M(X+Y) = M(X) + M(Y). \tag{1}$$

Доказательство. Рассмотрим отдельно случаи, когда X и Y являются соответственно дискретными или непрерывными случайными величинами*.

Начнем со случая дискретных случайных величин. Тогда, по определению математического ожидания, можно написать:

$$M(X+Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_i + y_j)\,p_{ij},$$

где p_{ij} означает вероятность того, что $X = x_i$ и $Y = y_j$. Разбивая сумму на отдельные слагаемые, получим:

$$M(X+Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}x_i + \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}y_j = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} p_{ij}\right)x_i + \sum_{j=1}^{n}\left(\sum_{i=1}^{m} p_{ij}\right)y_j. \tag{2}$$

Как было установлено в § 18, сумма $\sum_{j=1}^{n} p_{ij} = P_i$ выражает ве-

* Как уже указывалось на стр. 75, рассмотрением этих двух типов случайных величин мы и ограничиваемся.

роятность того, что случайная величина X примет значение $X=x_i$ независимо от того, какое значение принимает Y. Точно так же сумма $\sum_{i=1}^{m} p_{ij}=P_j$ есть вероятность равенства $Y=y_j$ независимо от значений X.

Таким образом, первое слагаемое в правой части (2) равно:

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} p_{ij}\right)x_i=\sum_{i=1}^{m} P_i x_i=M(X),$$

а второе:

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{m} p_{ij}\right)y_j=\sum_{j=1}^{n} P_j y_j=M(Y).$$

Обращаясь снова к формуле (2), находим:

$$M(X+Y)=M(X)+M(Y),$$

что и требовалось доказать.

Пусть теперь X и Y — непрерывные случайные величины. По определению,

$$M(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\,\varphi(x,y)\,dx\,dy=$$

$$=\int_{-\infty}^{+\infty} xdx\int_{-\infty}^{+\infty}\varphi(x,y)\,dy+\int_{-\infty}^{+\infty} ydy\int_{-\infty}^{+\infty}\varphi(x,y)\,dx. \qquad (3)$$

Но, как было указано в предыдущем параграфе (см. формулы (1) и (2) из § 20), интегралы

$$\int_{-\infty}^{+\infty}\varphi(x,y)\,dy=\varphi_1(x) \quad \text{и} \quad \int_{-\infty}^{+\infty}\varphi(x,y)\,dx=\varphi_2(y)$$

выражают соответственно плотность вероятности случайных величин X и Y. Поэтому слагаемые правой части равенства (3) равны:

$$\int_{-\infty}^{+\infty} xdx\int_{-\infty}^{+\infty}\varphi(x,y)\,dy=\int_{-\infty}^{+\infty} x\varphi_1(x)\,dx=M(X),$$

$$\int_{-\infty}^{+\infty} ydy\int_{-\infty}^{+\infty}\varphi(x,y)\,dx=\int_{-\infty}^{+\infty} y\varphi_2(y)\,dy=M(Y),$$

откуда и следует равенство (1) для непрерывных случайных величин.

Теорема 2. *Математическое ожидание произведения двух случайных величин равно произведению их математических ожиданий, сложенному с их корреляционным моментом.*

Доказательство. По определению корреляционного момента, имеем:

$$K_{xy} = M[(X-\bar{x})(Y-\bar{y})],$$

где $\bar{x} = M(X)$ и $\bar{y} = M(Y)$. Раскрывая скобки и пользуясь уже доказанной теоремой 1, находим *:

$$K_{xy} = M(XY) - \bar{x}M(Y) - \bar{y}M(X) + \bar{x}\bar{y} = M(XY) - \bar{x}\bar{y},$$

откуда следует:

$$M(XY) = M(X)M(Y) + K_{xy}. \qquad (4)$$

Так как для независимых случайных величин корреляционный момент равен нулю, то из формулы (4) непосредственно следует

Теорема 3. *Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:*

$$M(XY) = M(X)M(Y). \qquad (5)$$

С помощью доказанных теорем мы можем теперь установить указанные в предыдущем параграфе неравенства для коэффициента корреляции.

Теорема 4. *Коэффициент корреляции двух случайных величин по абсолютной величине не превосходит единицы:*

$$-1 \leqslant r_{xy} \leqslant +1. \qquad (6)$$

Доказательство. Пусть даны случайные величины X и Y с математическими ожиданиями соответственно $\bar{x}$ и $\bar{y}$ и дисперсиями σ_x^2 и σ_y^2. Образуем новую случайную величину Z, которая определяется через X и Y равенством

$$Z = \sigma_y X - \sigma_x Y. \qquad (7)$$

Найдем математическое ожидание и дисперсию Z. В силу свойств математического ожидания,

$$\bar{z} = M(Z) = M(\sigma_y X - \sigma_x Y) = \sigma_y \bar{x} - \sigma_x \bar{y}.$$

Далее,

$$D(Z) = M(Z-\bar{z})^2 = M[(\sigma_y X - \sigma_x Y) - (\sigma_y \bar{x} - \sigma_x \bar{y})]^2.$$

Возведя выражение под знаком математического ожидания в квадрат и воспользовавшись снова свойствами математического ожидания, получим:

$$D(Z) = \sigma_z^2 = \sigma_y^2 M(X-\bar{x})^2 - 2\sigma_x\sigma_y M[(X-\bar{x})(Y-\bar{y})] +$$
$$+ \sigma_x^2 M(Y-\bar{y})^2.$$

* Возможность вынесения постоянного множителя за знак математического ожидания была доказана еще в § 16.

Исходя из определений дисперсии и корреляционного момента, мы можем переписать полученное равенство в виде

$$\sigma_z^2 = \sigma_y^2 \sigma_x^2 - 2\sigma_x\sigma_y K_{xy} + \sigma_x^2 \sigma_y^2,$$

или окончательно:

$$\sigma_z^2 = 2\sigma_x^2 \sigma_y^2 - 2\sigma_x \sigma_y K_{xy}.$$

Поскольку дисперсия любой случайной величины есть величина неотрицательная, из последнего выражения получаем:

$$\sigma_x\sigma_y - K_{xy} \geqslant 0,$$

откуда

$$K_{xy} \leqslant \sigma_x\sigma_y. \tag{8}$$

Вводя вместо (7) новую случайную величину

$$Z = \sigma_y X + \sigma_x Y$$

и произведя с ней те же действия, получим вместо (8) неравенство:

$$K_{xy} \geqslant -\sigma_x\sigma_y. \tag{9}$$

Комбинация неравенств (8) и (9) дает:

$$-\sigma_x\sigma_y \leqslant K_{xy} \leqslant \sigma_x\sigma_y,$$

или, что то же самое,

$$|K_{xy}| \leqslant \sigma_x\sigma_y. \tag{10}$$

Остается разделить обе части неравенства (10) на положительное число $\sigma_x\sigma_y$ и воспользоваться определением коэффициента корреляции:

$$|r_{xy}| = \frac{|K_{xy}|}{\sigma_x \sigma_y} \leqslant 1,$$

что и утверждалось.

Рассмотрим теперь теорему о дисперсии суммы случайных величин.

Теорема 5. *Дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых*:

$$D(X+Y) = D(X) + D(Y). \tag{11}$$

Доказательство. Воспользуемся формулой (15) из § 16:

$$D(X+Y) = M(X+Y)^2 - (M(X+Y))^2.$$

Раскрыв скобки в правой части равенства, получим:

$$D(X+Y) = M(X^2 + 2XY + Y^2) - [M(X) + M(Y)]^2.$$

Теперь мы должны будем воспользоваться теоремой 1 о математическом ожидании суммы и теоремой 3 о математическом ожидании произведения независимых случайных величин. Получаем:

$$D(X+Y) = M(X^2) + 2M(X)M(Y) + M(Y^2) - (M(X))^2 - 2M(X)M(Y) - (M(Y))^2 = M(X^2) - (M(X))^2 + M(Y^2) - (M(Y))^2.$$

Таким образом,

$$D(X+Y) = D(X) + D(Y).$$

Обращаем внимание читателя на различие в формулировках теорем 1 и 5. В то время как в теореме 1 речь идет о сумме произвольных случайных величин, теорема 5 имеет место лишь в предположении, что случайные слагаемые независимы. В процессе доказательства было указано, где это обстоятельство используется.

§ 22. МНОГОМЕРНАЯ СЛУЧАЙНАЯ ВЕЛИЧИНА И СИСТЕМА СЛУЧАЙНЫХ ВЕЛИЧИН. СУММИРОВАНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН. КОМПОЗИЦИЯ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ

Изучение двумерных случайных величин, которым мы занимались в § 18—20, дает уже достаточно полное представление о многомерных случайных величинах. Поэтому мы можем ограничиться только тем, что приведем некоторые общие определения и примеры.

Под *многомерной случайной величиной* мы будем понимать случайную величину, значения которой представляют собою *точки n-мерного пространства*, или, что то же самое, *n-мерный вектор*. Многомерную случайную величину часто называют также *случайным вектором*. Функция n переменных

$$F(x_1, x_2, \ldots, x_n) = P(X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n) \tag{1}$$

называется *n-мерной функцией распределения случайного вектора*. Нетрудно по аналогии с введенным в § 18 понятием двумерной плотности вероятности ввести также *n-мерную плотность вероятности распределения* случайного вектора.

Координаты $(X_1, X_2, \ldots, X_n)$ точки, служащей значением n-мерной случайной величины, сами являются случайными величинами. Поэтому n-мерную случайную величину—как и двумерную—можно рассматривать как систему n случайных величин.

Если, например, все эти случайные величины независимы и каждая из них распределена по нормальному закону с математическим ожиданием a_k и дисперсией σ_k^2, то n-мерная плотность вероятности нормального закона определится формулой

$$\varphi(x_1, x_2, \ldots, x_n) = \frac{(2\pi)^{-\frac{n}{2}}}{\sigma_1\sigma_2\ldots\sigma_n} e^{-\frac{1}{2}\sum_{k=1}^{n}\left(\frac{x_k-a_k}{\sigma_k}\right)^2}$$

Точка с координатами $(a_1, a_2, \ldots, a_n)$ называется *центром рассеяния* многомерной случайной величины.

На многомерную случайную величину без всякого изменения переносятся понятия математического ожидания и дисперсии, а также центральных моментов, рассмотренные в § 20.

В дальнейшем нас будут интересовать суммы случайных величин, главным образом независимых. Теоремы 1 и 5 из § 21 позволяют находить математическое ожидание и дисперсию суммы случайных величин по соответствующим числовым характеристикам отдельных слагаемых.

Однако в ряде случаев требуется большее: нужно определять законы распределения суммы случайных величин по законам распределения слагаемых. Такую задачу называют *композицией законов распределения.*

Рассмотрим композицию законов распределения для случаев дискретных и непрерывных случайных величин. Пусть X и Y — две независимые дискретные случайные величины, имеющие множества возможных значений соответственно $\{x_i\}$ и $\{y_j\}$. Их сумма $Z = X + Y$ есть случайная величина, множество значений которой есть $\{x_i + y_j\}$ $(i, j = 1, 2, \ldots)$. Обозначим это множество через $\{z_k\}$ и определим вероятность равенства $Z = z_k$.

Допустим, что случайная величина X приняла значение $X = x_i$. Тогда для равенства $Z = z_k$ необходимо, чтобы случайная величина Y приняла значение $Y = Z_k - x_i$. Если эта разность не входит в число возможных значений Y, то при $X = x_i$ равенство $Z = z_k$ невозможно, т. е. имеет вероятность $P(Z = z_k) = 0$.

Если же равенство $Y = z_k - x_i$ возможно, то, очевидно, вероятность равенства $Z = z_k$ при условии $X = x_i$ может быть найдена как вероятность совмещения событий и равна произведению

$$P(X = x_i)\, P(Y = z_k - x_i), \tag{3}$$

так как значения слагаемых X и Y независимы.

Так как различные значения случайной величины X несовместны, то вероятность равенства $Z = z_k$ может быть получена по теореме сложения, путем суммирования произведений вида (3). Таким образом,

$$P(Z = z_k) = \sum_i P(X = x_i)\, P(Y = z_k - x_i), \tag{4}$$

где сумма распространяется на все возможные при данном z_k значения x_i. Аналогично можно показать, что справедливо равенство:

$$P(Z = z_k) = \sum_j P(Y = y_j)\, P(X = z_k - y_j). \tag{5}$$

Формулы (4) и (5) выражают закон распределения суммы случайных величин через законы распределения отдельных слагаемых, т. е. выражают композицию дискретных законов распределения.

Перейдем теперь к случаю, когда X и Y являются непрерывными случайными величинами соответственно с плотностями вероятности $\varphi_1(x)$ и $\varphi_2(y)$. Найдем плотность вероятности суммы $Z = X + Y$.

Обозначим через $\varphi(x, y)$ двумерную плотность распределения системы двух случайных величин (X, Y). Тогда вероятность неравенства $Z < z = x + y$ можно записать с помощью двумерной функции распределения в виде

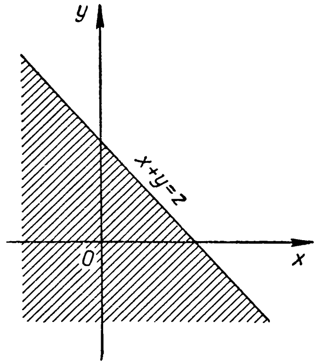


Рис. 30

$$P(Z < x + y) = F(z) = \iint_G \varphi(x,y)\,dx\,dy,$$

где G — область плоскости (x, y), ограниченная линией $x + y = z$ (см. рис. 30). Вычисляя этот двойной интеграл, находим:

$$F(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} \varphi(x, y)\,dy.$$

Случайные величины X и Y были предположены независимыми. Поэтому двумерная плотность вероятности $\varphi(x, y)$ равна произведению:

$$\varphi(x, y) = \varphi_1(x)\,\varphi_2(y).$$

Подставим это значение в выражение для $F(z)$. Тогда

$$F(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} \varphi_1(x)\,\varphi_2(y)\,dy = \int_{-\infty}^{+\infty} \varphi_1(x)\,dx \int_{-\infty}^{z-x} \varphi_2(y)\,dy.$$

Заменим во внутреннем интеграле переменную интегрирования, положив $y = u - x$. Тогда $dy = du$, и при $y = z - x$ будем иметь $u = z$, т. е.

$$\int_{-\infty}^{z-x} \varphi_2(y)\,dy = \int_{-\infty}^{z} \varphi_2(u - x)\,du.$$

Следовательно,

$$F(z) = \int_{-\infty}^{+\infty} \varphi_1(x)\,dx \int_{-\infty}^{z} \varphi_2(u - x)\,du = \int_{-\infty}^{z} du \int_{-\infty}^{+\infty} \varphi_1(x)\,\varphi_2(u - x)\,dx.$$

Последнее выражение получено изменением последовательности интегрирования.

Итак, окончательно для функции распределения получилось выражение

$$F(z) = \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{+\infty} \varphi_1(x)\, \varphi_2(u - x)\, dx \right) du. \tag{6}$$

Если исходную плотность вероятности случайной величины Z обозначить через $\varphi(z)$, то выражение для функции распределения должно, как известно, иметь вид:

$$\varphi(z) = F'(z).$$

Дифференцируя по z формулу (6), находим:

$$\varphi(z) = \int\limits_{-\infty}^{+\infty} \varphi_1(x)\, \varphi_2(z - x)\, dx. \tag{7}$$

Формула (7) для плотности вероятности напоминает формулу (4) для дискретных случайных величин. Тем же путем можно получить формулу, аналогичную формуле (5). Она будет иметь вид:

$$\varphi(z) = \int\limits_{-\infty}^{+\infty} \varphi_1(z - y)\, \varphi_2(y)\, dy. \tag{8}$$

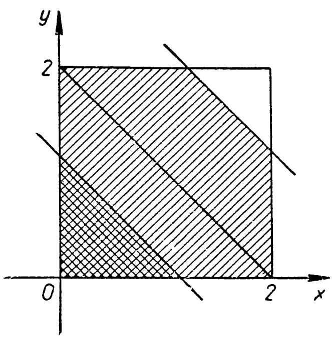


Рис. 31

Функция $\varphi(z)$, полученная из функций $\varphi_1(x)$ и $\varphi_2(y)$ с помощью формул (7) или (8), называется *сверткой* этих функций соответственно по переменной x или y.

Пример. Случайные величины X и Y независимы и равномерно распределены в интервале (0, 2). Найдем плотность распределения суммы $Z = X + Y$.

Для лучшего уяснения сути дела воспользуемся геометрическими соображениями. Система случайных величин (X, Y) распределена в квадрате $0 \leqslant x \leqslant 2$, $0 \leqslant y \leqslant 2$ (см. рис. 31). Поэтому возможные значения случайной величины Z заключены в пределах $0 \leqslant z \leqslant 4$. При $0 < z \leqslant 2$ вероятность неравенства $Z < z$ равна вероятности попадания в левый нижний дважды заштрихованный треугольник. Так как площадь этого треугольника равна $\frac{z^2}{2}$, а площадь всего квадрата равна 4, то эта вероятность есть $\frac{z^2}{8}$.

Если $2 < z \leqslant 4$, то вероятность неравенства $Z < z$ есть вероятность попадания во всю заштрихованную фигуру. Эта вероятность равна

$$1 - \frac{(4-z)^2}{8}.$$

Учитывая, что значения $Z < 0$ и $Z > 4$ невозможны, запишем функцию распределения суммы $Z = X + Y$ следующим образом:

$$F(z) = \begin{cases} 0 & \text{при} \quad -\infty < z \leqslant 0, \\ \frac{z^2}{8} & \text{при} \quad 0 < z \leqslant 2, \\ 1 - \frac{(4-z)^2}{8} & \text{при} \quad 2 < z \leqslant 4, \\ 1 & \text{при} \quad 4 < z < +\infty. \end{cases} \tag{9}$$

Дифференцируя функцию (9) по z, находим для плотности вероятности:

$$\varphi(z) = \begin{cases} 0 & \text{при} \quad -\infty < z < 0, \\ \frac{z}{4} & \text{при} \quad 0 < z < 2, \\ \frac{4-z}{4} & \text{при} \quad 2 < z < 4, \\ 0 & \text{при} \quad 4 < z < +\infty. \end{cases} \tag{10}$$

Закон распределения (10) есть частный случай *закона распределения Симпсона*. Его выражение можно было получить и формально по формулам свертки (7) или (8). Действительно, для случайных величин X и Y имеем плотности вероятности:

$$\varphi_1(x) = \varphi_2(y) = \begin{cases} 0 & \text{при} \ -\infty < x, \quad y < 0, \\ \frac{1}{2} & \text{при} \quad 0 < x, \quad y < 2, \\ 0 & \text{при} \quad 2 < x, \quad y < +\infty. \end{cases}$$

По формуле (7) получаем:

$$\varphi(z) = \int_{-\infty}^{+\infty} \varphi_1(x)\, \varphi_2(z-x)\, dx =$$

$$= \int_0^2 \varphi_1(x)\, \varphi_2(z-x)\, dx = \frac{1}{2} \int_0^2 \varphi_2(z-x)\, dx,$$

При $z < 0$ находим $z - x < 0$ и, значит, $\varphi_2(z - x) = 0$, так что $\varphi(z) = 0$. Далее, при $0 < z < 2$ функция $\varphi(z - x)$ отлична от нуля только при тех x, которые удовлетворяют неравенству $0 < z - x < 2$, или, что то же самое,

$$z - 2 < x < z.$$

Отсюда следует, что если $0 < z < 2$, то

$$\varphi(z) = \frac{1}{2}\int_0^z \frac{1}{2}\,dx = \frac{z}{4}.$$

Если же $2 < z < 4$, то

$$\varphi(z) = \frac{1}{2}\int_{z-2}^{2} \frac{1}{2}\,dx = \frac{1}{4}(2 - (z - 2)) = \frac{4 - z}{4}.$$

Наконец, при $z > 4$ имеем $z - x > 2$, откуда следует $\varphi_2(z - x) = 0$ и $\varphi(z) = 0$. Таким образом, мы снова получили формулы (10).

Для различных приложений теории вероятностей очень важен факт, выражаемый следующей теоремой.

Теорема. *Если две случайные величины X и Y распределены по нормальному закону, то их сумма $Z = X + Y$ также распределена по нормальному закону.*

Доказательство этой теоремы можно получить, вычисляя закон распределения суммы с помощью формулы свертки (7) или (8). Подробнее мы на этом останавливаться не будем.

§ 23. ЗАКОН БОЛЬШИХ ЧИСЕЛ И ЕГО ОБОБЩЕНИЯ. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА А. М. ЛЯПУНОВА

С термином *закон больших чисел* мы уже встречались в § 12 при формулировке теоремы Бернулли. В ней речь шла о том, что при достаточно большом числе испытаний частота появлений некоторого события сколь угодно мало отличается от его вероятности. Вообще законом больших чисел, в достаточно широком смысле этого слова, естественно называть теоремы, где устанавливается, что с вероятностью, близкой к единице, наступит некоторое событие, зависящее от неограниченно увеличивающегося числа событий, каждое из которых в отдельности играет незначительную роль.

В основе доказательств теорем такого рода лежит очень важное неравенство, установленное П. Л. Чебышевым в 1845—1846 годах.

Неравенство Чебышева. *Пусть случайная величина X имеет конечное математическое ожидание и конечную дисперсию. Тогда для любого $\varepsilon > 0$ вероятность отклонения X от сво-*

его математического ожидания меньше чем на ε *отличается от* 1 *на величину, не превосходящую* $\frac{D(x)}{\varepsilon^2}$:

$$P\{|X - M(X)| < \varepsilon\} \geqslant 1 - \frac{D(X)}{\varepsilon^2}. \qquad (1)$$

Доказательство. Неравенство (1) является следствием другого неравенства, также принадлежащего Чебышеву: *если случайная величина* X *может принимать только неотрицательные значения и имеет конечное математическое ожидание, то вероятность того, что принятое ею значение окажется не меньше единицы, не превосходит математического ожидания этой величины*:

$$P(X \geqslant 1) \leqslant M(X). \qquad (2)$$

В самом деле, пусть X — дискретная случайная величина. Тогда вероятность $P(X \geqslant 1)$ есть сумма вероятностей, с которыми X принимает отдельные значения, большие единицы, т. е.

$$P(X \geqslant 1) = \sum_{x_i \geqslant 1} P(X = x_i).$$

Если каждое слагаемое справа умножить на соответствующее значение x_i, то правая часть возрастет, так как значения $x_i \geqslant 1$. Мы придем к неравенству:

$$\sum_{x_i \geqslant 1} P(X = x_i) \leqslant \sum_{x_i \geqslant 1} x_i P(X = x_i).$$

Это неравенство лишь усилится, если распространить стоящую справа сумму на все возможные значения X, которые по предположению неотрицательны. Следовательно,

$$P(X \geqslant 1) \leqslant \sum_i x_i P(X = x_i) = M(X),$$

поскольку последняя сумма, по определению, совпадает с математическим ожиданием X.

Таким образом мы приходим к цепочке равенств и неравенств:

$$P(X \geqslant 1) = \sum_{x_i \geqslant 1} P(X = x_i) \leqslant \sum_{x_i \geqslant 1} x_i P(X = x_i) \leqslant$$
$$\leqslant \sum_i x_i P(X = x_i) = M(X), \qquad (3)$$

которая и доказывает неравенство (2) для случая дискретной случайной величины.

Пусть теперь X — непрерывная случайная величина с плотностью вероятности $\varphi(x)$. В силу предположения о неотрицательности значений X для $-\infty < x < 0$ справедливо тождество $\varphi(x) \equiv 0$. Как и для дискретной случайной величины, мы можем написать цепочку равенств и неравенств, вполне аналогичную (3), в которой лишь суммы заменяются интегралами:

$$P(X \geqslant 1) = \int_1^{+\infty} \varphi(x)\,dx \leqslant \int_1^{+\infty} x\varphi(x)\,dx \leqslant \int_0^{+\infty} x\varphi(x)\,dx = $$
$$= \int_{-\infty}^{+\infty} x\varphi(x)\,dx = M(X). \tag{4}$$

Формулы (4) доказывают неравенство (2) и для случая непрерывной случайной величины.

Мы можем теперь возвратиться к рассмотрению неравенства (1). Взамен написанного там неравенства $|X - M(X)| < \varepsilon$ возьмем противоположное неравенство $|X - M(X)| \geqslant \varepsilon$, которое, как легко видеть, равносильно следующему:

$$\frac{[X - M(X)]^2}{\varepsilon^2} \geqslant 1.$$

Случайная величина $\frac{[X - M(X)]^2}{\varepsilon^2}$ принимает лишь неотрицательные значения, и к ней можно поэтому применить неравенство (2). Тогда находим:

$$P\{|X - M(X)| \geqslant \varepsilon\} = P\left\{\frac{[X - M(X)]^2}{\varepsilon^2} \geqslant 1\right\} \leqslant$$
$$\leqslant M\left\{\frac{[X - M(X)]^2}{\varepsilon^2}\right\} = \frac{D(X)}{\varepsilon^2}.$$

Последнее равенство получено вынесением за знак математического ожидания постоянного множителя $\frac{1}{\varepsilon^2}$ и вследствие определения дисперсии.

Итак, мы получили:

$$P\{|X - M(X)| \leqslant \varepsilon\} \leqslant \frac{D(X)}{\varepsilon^2},$$

откуда, возвращаясь к первоначальному неравенству, получаем:

$$P\{|X - M(X) < \varepsilon\} \leqslant 1 - \frac{D(X)}{\varepsilon^2},$$

что и требовалось доказать.

Неравенство Чебышева (1) позволяет легко доказать теорему Бернулли (см. стр. 73); для этого достаточно применить это неравенство к случайной величине $\frac{m}{n}$, где m — число наступлений некоторого события при n испытаниях. Действительно, как было установлено в § 16, если вероятность наступления события при отдельном испытании равна p, то $M(m) = np$, а $D(m) = npq$. Но тогда $M\left(\frac{m}{n}\right) = p$ и $D\left(\frac{m}{n}\right) = \frac{npq}{n^2} = \frac{pq}{n}$. Поэтому

$$P\left\{\left|\frac{m}{n} - p\right| < \varepsilon\right\} = P\left\{\left|\frac{m}{n} - M\left(\frac{m}{n}\right)\right| < \varepsilon\right\} \geqslant 1 - \frac{D\left(\frac{m}{n}\right)}{\varepsilon^2} = 1 - \frac{pq}{n\varepsilon^2},$$

откуда и следует, что при $n \to \infty$ вероятность неравенства $\left|\frac{m}{n} - p\right| < \varepsilon$ стремится к единице.

Как уже было указано, теорема Бернулли является частным случаем более общих теорем. Основной формой закона больших чисел обычно принято считать теорему Чебышева, которую собственно и называют законом больших чисел в узком смысле этого слова.

Теорема Чебышева (*Закон больших чисел в форме Чебышева*). *Пусть $X_1, X_2, \ldots, X_n, \ldots$ — последовательность попарно независимых случайных величин с одним и тем же математическим ожиданием a и дисперсиями, ограниченными одной и той же постоянной*:

$$M(X_1) = \ldots = M(X_n) = \ldots = a,$$

$$D(X_1) < C,\ D(X_2) < C,\ \ldots,\ D(X_n) < C,\ \ldots.$$

Тогда для любого положительного ε справедливо соотношение

$$\lim_{n \to \infty} P\left\{\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - a\right| < \varepsilon\right\} = 1. \tag{5}$$

Иначе говоря, если все случайные величины с ограниченными дисперсиями имеют одно и то же математическое ожидание, то при достаточно большом n с вероятностью, близкой к единице, среднее арифметическое этих случайных величин как угодно мало отличается от их математического ожидания.

Смысл теоремы Чебышева можно пояснить следующим примером. Пусть требуется измерить некоторую физическую постоянную a. Результат каждого отдельного измерения есть случайная величина X_i. Предположение $M(X_i) = a$ означает, что математическое ожидание результата измерения равно измеряемой величине, т. е. что измерения свободны от систематических ошибок. Условие $D(X_i) < C$ означает, что выполняемые измерения имеют некоторую гарантированную минимальную точность, в результате чего случайные рассеяния результатов измерений не могут возрастать неограниченно.

В этих условиях теорема Чебышева утверждает, что при достаточно большом числе измерений с вероятностью, очень близкой к единице, среднее арифметическое полученных результатов будет как угодно мало отличаться от измеряемой величины. Тем самым оправдывается обычно рекомендуемый в физике способ получения более точных результатов измерений: одна и та же физическая величина измеряется многократно и в качестве ее значения берется среднее арифметическое полученных результатов измерений.

Доказательство теоремы Чебышева. Рассмотрим случайную величину $\frac{X_1+X_2+\ldots+X_n}{n}$. В силу свойств математического ожидания, имеем:

$$M\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)=\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}=a.$$

Применяя к этой случайной величине неравенство (1), получаем:

$$P\left\{\left|\frac{X_1+X_2+\ldots+X_n}{n}-a\right|<\varepsilon\right\}=P\left\{\left|\frac{X_1+X_2+\ldots+X_n}{n}-\right.\right.$$

$$\left.\left.-M\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)\right|<\varepsilon\right\}\geqslant 1-\frac{D\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)}{\varepsilon^2}=$$

$$=1-\frac{D(X_1)+D(X_2)+\ldots+D(X_n)}{n^2\varepsilon^2}\geqslant 1-\frac{C}{n\varepsilon^2}.$$

Полученное неравенство доказывает теорему, так как из него видно, что при неограниченно возрастающем n вероятность требуемого неравенства становится как угодно близкой к единице.

Легко видеть, что теорема Бернулли есть частный случай теоремы Чебышева. Достаточно рассмотреть случайные величины m_i, принимающие значения 1 или 0 соответственно, в зависимости от того, наступило или не наступило рассматриваемое событие в i-м испытании. Тогда при n испытаниях число m наступления события равно $m=m_1+m_2+\ldots+m_n$. Как было показано в § 16, $M(m_i)=p$, $D(m_i)=pq$ и

$$\left|\frac{m_1+m_2+\ldots+m_n}{n}-p\right|=\left|\frac{m}{n}-p\right|,$$

так что неравенство, рассматриваемое в теореме Чебышева, переходит в неравенство теоремы Бернулли.

С другой стороны, нетрудно получить и более общую теорему, отказавшись от того, чтобы все случайные величины имели одинаковые математические ожидания. Более того, можно даже отказаться и от ограниченности дисперсий, потребовав лишь, чтобы они возрастали не слишком быстро. Именно, справедлива следующая более общая теорема.

Т е о р е м а М а р к о в а. *Пусть* $X_1, X_2, \ldots, X_n, \ldots$ — *попарно независимые случайные величины с конечными математическими ожиданиями* $M(X_i) = a_i$ *и конечными дисперсиями* $D(X_i) = d_i$, *которые удовлетворяют условию*

$$\lim_{n \to \infty} \frac{d_1 + d_2 + \ldots + d_n}{n^2} = 0. \qquad (6)$$

Тогда для любого положительного ε *справедливо соотношение*

$$\lim_{n \to \infty} P\left\{\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{a_1 + a_2 + \ldots + a_n}{n}\right| < \varepsilon\right\} = 1. \qquad (7)$$

Д о к а з а т е л ь с т в о теоремы Маркова, как и предыдущей теоремы Чебышева, сводится к применению неравенства (1) к случайной величине $\frac{X_1 + X_2 + \ldots + X_n}{n}$, после чего остается только воспользоваться условием (6), которому удовлетворяют дисперсии d_i. Рекомендуем читателю самостоятельно проделать все требуемые выкладки, а также убедиться, что теорема Чебышева получается из теоремы Маркова как частный случай.

Во всех трех формулировках закона больших чисел утверждалось, что вероятности соответствующих неравенств стремятся при неограниченном росте n к единице. Однако это не исключает возможности того, что, хотя бы и с очень малой вероятностью, при длинной серии испытаний частота $\frac{m}{n}$ будет заметно отклоняться от вероятности.

Иными словами, из теоремы Бернулли, например, н е л ь з я с д е л а т ь в ы в о д а, что частота $\frac{m}{n}$ наступления события при неограниченном возрастании n с т р е м и т с я к в е р о я т н о с т и p. Такое утверждение было бы просто н е в е р н ы м. Однако, оказывается, можно сделать утверждение, более сильное, чем теорема Бернулли.

Это усиление теоремы Бернулли, как и аналогичные усиления других форм закона больших чисел — теорем Чебышева и Маркова, — получили название *усиленного закона больших чисел*. Приведем одну из таких теорем.

Т е о р е м а Б о р е л я. *Пусть* m — *число наступлений события* A *при* n *испытаниях, причем вероятность наступления* A *при отдельном испытании постоянна и равна* p. *Тогда вероятность того, что частота* $\frac{m}{n}$ *стремится к* p, *равна единице, т. е.*

$$P\left\{\lim \frac{m}{n} = p\right\} = 1. \qquad (8)$$

Доказательства теоремы Бореля мы приводить не будем. Укажем лишь, что его, как и ряд других форм усиленного закона больших чисел (теоремы Колмогорова, Хинчина, Кантелли и другие), можно получить, используя одно замечательное обобщение неравенства Чебышева (1).

Н е р а в е н с т в о К о л м о г о р о в а. *Если взаимно независимые случайные величины* $X_1, X_2, \ldots, X_n, \ldots$ *имеют конечные дисперсии, то вероятность совместного осуществления неравенств*

$$\left|\sum_{i=1}^{k} X_i - M(X_i)\right| < \varepsilon \quad (k = 1, 2, \ldots, n) \qquad (9)$$

н е м е н ь ш е ч и с л а $\quad 1 - \frac{1}{\varepsilon^2} \sum_{i=1}^{n} D(X_i).$

До сих пор речь шла лишь о числовых характеристиках сумм случайных величин. Известны также теоремы, позволяющие при соответствующих условиях установить законы распределения для сумм случайных величин. Такие теоремы называют *предельными теоремами теории вероятностей*.

Собственно говоря, мы уже имели дело с теоремами такого рода: асимптотические теоремы, рассмотренные нами в главе II, могут быть истолкованы как теоремы о предельных законах распределения.

Обратимся, например, к схеме Бернулли повторения независимых испытаний (см. § 4) и рассмотрим случайные величины m_i, означающие наступление или ненаступление события A при i-м испытании. Их мы вводили уже в § 15 (см. стр. 103), где было указано, что m_i принимает значение 1 с вероятностью p и значение 0 с вероятностью $q = 1 - p$. Число m наступлений события A при n испытаниях есть сумма независимых случайных величин: $m = m_1 + m_2 + \ldots + m_n$.

В этих условиях интегральная предельная теорема Муавра — Лапласа, доказанная в § 12, означает, что при неограниченном возрастании числа испытаний закон распределения случайной величины m стремится к нормальному закону*. Аналогичную роль играет и теорема Пуассона (§ 11), показывающая, что при других условиях предельным законом распределения оказывается закон Пуассона.

Чаще всего в качестве предельных законов распределения фигурирует некоторое небольшое число уже рассмотренных и хорошо известных законов. Наиболее часто таким предельным законом, к которому стремится закон распределения суммы случайных величин, оказывается *нормальный закон распределения*.

Наиболее важную роль среди предельных теорем теории вероятностей играет *центральная предельная теорема А. М. Ляпунова*. Эта теорема устанавливает достаточные условия того, что закон распределения суммы случайных величин при неограниченном возрастании числа слагаемых стремится к нормальному закону распределения. Приведем формулировку этой теоремы.

Центральная предельная теорема А.М. Ляпунова. *Пусть $X_1, X_2, \ldots, X_n$ — независимые случайные величины с конечными математическими ожиданиями $M(X_i) = a_i$ и дисперсиями $D(X_i) = d_i$. Если существуют конечные моменты*

---

* На самом деле в теореме Муавра — Лапласа рассматривался закон распределения нормированного отклонения $\frac{m - np}{\sqrt{npq}}$. Однако нетрудно доказать, что из нормальности распределения одной из этих величин вытекает нормальность распределения другой.

третьего порядка $M|X_i - a_i|^3 = k_i$ *и при неограниченном возрастании* n *выполняется условие*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} k_i}{\left(\sum_{i=1}^{n} d_i\right)^{3/2}} = 0, \qquad (11)$$

то справедливо соотношение

$$\lim_{n\to\infty} \left\{ \frac{\sum_{i=1}^{n} X_i - a}{\left(\sum_{i=1}^{n} d_i\right)^{1/2}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt. \qquad (12)$$

Условие (11) называют *условием Ляпунова**.

Для доказательства этой теоремы А. М. Ляпунов разработал новый метод, носящий название *метода характеристических функций*. В настоящее время метод характеристических функций играет очень важную роль не только в теории вероятностей, но и в ряде других областей математики, однако его рассмотрение потребовало бы чересчур много места, времени и усилий. Поэтому мы ограничимся только тем, что уже сказано.

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ К ГЛАВЕ IV

1. Что такое двумерная случайная величина?
2. Опишите закон распределения дискретной двумерной случайной величины.
3. Что такое условное распределение вероятностей?
4. Как определяется функция распределения и плотность вероятности двумерной случайной величины?
5. В каком случае двумерная случайная величина называется распределенной по нормальному закону?
6. Что называется центром рассеяния двумерной случайной величины?
7. Как определяются математическое ожидание и дисперсия системы двух случайных величин?
8. Каков вероятностный смысл центральных моментов первого и второго порядков системы двух случайных величин?

---

* Мы привели условие Ляпунова в несколько более узкой форме для упрощения формулировки. В дальнейшем это условие было еще ослаблено Линдебергом, после чего Феллер доказал, что это новое условие является не только достаточным, но и необходимым.

9. Что характеризует корреляционный момент системы двух случайных величин?

10. Что такое коэффициент корреляции? Каковы границы его изменения?

11. Чему равен коэффициент корреляции для двух независимых случайных величин? Чему равен коэффициент корреляции в том случае, когда случайные величины связаны линейной зависимостью?

12. Можно ли из равенства нулю коэффициента корреляции заключить, что случайные величины независимы?

13. Одинаковы ли условия в теоремах о математическом ожидании и о дисперсии суммы случайных величин ?

14. Что такое многомерная случайная величина?

15. Как по законам распределения двух дискретных случайных величин найти закон распределения их суммы?

16. Приведите пример композиции законов распределения для случая двух непрерывных случайных величин.

17. Что называют законом больших чисел? Какой смысл имеет это название?

18. В чем состоит закон больших чисел в форме Чебышева?

19. Какова роль предельных теорем теории вероятностей?

20. Какой из законов распределения чаще всего фигурирует в качестве предельного закона?

21. Как можно истолковать теоремы Муавра — Лапласа в качестве предельных теорем теории вероятностей?

22. В чем состоит центральная предельная теорема Ляпунова?

ТАБЛИЦЫ ПЛОТНОСТИ ВЕРОЯТНОСТИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ И ФУНКЦИИ ЛАПЛАСА

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \text{ и } \Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{1}{2}t^2}dt.$$

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
0,00	0,3989	0,0000	0,40	0,3683	0,1554	0,80	0,2897	0,2881
01	3989	0040	41	3668	1591	81	2874	2910
02	3989	0080	42	3653	1628	82	2850	2939
03	3988	0120	43	3637	1664	83	2827	2967
04	3986	0160	44	3621	1700	84	2803	2995
05	3984	0199	45	3605	1736	85	2780	3023
06	3982	0239	46	3589	1772	82	2756	3051
07	3980	0279	47	3572	1808	87	2732	3078
08	3977	0319	48	3555	1844	88	2709	3106
09	3973	0359	49	3538	1879	89	2685	3133
0,10	0,3970	0,0398	0,50	0,3521	0,1915	0,90	0,2661	0,3159
11	3965	0438	51	3503	1950	91	2637	3186
12	3961	0478	52	3485	1985	92	2613	3212
13	3956	0517	53	3467	2019	93	2589	3238
14	3951	0557	54	3448	2054	94	2565	3264
15	3945	0596	55	3429	2088	95	2541	3289
16	3939	0636	56	3410	2123	96	2516	3315
17	3932	0675	57	3391	2157	97	2492	3340
18	3925	0714	58	3372	2190	98	2468	3365
19	3918	0753	59	3352	2224	99	2444	3389
0,20	0,3910	0,0793	0,60	0,3332	0,2257	1,00	0,2420	0,3413
21	3902	0832	61	3312	2291	01	2396	3438
22	3894	0871	62	3292	2324	02	2371	3461
23	3885	0910	63	3271	2357	03	2347	3485
24	3876	0948	64	3251	2389	04	2323	3508
25	3867	0987	65	3230	2422	05	2299	3531
26	3857	1026	66	3209	2454	06	2275	3554
27	3847	1064	67	3187	2486	07	2251	3577
28	3836	1103	68	3166	2517	08	2227	3599
29	3825	1141	69	3144	2549	09	2203	3621
0,30	0,3814	0,1179	0,70	0,3123	0,2580	1,10	0,2179	0,3643
31	3802	1217	71	3101	2611	11	2155	3665
32	3790	1255	72	3079	2642	12	2131	3686
33	3778	1293	73	3056	2673	13	2107	3708
34	3765	1331	74	3034	2703	14	2083	3729
35	3752	1368	75	3011	2734	15	2059	3749
36	3739	1406	76	2989	2764	16	2036	3770
37	3725	1443	77	2966	2794	17	2012	3790
38	3712	1480	78	2943	2823	18	1989	3810
39	3697	1517	79	2920	2852	19	1965	3830

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
1,20	0,1942	0,3849	1,70	0,0940	0,4554	2,40	0,0224	0,4918
21	1919	3869	71	0925	4564	42	0213	4922
22	1895	3888	72	0909	4573	44	0203	4927
23	1872	3907	73	0898	4582	46	0194	4931
24	1849	3925	74	0878	4591	48	0184	4934
25	1826	3944	75	0863	4599	50	0175	4938
26	1804	3962	76	0848	4608	52	0167	4941
27	1781	3980	77	0833	4616	54	0158	4945
28	1758	3997	78	0818	4625	58	0151	4948
29	1736	4015	79	0804	4633	53	0143	4951
1,30	0,1714	0,4032	1,80	0,0790	0,4641	2,60	0,0136	0,4053
31	1691	4049	81	0775	4649	62	0129	4956
32	1669	4066	82	0761	4656	64	0122	4959
33	1647	4082	83	0748	4664	66	0116	4961
34	1626	4099	84	0734	4671	68	0110	4963
35	1604	4115	85	0721	4678	70	0104	4965
36	1582	4131	86	0707	4686	72	0099	4667
37	1561	4147	87	0694	4693	74	0093	4969
38	1539	4162	88	0681	4699	76	0088	4971
39	1518	4177	89	0669	4706	78	0084	4973
1,40	0,1497	0,4192	1,90	0,0656	0,4713	0,80	0,0079	0,4974
41	1476	4207	91	0644	4719	82	0075	4976
42	1456	4222	92	0632	4726	84	0071	4977
43	1435	4236	93	0620	4732	86	0067	4979
44	1415	4251	94	0608	4738	88	0063	4980
45	1394	4265	95	0596	4744	90	0,0060	0,4981
46	1374	4279	96	0584	4750	92	0056	4982
47	1354	4292	97	0573	4756	94	0053	4984
48	1334	4306	98	0562	4761	96	0050	4985
49	1315	4319	99	0551	4767	98	0047	4986
1,50	0,1295	0,4332	2,00	0,0540	0,4772	3,00	0,00443	0,49865
51	1276	4345	02	0519	4783	3,10	00327	49903
52	1257	4357	04	0498	4793	3,20	00238	49931
53	1238	4370	06	0478	4803	3,30	00172	49952
54	1219	4382	08	0459	4812	3,40	00123	49966
55	1200	4394	10	0440	4821	3,50	00087	49977
56	1182	4406	12	0422	4830	3,60	00061	49984
57	1163	4418	14	0404	4838	3,70	00042	49989
58	1145	4429	16	0387	4846	3,80	00029	49993
59	1127	4441	18	0,0371	4854	3,90	00020	49995
1,60	0,1109	0,4452	2,20	0355	0,4861	4,00	0,0001338	0,499968
61	1092	4463	22	0339	4868	4,50	0000160	499997
62	1074	4474	24	0325	4875	5,00	0000015	49999997
63	1057	4484	26	0310	4881			
64	1040	4495	28	0297	4887			
65	1023	4505	30	0283	4893			
66	1006	4515	32	0270	4898			
67	0989	4525	34	0258	4904			
68	0973	4535	36	0246	4909			
69	0957	4545	38	0235	4913			

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Бернштейн С. Н., Теория вероятностей, ГИТТЛ, 1946.
2. Вентцель Е. С., Теория вероятностей, изд. «Наука», 1964.
3. Гливенко В. И., Курс теории вероятностей, ГОНТИ, 1939.
4. Гнеденко Б. В., Курс теории вероятностей, изд. «Наука», 1965.
5. Гончаров В. Л., Теория вероятностей, Оборонгиз, 1939.
6. Феллер В., Введение в теорию вероятностей и ее приложения, изд. «Мир», 1964.
7. Яглом А. М. и Яглом И. М., Вероятность и информация, Физматгиз, 1960.

Задачники

1. Гюнтер Н. М. и Кузьмин Р. О., Сборник задач по высшей математике, т. III, ГИТТЛ, 1947.
2. Володин Б. Г., Ганин М. П. и др., Сборник задач по теории вероятностей, математической статистике и теории случайных функций, под ред. А. А. Свешникова, изд. «Наука», 1965.

СОДЕРЖАНИЕ

Предисловие 2

Глава I. **Событие и вероятность**

§ 1. Основные понятия. Классическое определение вероятности . 3
§ 2. Сложные вероятности. Теоремы сложения и умножения. Условные вероятности 7
§ 3. Полная вероятность. Формула Бейеса 16
§ 4. Повторение испытаний. Схема Бернулли 20
§ 5. Примеры вычисления вероятностей 26
§ 6. Обобщение схемы Бернулли. Задача о безвозвратной выборке 34
§ 7. Цепь Маркова как обобщение схемы Бернулли 40
§ 8. Другие определения вероятности. Аксиомы теории вероятностей 47
Вопросы для самопроверки к главе I 54

Глава II. **Асимптотические формулы**

§ 9. Локальная теорема Муавра — Лапласа 56
§ 10. Нормальная функция распределения 62
§ 11. Теорема Пуассона 65
§ 12. Интегральная теорема Муавра — Лапласа. Теорема Бернулли 67
Вопросы для самопроверки к главе II 73

Глава III. **Случайные величины**

§ 13. Случайная величина и ее закон распределения 74
§ 14. Функция распределения и плотность вероятности 77
§ 15. Основные примеры дискретных и непрерывных распределений 84
§ 16. Числовые характеристики случайных величин. Математическое ожидание и дисперсия 92
§ 17. Степень неопределенности дискретного распределения. Понятие об энтропии 105
Вопросы для самопроверки к главе III 112

Глава IV. **Многомерные случайные величины и системы случайных величин**

§ 18. Двумерная случайная величина. Функция распределения и плотность вероятности 114
§ 19. Нормальное распределение двумерной случайной величины . 122
§ 20. Числовые характеристики системы двух случайных величин 129
§ 21. Теоремы о математическом ожидании и дисперсии случайных величин 138
§ 22. Многомерная случайная величина и система случайных величин. Суммирование случайных величин. Композиция законов распределения 142
§ 23. Закон больших чисел и его обобщения. Центральная предельная теорема А. М. Ляпунова 147
Вопросы для самопроверки к главе IV 154
Приложение. Таблица функций $\varphi(x)$ и $\Phi(x)$ 156
Рекомендуемая литература 158

Рафаил Самойлович Гутер
Борис Владимирович Овчинский

ОСНОВА ТЕОРИИ ВЕРОЯТНОСТЕЙ

Редактор *Ю. А. Гастев*
Обложка художника *И. А. Тарасова*
Художественный редактор *В. С. Эрденко*
Технический редактор *Н. Н. Махова*
Корректор *М. В. Голубева*

Сдано в набор 24/I 1967 г. Подписано к печати 27/VI 1967 г. $60 \times 90^1/_{16}$. Типографская № 2. Печ. л. 10. Уч.-изд. л. 8,6. Тираж 30 тыс. экз. А 10059. Заказ № 26.

Издательство «Просвещение» Государственного комитета по печати при Совете Министров РСФСР Москва, 3-й проезд Марьиной рощи, 41.

Саратовский полиграфический комбинат Росглавполиграфпрома Комитета по печати при Совете Министров РСФСР. г. Саратов, ул. Чернышевского, 59.

Цена без переплета 24 к., переплет 18 к.

42 к.