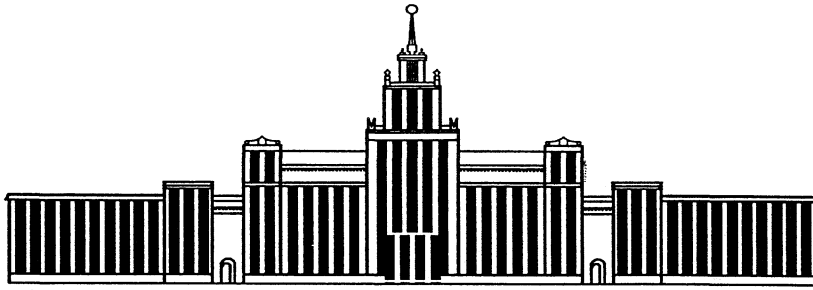
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ



ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

519.2(07)
К707

М.А. Корытова, С.А. Шунайлова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Челябинск
2019

Министерство науки и высшего образования Российской Федерации
Южно-Уральский государственный университет
Институт естественных и точных наук

519.2(07)
К707

М.А. Корытова, С.А. Шунайлова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Челябинск
Издательский центр ЮУрГУ
2019

УДК 519.21(075.8)
К707

Одобрено
учебно-методической комиссией
института естественных и точных наук

Рецензенты:
Кипнис М.М., Лепчинский М.Г.

Корытова, М.А.
К707 Теория вероятностей: учебное пособие / М.А. Корытова, С.А. Шунайлова. – Челябинск: Издательский центр ЮУрГУ, 2019. – 105 с.

Учебное пособие содержит теоретический материал, соответствующий разделу математики «Теория вероятностей и математическая статистика», изучаемому студентами экономических и технических направлений. После каждого раздела приведены контрольные вопросы для проверки усвоения учебного материала. Пособие составлено на основе работ из библиографического списка и преподавательского опыта авторов.

Учебное пособие предназначено для студентов следующих направлений подготовки бакалавров и специалистов: 29.03.04, 38.03.04, 38.05.01, 43.03.01, 43.03.02 и других, а также для преподавателей, читающих лекции для этих студентов.

УДК 519.21(075.8)

КОМБИНАТОРИКА

Пусть имеется n различных объектов произвольной природы (множество объектов). Выберем из них k объектов. Полученное подмножество называется *выборкой*.

Комбинаторика – это раздел математики, который изучает, сколькими способами можно осуществить такой выбор согласно заданным условиям.

Выборки будем различать по двум признакам.

1. Упорядоченность – неупорядоченность.

Выборку будем считать *упорядоченной*, если при перестановке двух элементов местами получится другая, отличная от данной, выборка. Например, число, номер, слово.

Выборку будем считать *неупорядоченной*, если при перестановке двух элементов местами выборка не изменится. Например, группа людей, все студенты в аудитории.

2. Повторность – бесповторность.

Выборку будем считать *повторной*, если она может содержать одинаковые элементы. Такие выборки получаются, если после выбора одного элемента из данного множества, перед следующим выбором его возвращают обратно.

Выборку будем считать *бесповторной*, если все элементы в ней различны. Такие выборки получаются, если элементы из данного множества выбираются без возвращения.

Правило умножения

Если объект А можно выбрать k_1 различными способами, а после этого объект В – k_2 различными способами, то пару объектов А и В именно в таком порядке можно выбрать $k_1 \cdot k_2$ различными способами.

Другая интерпретация этого правила такова. Пусть первое действие можно совершить k_1 различными способами, второе – k_2 различными способами. Тогда оба действия можно совершить $k_1 \cdot k_2$ различными способами.

Пример. Пусть из пункта А в пункт В имеется 5 дорог, а из пункта В в пункт С – 6 дорог.

1) Сколько существует различных вариантов проезда из пункта А в пункт С?

Решение. Существует 5 различных путей из пункта А в пункт В – это 5 способов выполнения первого действия, при этом существует 6 различных путей из пункта В в пункт С – это 6 различных способов выполнения

второго действия. Согласно правилу умножения, число различных способов выбора пути из пункта А в пункт С равно $5 \cdot 6 = 30$.

2) Сколько существует различных вариантов проезда из пункта А в пункт В и обратно?

Решение. Из пункта А в пункт В ведет 5 дорог, значит, имеется 5 способов проезда туда и 5 способов проезда обратно. По правилу умножения число всех способов проезда туда и обратно равно $5 \cdot 5 = 25$.

3) Сколько существует различных вариантов проезда из пункта А в пункт В и обратно при условии, что дороги туда и обратно будут разными?

Решение. Рассуждаем аналогично пункту 2), но учитываем, что дороги туда и обратно не должны совпадать, т.е. при выборе одного из 5-ти способов проезда «туда» обратно можно вернуться одним из 4-х способов. Поэтому число различных способов проехать из пункта А в пункт В и вернуться обратно, но обязательно другой дорогой, равно $5 \cdot 4 = 20$.

Пример. Сколько трехзначных чисел можно составить из цифр 1, 2, 3, 4, 5, 6 при дополнительном условии.

1) Цифры в числе не повторяются.

Решение. Первую цифру выбираем из всего представленного набора, значит, существует 6 способов выбора. Вторую – из набора цифр на одну, выбранную первой, цифру меньше, т.е. 5 способов выбора. Третью из набора, который меньше еще на одну цифру. Получаем ответ: $6 \cdot 5 \cdot 4 = 120$.

2) Цифры в числе могут повторяться.

Решение. И первую, и вторую, и третью цифры выбираем из всего представленного набора. Получаем ответ: $6 \cdot 6 \cdot 6 = 216$.

3) Число должно быть четным.

Решение. Первую и вторую цифры выбираем из всего представленного набора, третью – только из четных, т.к. число является четным, если оно заканчивается на четное число. Получаем ответ: $6 \cdot 6 \cdot 3 = 108$.

Правило сложения

Если объект А можно выбрать k_1 различными способами, а объект В – k_2 различными способами, то выбрать один объект А **или** В можно $k_1 + k_2$ различными способами.

Пример. В ящике лежат 2 красных, 3 зеленых и 4 черных шара. Сколькими способами можно выбрать цветной шар?

Решение. Выбрать цветной шар – значит выбрать красный или зеленый шар. Количество способов равно $2 + 3 = 5$.

Пример. На столе лежит 4 яблока, 6 груш и 2 морковки. Сколькими способами можно выбрать один фрукт?

Решение. Выбрать фрукт – значит выбрать яблоко (4 способа) или грушу (6 способов). Общее количество способов равно $4 + 6 = 10$.

Размещения

Пусть имеется множество из n различных объектов (элементов), т.е. объекты имеют или разные названия или разные номера. Пусть $k \le n,\ k \in \mathbf{N}$.

Размещением из n элементов по k называется любое подмножество, содержащее k элементов, взятых из данных n элементов с учетом порядка выбора. То есть подмножества различаются или элементами, входящими в них, или порядком, в котором расположены элементы.

Например, размещениями из трех цифр (1, 2, 3) по 2 элемента будут 1-2, 2-1, 1-3, 3-1, 2-3 и 3-2. Их количество равно 6.

Число различных размещений из n элементов по k можно найти по правилу умножения:

$$A_n^k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}, \quad k = 0, \ldots, n.$$

Здесь $n! = 1 \cdot 2 \cdot \ldots \cdot n$ (читается как «*н факториал*»).

Полагают $0! = 1$.

Например, $1! = 1$; $2! = 1 \cdot 2 = 2$; $3! = 1 \cdot 2 \cdot 3 = 6$; $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$ и т.д.

Пример. Сколькими способами можно выбрать старосту и профорга из группы в 10 человек?

Решение. При выборе двух человек на должности старосты и профорга имеет значение порядок, в котором это делается. Значит, искомое число – это число размещений из 10 элементов по 2 элемента:

$$A_{10}^2 = \frac{10!}{(10-2)!} = \frac{10!}{8!} = 9 \cdot 10 = 90.$$

Размещения с повторениями

Два размещения с повторениями одинаковы тогда и только тогда, когда на одинаковых местах находятся одни и те же элементы.

Например, все возможные размещения с повторениями из трех элементов a, b, c по 2 элемента:

$$aa, ab, ac, ba, bb, bc, ca, cb, cc.$$

Число размещений с повторениями из n элементов по k элементов равно

$$A_n^k = n^k, \quad k = 0, 1, \ldots .$$

Пример. На пути школьника от дома до школы три светофора. Сколько может быть различных комбинаций их сигналов?

Решение. Например, могут быть следующие варианты (К – красный, З – зеленый свет): ККК, ККЗ, КЗЗ, …. Т.е. из двух возможных сигналов выбираются три. Общее количество $A_2^3 = 2^3 = 8$.

Перестановки

Перестановки – это размещения из n элементов по n элементов. Их число равно

$$A_n^n = P_n = n \cdot (n-1) \cdot \ldots \cdot 1 = n!.$$

Пример. На полке наудачу располагаются 10 книг. Сколько существует различных способов их расположения?

Решение. 10 книг образуют множество из $n = 10$ различных элементов, так как книги разные. Расположение книг на полке – это упорядочивание книг слева направо. Таким образом, расположение книг на полке – перестановка из 10-ти элементов. Поэтому число различных расположений 10-ти книг на полке совпадает с числом различных перестановок из 10-ти элементов и находится по формуле $P_{10} = 10! = 3628800$.

Пример. На шести карточках написаны буквы, из которых можно составить слово РОДНИК. Сколько существует различных шестибуквенных «слов», которые можно составить при помощи этих шести карточек?

Решение. $P_6 = 6! = 720$.

Перестановки с повторениями

Пусть имеются k групп элементов, причем в первой группе n_1 неразличимых элементов, во второй группе n_2 неразличимых элементов, …, в k-ой группе – n_k неразличимых элементов. Элементы из разных групп различимы. Таким образом, имеем всего $n_1 + n_2 + \ldots + n_k = n$ элементов.

Если бы все n элементов были различны, то число их перестановок равно $n!$. Но т.к. некоторые элементы в выборке повторяются, то при их перестановке друг с другом выборка не изменится. В этом случае имеют место *перестановки с повторениями*. Их число равно

$$P_n(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}.$$

Пример. На пяти карточках написаны буквы, из которых можно составить слово ГОРОД. Сколько существует различных пятибуквенных слов, которые можно составить при помощи этих пяти карточек?

Решение. Пять карточек разобьем на 4 группы. Первая группа образована карточками с буквой О. Число таких карточек равно 2. Они неразличимы по буквам, на всех одна и та же буква О, $n_1 = 2$. Остальные группы образованы каждая одной карточкой, $n_2 = n_3 = n_4 = 1$.

Таким образом, мы имеем дело с перестановками с повторениями и число слов из пяти букв равно: $P_5(2, 1, 1, 1) = \frac{(2+1+1+1)!}{2! \cdot 1! \cdot 1! \cdot 1!} = \frac{5!}{2!} = 60$.

Пример. На пяти карточках написаны буквы, из которых можно составить слово ТОПОТ. Сколько существует различных пятибуквенных слов, которые можно составить при помощи этих пяти карточек?

Решение. Пять карточек разобьем на 3 группы. Первая группа образована карточками с буквой О. Число таких карточек равно 2. Вторая – карточками с буквой Т. Третья карточкой с буквой П. $n_1 = n_2 = 2$, $n_3 = 1$.

Таким образом, $P_5(2, 2, 1) = \dfrac{(2+2+1)!}{2!\cdot 2!\cdot 1!} = \dfrac{120}{4} = 30$.

Сочетания

Сочетанием из n элементов по k называется любое подмножество, содержащее k элементов, взятых из данных n элементов без учета порядка выбора.

То есть подмножества различаются только элементами, входящими в них, а порядок, в котором они расположены, не имеет значения.

Число различных сочетаний из n элементов по k можно найти по формуле:

$$C_n^k = \frac{n!}{k!(n-k)!}, \quad k = 0,\ldots,n.$$

Например, сочетаниями из трех цифр (1, 2, 3) по 2 элемента будут 1-2, 1-3 и 2-3. Их количество равно трем.

$$C_5^2 = \frac{5!}{2!\cdot 3!} = \frac{\cancel{1}\cdot\cancel{2}\cdot\cancel{3}\cdot 4\cdot 5}{\cancel{1}\cdot\cancel{2}\cdot 1\cdot 2\cdot\cancel{3}} = \frac{4\cdot 5}{2} = 10,$$

$$C_3^0 = \frac{3!}{0!\cdot 3!} = 1,\ C_4^1 = \frac{4!}{1!3!} = \frac{\cancel{1}\cdot\cancel{2}\cdot\cancel{3}\cdot 4}{\cancel{1}\cdot 1\cdot\cancel{2}\cdot\cancel{3}} = 4,\ C_4^3 = \frac{4!}{1!\cdot 3!} = \frac{4!}{3!\cdot 1!} = C_4^1 = 4,$$

$$C_{40}^{37} = \frac{40!}{37!\cdot 3!} = \frac{38\cdot\overset{13}{\cancel{39}}\cdot\overset{20}{\cancel{40}}}{1\cdot\cancel{2}\cdot\cancel{3}} = 9880,\ C_7^2 = \frac{7!}{2!5!} = \frac{\overset{3}{\cancel{6}}\cdot 7}{1\cdot\cancel{2}} = 21.$$

Справедливы следующие свойства: $C_n^0 = C_n^n = 1$, $C_n^1 = C_n^{n-1} = n$, $C_n^k = C_n^{n-k}$.

Пример. В урне находятся 10 белых, 15 черных, 20 красных шаров. Найти:

1) сколькими различными способами можно вынуть девять шаров;

2) сколькими различными способами можно вынуть девять шаров, среди которых 6 белых и 3 черных.

Решение. 1) Всего в урне 45 шаров. Считаем, что шары различимы, например, пронумерованы. Следовательно, имеем множество из $n = 45$ различных объектов.

Наудачу взятые девять шаров образуют подмножество из $k = 9$ элементов. Это подмножество определяется лишь элементами, попавшими в него, порядок не имеет значения.

Следовательно, это сочетание из 45 элементов по 9:

$$C_{45}^{9} = \frac{45!}{9!36!} = \frac{37 \cdot 38 \cdot 39 \cdot 40 \cdot 41 \cdot 42 \cdot 43 \cdot 44 \cdot 45}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9} = 886\,163\,135.$$

2) Взятие девяти шаров, из которых 6 белых и 3 черных, можно разбить на два действия: первое действие – возьмем 6 белых шаров из 10 белых шаров, находящихся в урне (это можно сделать C_{10}^{6} различными способами); второе действие – возьмем 3 черных шара из общего числа 15 черных шаров (это можно сделать C_{15}^{3} различными способами).

Тогда число различных способов вынуть девять шаров нужного состава по правилу умножения равно

$$C_{10}^{6} \cdot C_{15}^{3} = \frac{10!}{6!4!} \cdot \frac{15!}{3!12!} = \frac{7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4} \cdot \frac{13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3} = 95\,550.$$

Формулы для подсчета количества выборок сведены в таблицу (табл. 1).

Таблица 1

Виды выборок	Упорядоченная	Неупорядоченная
Бесповторная	Размещения $A_n^k = \frac{n!}{(n-k)!}, \quad k = 0, \ldots, n.$ Перестановки $P_n = n!$	Сочетания $C_n^k = \frac{n!}{k!(n-k)!}, \quad k = 0, \ldots, n$
Повторная	Размещения с повторениями $A_n^k = n^k, \quad k = 0, 1, \ldots.$ Перестановки с повторениями $P_n(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$	Сочетания с повторениями $C_n^k = C_{n+k-1}^k = \frac{(n+k-1)!}{k!(n-1)!}, \quad k = 0, 1, \ldots$

Контрольные вопросы

1. Что изучает комбинаторика?
2. Приведите два признака, по которым одна выборка отличается от другой.
3. Какие выборки называются сочетаниями? Приведите формулу для подсчета количества сочетаний.

4. Какие выборки называются размещениями? Приведите формулу для подсчета количества размещений.

5. Какие выборки называются перестановками? Приведите формулу для подсчета количества перестановок.

6. Какие выборки называются перестановками с повторениями? Приведите формулу для подсчета количества перестановок с повторениями.

7. Какие выборки называются размещениями с повторениями? Приведите формулу для подсчета количества размещений с повторениями.

8. Сформулируйте правило умножения.

9. Сформулируйте правило сложения.

ВЕРОЯТНОСТЬ СОБЫТИЯ

Случайные события. Вероятность

Теория вероятностей, как следует из названия, имеет дело с вероятностями. Нас окружают множество вещей и явлений, о которых, как бы ни была развита наука, нельзя сделать точных прогнозов. Мы не знаем, какую карту вытянем из колоды наугад или сколько дней в мае будет идти дождь, но, имея некоторую дополнительную информацию, можем строить прогнозы и вычислять вероятности этих случайных событий.

Пусть производится некоторая совокупность действий с неизвестным заранее результатом. Такая совокупность в теории вероятностей называется *экспериментом (опытом, испытанием)*.

Теория вероятностей изучает закономерности, которые проявляются при массовой реализации эксперимента.

Основным объектом классической теории вероятности является так называемое *случайное событие*. Случайное событие – это любой из возможных результатов эксперимента.

Например, эксперимент – бросание игрального кубика с шестью гранями.

События:

A – выпадение «1» на верхней грани;

B – выпадение четного числа на верхней грани;

C – выпадение числа, делящегося на 3;

D – выпадение числа, большего 10;

E – выпадение числа, меньшего 10.

Событие D не может произойти ни при какой реализации эксперимента. Оно называется *невозможным*. Будем его обозначать буквой V.

Событие E напротив происходит при любой реализации эксперимента. Оно называется *достоверным*. Будем его обозначать буквой U.

Пример. Стрелок стреляет по мишени, разделенной на четыре области. Выстрел – это испытание. Попадание в определенную область мишени – событие.

Пример. В урне имеются цветные шары. Из урны наудачу берут один шар. Извлечение шара из урны есть испытание. Появление шара определенного цвета – событие.

Замечание. В приведенном выше примере с игральным кубиком можно выделить шесть так называемых элементарных событий (элементарных исходов): выпадение конкретного числа очков от 1 до 6. Обозначим их через e_1, e_2, e_3, e_4, e_5, e_6, где индекс указывает выпавшее число очков. Множество этих событий называется множеством элементарных исходов. Тогда любое событие можно представить как подмножество этого множества.

В примере выше: $A=\{e_1\}$, $B=\{e_2, e_4, e_6\}$, $C=\{e_3, e_6\}$, $D=\varnothing$, $E=\{e_1, e_2, e_3, e_4, e_5, e_6\}$.

Несовместные события

События называют *несовместными*, если появление одного из них исключает появление других событий в одном и том же испытании.

Пример. Из ящика с деталями наудачу извлечена деталь. Появление стандартной детали исключает появление нестандартной детали. События «появилась стандартная деталь» и «появилась нестандартная деталь» – несовместные.

Пример. Брошена монета. Появление «герба» исключает появление цифры (решки). События «появился герб» и «появилась цифра» – несовместные.

Полная группа событий

Несколько событий образуют *полную группу*, если в результате испытания появится хотя бы одно из них.

Другими словами, появление хотя бы одного из событий полной группы есть достоверное событие.

В частности, если события, образующие полную группу, попарно несовместны, то в результате испытания появится одно и только одно из этих событий.

Пример. Приобретены два билета денежно-вещевой лотереи. Обязательно произойдет одно и только одно из следующих событий: «выигрыш выпал на первый билет и не выпал на второй», «выигрыш не выпал на первый билет и выпал на второй», «выигрыш выпал на оба билета», «на оба билета выигрыш не выпал». Эти события образуют полную группу попарно несовместных событий.

Пример. Стрелок произвел выстрел по цели. Обязательно произойдет одно из следующих двух событий: попадание, промах. Эти два несовместных события образуют полную группу.

Равновозможные события

События называют *равновозможными*, если есть основания считать, что ни одно из них не является более возможным, чем другое.

Пример. Появление «герба» и появление надписи при бросании монеты – равновозможные события. Действительно, предполагается, что монета изготовлена из однородного материала, имеет правильную цилиндрическую форму, и наличие чеканки не оказывает влияния на выпадение той или иной стороны монеты.

Пример. Появление того или иного числа очков на брошенной игральной кости – равновозможные события. Предполагается, что игральная кость изготовлена из однородного материала и имеет форму правильного многогранника.

Классическое определение вероятности

Числовая величина, характеризующая степень возможности наступления данного события, называется его *вероятностью*.

Вероятность события должна обладать следующими свойствами:

1. $0 \le P(A) \le 1$, т.к. $1 \le n \le m$.
2. $P(V) = 0$, где V – невозможное событие.
3. $P(U) = 1$, где U – достоверное событие.

Если можно пересчитать все возможные исходы проводимого опыта (элементарные исходы) и если эти исходы равновозможны, то говорят, что имеет место схема случаев.

Пусть n – число всех равновозможных исходов данного опыта, а m – число его исходов, при которых происходит некоторое событие A (благоприятные или благоприятствующие событию A).

Тогда вероятность события A определяется как отношение числа благоприятных исходов к числу всех возможных исходов:

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

Например, эксперимент – бросание игрального кубика с шестью гранями. Общее число элементарных исходов $n = 6$.

События:

A – выпадение «1» на верхней грани;

B – выпадение четного числа на верхней грани;

C – выпадение числа, делящегося на 3;

D – выпадение числа, большего 10;

E – выпадение числа, меньшего 10.

Вероятности:

$$P(A)=\frac{1}{6},\ P(B)=\frac{3}{6}=\frac{1}{2},\ P(C)=\frac{2}{6}=\frac{1}{3},\ P(D)=\frac{0}{6}=0,\ P(E)=\frac{6}{6}=1.$$

Пример. Набирая номер телефона, абонент забыл одну цифру и набрал ее наудачу. Найти вероятность того, что набрана нужная цифра.

Решение. Обозначим через A событие – набрана нужная цифра. Абонент мог набрать любую из 10 цифр, поэтому общее число возможных элементарных исходов равно 10.

Эти исходы несовместны, равновозможны и образуют полную группу. Благоприятствует событию A лишь один исход (нужная цифра лишь одна). Искомая вероятность равна отношению числа исходов, благоприятствующих событию, к числу всех элементарных исходов:

$$P(A)=\frac{1}{10}.$$

Пример. Набирая номер телефона, абонент забыл последние две цифры и, помня лишь, что эти цифры различны, набрал их наудачу. Найти вероятность того, что набраны нужные цифры.

Решение. Обозначим через B событие – набраны две нужные цифры.

Всего можно набрать столько различных цифр, сколько может быть составлено размещений из десяти цифр по две, т.е. $n=A_{10}^{2}=10\cdot 9=90$. Таким образом, общее число возможных элементарных исходов равно $n=90$.

Эти исходы несовместны, равновозможны и образуют полную группу. Благоприятствует событию B лишь один исход. Искомая вероятность равна отношению числа исходов, благоприятствующих событию, к числу всех элементарных исходов:

$$P(B)=\frac{1}{90}.$$

Пример. Указать ошибку «решения» задачи. Брошены две игральные кости. Найти вероятность того, что сумма выпавших очков равна 4 (событие A).

Решение. Всего возможны 2 исхода испытания: сумма выпавших очков равна 4, сумма выпавших очков не равна 4. Событию A благоприятствует один исход: общее число исходов равно двум. Следовательно, искомая вероятность

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

Ошибка этого решения состоит в том, что рассматриваемые исходы не являются равновозможными.

Правильное решение. Общее число равновозможных исходов испытания равно $n = 6 \cdot 6 = 36$ (каждое число выпавших очков на одной кости может сочетаться со всеми числами очков другой кости). Среди этих исходов благоприятствуют событию A только 3 исхода: (1; 3), (3; 1), (2; 2) (в скобках указаны количества выпавших очков). Следовательно, искомая вероятность

$$P(A) = \frac{3}{36} = \frac{1}{12}.$$

Пример. В партии из 10 деталей 7 стандартных. Найти вероятность того, что среди шести взятых наудачу деталей 4 стандартных.

Решение. Общее число возможных элементарных исходов испытания равно числу способов, которыми можно извлечь 6 деталей из 10, т.е. числу сочетаний из 10 элементов по 6 элементов $\left(n = C_{10}^{6}\right)$.

Четыре стандартные детали можно взять из семи стандартных деталей C_7^4 способами; при этом остальные 6 – 4 = 2 детали должны быть нестандартными; взять же 2 нестандартные детали из 10 – 7 = 3 нестандартных деталей можно C_3^2 способами. Следовательно, число благоприятствующих исходов равно $m = C_7^4 \cdot C_3^2$.

Искомая вероятность равна отношению числа исходов, благоприятствующих событию, к числу всех элементарных исходов:

$$P(A) = \frac{C_7^4 \cdot C_3^2}{C_{10}^6} = \frac{1}{2}.$$

Статистическое определение вероятности

Относительной частотой события A называется отношение числа опытов, в результате которых произошло событие A к общему числу опытов.

Отличие относительной частоты от вероятности заключается в том, что вероятность вычисляется без непосредственного произведения опытов, а относительная частота – после опыта.

Пусть в коробке находится 10 шаров: 3 красных, 2 зеленых, остальные белые. Если наугад извлечено 5 шаров и 2 из них оказались красными, то относительная частота появления красного шара равна: $W(A) = \frac{2}{5}$.

Вероятность же этого события равна $P(A)=\frac{3}{10}$. Как видно, относительная частота не совпадает с найденной вероятностью.

Однако при достаточно большом числе произведенных опытов относительная частота изменяется мало, колеблясь около одного числа. Это число может быть принято за вероятность события (статистическая вероятность).

Длительные наблюдения показали, что если в одинаковых условиях производят опыты, в каждом из которых число испытаний достаточно велико, то относительная частота обнаруживает свойство устойчивости. Это свойство состоит в том, что в различных опытах, относительная частота изменяется мало (тем меньше, чем больше произведено испытаний), колеблясь около некоторого постоянного числа. Это постоянное число и есть вероятность появления события.

Пример. По данным шведской статистики, относительная частота рождения девочек за 1935 г. по месяцам характеризуется следующими числами (числа расположены в порядке следования месяцев, начиная с января): 0,486; 0,489; 0.490; 0,471; 0,478; 0,482; 0,462; 0,484; 0,485; 0,491; 0,482; 0,473.

Относительная частота колеблется около числа 0,482, которое можно принять за приближенное значение вероятности рождения девочек.

Статистические данные различных стран дают примерно такие же значения относительной частоты.

Пример. Многократно проводились опыты по бросанию монеты, в которых подсчитывали число появления «герба». Результаты нескольких опытов приведены в табл. 2.

Таблица 2

Число бросаний	Число появлений «герба»	Относительная частота
4040	2048	0,5069
12000	6019	0,5016
24000	12012	0,5005

Здесь относительные частоты незначительно отклоняются от числа 0,5, причем тем меньше, чем больше число испытаний.

Например, при 4040 испытаниях отклонение равно 0,0069, а при 24000 испытаний – лишь 0,0005. Приняв во внимание, что вероятность появления «герба» при бросании монеты равна 0,5, мы видим, что относительная частота колеблется около вероятности.

Геометрическая вероятность

Классическое определение вероятности неприменимо к испытаниям с бесконечным числом исходов. Чтобы преодолеть этот недостаток вводится понятие *геометрической вероятности*, т.е. вероятности попадания точки в какой-либо отрезок, часть плоскости или часть пространства.

Множество всех элементарных исходов рассматривается как множество точек на прямой, на плоскости или в пространстве, а событие – его подмножество.

Если на отрезке длиной L выделен отрезок длины l, то за вероятность попадания наугад взятой точки в этот отрезок принимают отношение $\frac{l}{L}$.

Если внутри плоской области площадью S выделена область площадью s, то за вероятность попадания наугад взятой точки в меньшую область принимают отношение $\frac{s}{S}$.

Если внутри тела объемом V выделена часть объема v, то за вероятность попадания наугад взятой точки в меньшую часть принимают отношение $\frac{v}{V}$.

Пример. В круг радиуса R случайным образом брошена точка, положение которой равновозможно в любом месте круга. Какова вероятность, что эта точка окажется внутри вписанного в этот круг правильного треугольника (рис. 1)?

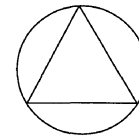


Рис. 1

Решение. Площадь круга равна $S=\pi R^2$. Площадь вписанного треугольника равна $s=\frac{3\sqrt{3}R^2}{4}$ Вероятность равна

$$P=\frac{3\sqrt{3}R^2}{4}:\pi R^2=\frac{3\sqrt{3}R^2}{4\pi R^2}=\frac{3\sqrt{3}}{4\pi}\approx 0,4137.$$

Пример. В куб с ребром 1 случайным образом брошена точка, положение которой равновозможно в любом месте куба. Какова вероятность, что расстояние от нее до центра куба больше 0,5?

Решение. Объем куба $V=1^3=1$. Точки, расстояние от которых до центра куба больше 0,5, находятся вне вписанного в этот куб шара. Его объем $v=1-\frac{4}{3}\cdot\pi\cdot(0,5)^3=1-\frac{\pi}{6}=\frac{6-\pi}{6}$. Вероятность равна

$$P=\frac{6-\pi}{6}:1=\frac{6-\pi}{6}\approx 0,4767.$$

Пример (задача о встрече). Два лица условились встретиться в определенном месте между двумя и тремя часами дня. Пришедший первым ждет другого в течение 10 минут, после чего уходит. Чему равна вероятность встречи этих лиц, если каждый из них может прийти в любое время в течение указанного часа независимо от другого?

Решение. Будем считать интервал с 14 до 15 часов отрезком [0, 1] длиной в 1 час. Пусть x и y – моменты прихода лиц. Все возможные результаты эксперимента – точки квадрата со стороной 1.

Можно считать, что эксперимент сводится к бросанию точки наудачу в квадрат. При этом благоприятными исходами являются точки множества $A=\left\{(x,y):\ x\in[0;1],\ y\in[0;1],\ |x-y|\le\frac{1}{6}\right\}$ (рис. 2).

Тогда вероятность встречи равна

$$P(A)=\frac{1-\left(5/6\right)^2}{1}=\frac{11}{36}.$$

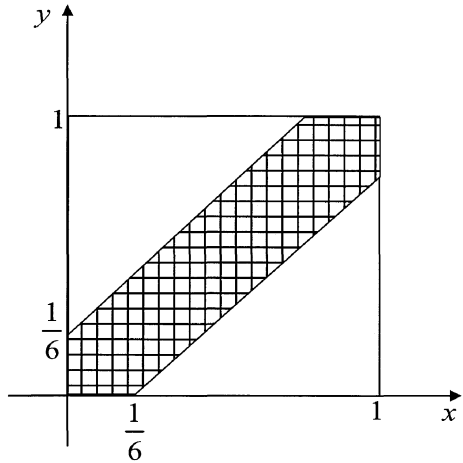


Рис. 2

Действия над событиями

Событие, которое заключается в том, что событие A не произошло в результате эксперимента, называется *противоположным* событию A и обозначается $\overline{A}$.

Суммой событий A_1, A_2, …, A_n называется событие, заключающееся в том, что произошло хотя бы одно из событий A_1, A_2, …, A_n (или A_1, или A_2, …, или A_n, или несколько из них, или все).

Обозначается: $A_1 + A_2 + ... + A_n$.

Суммой $A+B$ событий A и B называется событие, заключающееся в том, что произошло хотя бы одно из двух событий (A или B или оба).

Произведением событий A_1, A_2, ..., A_n называется событие, заключающееся в том, что произошли одновременно все события A_1, A_2, ..., A_n (и A_1, и A_2, ..., и A_n).

Обозначается: $A_1 A_2 ... A_n$.

Произведением AB событий A и B называется событие, состоящее в том, что произошли оба данных события одновременно (A и B).

Говорят, что событие A *влечет* событие B или B *является следствием* события A $(A \subset B)$, если при наступлении события A событие B также обязательно наступит.

Пример. Подбрасывается игральный кубик. A – выпадение 6 очков, B – выпадение трех очков, C – выпадение четного числа очков, D – выпадение числа очков, кратного трем.

Между этими событиями есть следующие соотношения:

$$A \subset D,\ B \subset D,\ A+B=D,\ CD=A.$$

Пример. Подбрасывается игральный кубик. Обозначим A_k – выпадение k очков $(k=1,2,3,4,5,6)$, A – выпадение четного числа очков, B – выпадение нечетного числа очков, C – выпадение числа очков, кратного трем, D – выпадение числа очков, большего трех. Выразить события A, B, C, D через A_k.

Решение.

$$A = A_2 + A_4 + A_6,\ B = A_1 + A_3 + A_5,\ C = A_3 + A_6,\ D = A_4 + A_5 + A_6.$$

Пример. Стрелок производит три выстрела по мишени. Обозначим через A_k – попадание при выстреле №k $(k=1, 2, 3)$, A – хотя бы одно попадание, B – три попадания, C – три промаха, D – хотя бы один промах, E – не меньше двух попаданий, F – не более одного попадания, G – попадание после первого выстрела. Выразить события A, B, C, D, E, F, G через A_k.

Решение.

$$A = A_1 + A_2 + A_3 = \overline{\overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3}},\ B = A_1 \cdot A_2 \cdot A_3,\ C = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3},$$

$$D = \overline{A_1} + \overline{A_2} + \overline{A_3} = \overline{A_1 \cdot A_2 \cdot A_3},$$

$$E = A_1 \cdot A_2 \cdot A_3 + \overline{A_1} \cdot A_2 \cdot A_3 + A_1 \cdot \overline{A_2} \cdot A_3 + A_1 \cdot A_2 \cdot \overline{A_3},$$

$$F = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} + \overline{A_1} \cdot \overline{A_2} \cdot A_3 + A_1 \cdot \overline{A_2} \cdot \overline{A_3} + \overline{A_1} \cdot A_2 \cdot \overline{A_3},$$

$$G = \overline{A_1} \cdot A_2 + \overline{A_1} \cdot \overline{A_2} \cdot A_3.$$

Вероятность наступления события хотя бы один раз

Пусть n – общее число элементарных исходов эксперимента, m – число элементарных исходов, благоприятствующих событию A. Тогда событию $\overline{A}$ благоприятствует $n-m$ элементарных исходов. Поэтому $P(A)+P(\overline{A})=\frac{m}{n}+\frac{n-m}{n}=1$. Тогда $P(A)=1-P(\overline{A})$.

Эту формулу можно использовать при нахождении вероятности наступления некоторого исхода хотя бы один раз. Противоположный исход – событие не наступило ни одного раза.

Пример. В урне находятся 25 белых и 5 черных шаров. Найти вероятность того, что среди девяти извлеченных шаров имеется хотя бы один шар черного цвета.

Решение. Всего в урне 30 шаров. Будем считать, что все они пронумерованы. Эти 30 шаров разделяются на две группы. Первая группа состоит из двадцати пяти белых шаров, вторая группа состоит из пяти черных шаров. Эксперимент состоит в изъятии наудачу девяти шаров из тридцати шаров (их порядок не имеет значения). Элементарным исходом в этом эксперименте является любое сочетание из тридцати элементов по девять. Тогда число таких элементарных исходов равно

$$n=C_{30}^{9}=\frac{30!}{9!\,21!}=14\,307\,150.$$

Событие A – среди девяти вынутых шаров имеется хотя бы один черный шар.

Событие $\overline{A}$ означает, что нет ни одного черного шара среди вынутых или что все девять шаров – белые.

$$m_{\overline{A}}=C_{25}^{9}=\frac{25!}{9!\cdot 16!}=\frac{17\cdot 18\cdot 19\cdot 20\cdot 21\cdot 22\cdot 23\cdot 24\cdot 25}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9}=2\,042\,975,$$

$$P(\overline{A})=\frac{C_{25}^{9}}{C_{30}^{9}}\approx 0{,}14.$$

Поэтому $P(A)=1-P(\overline{A})\approx 1-0{,}14=0{,}86$.

Теоремы сложения и умножения

Пусть в ящике лежит a белых и b черных шаров. Наугад вынимаем два шара последовательно один за другим. Тогда вероятность того, что первый шар оказался белым (событие A), равна

$$P(A)=\frac{a}{a+b},$$

а вероятность того, что второй шар оказался белым (событие B), равна $\frac{a-1}{a+b-1}$, т.к. после того, как событие A произошло, белых шаров осталось на один меньше, чем было вначале, и общее число шаров уменьшилось также на один. Если бы событие A не произошло, т.е. первым был черный шар, то вероятность события B была бы равна $\frac{a}{a+b-1}$.

Таким образом, вероятность события B зависит от того, произошло ли до этого событие A. Такие события называются зависимыми. Вероятность события B при условии, что произошло событие A, называется условной вероятностью B при условии A. Обозначается $P_A(B)$.

Если же наступление события A не изменяет вероятности события B, то A и B называются независимыми.

Найдем вероятность произведения двух событий. Пусть n – общее число элементарных исходов эксперимента, m – число элементарных исходов, благоприятствующих событию A, k – событию B, s – событию AB. При этом $s \le m$ и $s \le k$.

Если событие A произошло, то это означает, что наступил один из m элементарных исходов, ему благоприятствующих. Из них благоприятствуют событию B те элементарные исходы, которые благоприятствуют событию AB, т.е. s элементарных исходов.

Таким образом,

$$P_A(B) = \frac{s}{m} = \frac{\frac{s}{n}}{\frac{m}{n}} = \frac{P(AB)}{P(A)}.$$

Аналогично можно показать, что

$$P_B(A) = \frac{s}{k} = \frac{\frac{s}{n}}{\frac{k}{n}} = \frac{P(AB)}{P(B)}.$$

На основании этих формул делаем вывод о вероятности произведения двух событий.

Теорема (умножения). *Вероятность произведения двух событий равна*

$$P(AB) = P(A) \cdot P_A(B)$$

или

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

Следствие 1. *Если события A и B независимы, то вероятность их произведения равна произведению вероятностей сомножителей:*

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

Следствие 2. *Вероятность произведения n событий равна*

$$P(A_1A_2...A_n)=P(A_1)\cdot P_{A_1}(A_2)\cdot ...\cdot P_{A_1A_2...A_{n-1}}(A_n).$$

Следствие 3. *Вероятность произведения n независимых событий (каждое событие независимо от любой совокупности остальных) равна*

$$P(A_1A_2...A_n)=P(A_1)\cdot P(A_2)\cdot ...\cdot P(A_n).$$

Пример. В библиотеке на стеллаже в случайном порядке расставлены десять учебников по экономике и пять – по математике. Библиотекарь наудачу по одному берет три учебника. Найти вероятность того, что а) все три учебника по математике; б) хотя бы один из взятых учебников будет по математике.

Решение. Введем события A – все три учебника по математике,

B – хотя бы один из взятых учебников будет по математике,

A_k – учебник № k по математике $(k=1, 2, 3)$.

Тогда $A=A_1\cdot A_2\cdot A_3$, $\overline{B}=\overline{A_1}\cdot\overline{A_2}\cdot\overline{A_3}$.

События A_1, A_2, A_3 зависимы, значит,

$$P(A)=P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3)=\frac{5}{15}\cdot\frac{4}{14}\cdot\frac{3}{13}=\frac{2}{91}.$$

$$P(\overline{B})=P(\overline{A_1})P_{\overline{A_1}}(\overline{A_2})P_{\overline{A_1}\,\overline{A_2}}(\overline{A_3})=\frac{10}{15}\cdot\frac{9}{14}\cdot\frac{8}{13}=\frac{24}{91}.$$

Тогда $P(B)=1-P(\overline{B})=1-\frac{24}{91}=\frac{67}{91}$.

Пусть n – общее число элементарных исходов эксперимента, m – число элементарных исходов, благоприятствующих событию A, k – событию B, s – событию AB, т.е. и событию A, и событию B. Тогда событию $A+B$ благоприятствует $m+k-s$ элементарных исходов.

$$P(A+B)=\frac{m+k-s}{n}=\frac{m}{n}+\frac{k}{n}-\frac{s}{n}=P(A)+P(B)-P(AB).$$

Теорема (сложения). *Вероятность суммы двух событий равна*

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

Следствие 1. *Если события A и B несовместны, то*

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

Следствие 2. *Если события* A_1, A_2, ..., A_n *попарно несовместны, то*

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

Пример. Бросают две игральные кости. Какова вероятность выпадения хотя бы на одной кости четырех очков?

Решение.

Введем события:

A – выпадение четырех очков при бросании первой кости,

B – выпадение четырех очков при бросании второй кости.

Тогда найти требуется вероятность события $A+B$. События A и B совместны. По теореме сложения:

$$P(A+B)=P(A)+P(B)-P(AB)=\frac{1}{6}+\frac{1}{6}-\frac{1}{36}=\frac{11}{36}.$$

Часто эти теоремы применяются вместе, в одной и той же задаче.

Пример. Два стрелка стреляют по одной мишени. Вероятности попадания в мишень при одном выстреле для стрелков соответственно равны $p_1=0,7$ и $p_2=0,8$. Найти вероятность того, что при одном залпе в мишени будет: а) одно попадание; б) не менее одного попадания.

Решение. а) Введем обозначения для событий из задачи:

A – в мишени одно попадание,

B – в мишени не менее одного попадания.

И введем еще два события:

C_1 – в мишень попал первый стрелок,

C_2 – в мишень попал второй стрелок.

Тогда $A=C_1\overline{C_2}+\overline{C_1}C_2$, $B=C_1\overline{C_2}+\overline{C_1}C_2+C_1C_2$.

В обеих формулах случайные события-слагаемые – несовместны, а случайные события-сомножители – независимы, так как вероятность попадания в мишень каждого из стрелков не зависит от результата другого стрелка. Поэтому

$$P(A)=P\left(C_1\overline{C_2}\right)+P\left(\overline{C_1}C_2\right)=P(C_1)\cdot P\left(\overline{C_2}\right)+P\left(\overline{C_1}\right)\cdot P(C_2)=$$
$$=0,7\cdot 0,2+0,3\cdot 0,8=0,38,$$
$$P(B)=P(C_1)\cdot P\left(\overline{C_2}\right)+P\left(\overline{C_1}\right)\cdot P(C_2)+P(C_1)\cdot P(C_2)=$$
$$=0,7\cdot 0,2+0,3\cdot 0,8+0,7\cdot 0,8=0,94.$$

Вероятность события B можно определить также, вычислив сначала вероятность противоположного события $\overline{B}$. Т.к. $\overline{B}=\overline{C_1}\cdot\overline{C_2}$, то

$$P\left(\overline{B}\right)=P\left(\overline{C_1}\right)\cdot P\left(\overline{C_2}\right)=0,3\cdot 0,2=0,06.$$

Тогда $P(B)=1-0,06=0,94$.

Пример. Абонент забыл последнюю цифру номера телефона и набирает ее наудачу. Определить вероятность того, что он наберет нужный номер не более, чем за три попытки.

Решение. Введем случайные события:

A – абонент дозвонился не более, чем за три попытки;

A_i – абонент дозвонился при i-м наборе номера (i = 1, 2, 3).

$$A = A_1 + \overline{A_1}A_2 + \overline{A_1}\,\overline{A_2}A_3,\ P(A) = P(A_1) + P\left(\overline{A_1}A_2\right) + P\left(\overline{A_1}\,\overline{A_2}A_3\right) =$$

$$P(A_1) + P\left(\overline{A_1}\right)P_{\overline{A_1}}(A_2) + P\left(\overline{A_1}\right)P_{\overline{A_1}}\left(\overline{A_2}\right)P_{\overline{A_1}\,\overline{A_2}}(A_3) =$$

$$= \frac{1}{10} + \frac{9}{10}\cdot\frac{1}{9} + \frac{9}{10}\cdot\frac{8}{9}\cdot\frac{1}{8} = \frac{3}{10}.$$

Формула полной вероятности

Предположим, что событие A может наступить только вместе с одним из нескольких попарно несовместных событий $H_1, H_2, ..., H_n$, образующих полную группу. Будем называть события H_i (i = 1, 2, …, n) *гипотезами*.

Для таких событий выполняется условие

$$P(H_1) + P(H_2) + ... + P(H_n) = 1.$$

Событие A можно представить как сумму попарно несовместных событий

$$A = AH_1 + AH_2 + ... + AH_n.$$

Тогда по теореме сложения для несовместных событий:

$$P(A) = P(AH_1) + P(AH_2) + ... + P(AH_n).$$

Применив теорему умножения, получаем формулу

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

которая называется *формулой полной вероятности*.

Пример. Имеются три урны. В первой урне находятся 5 белых и 3 черных шара, во второй – 4 белых и 4 черных шара, а в третьей – 8 белых шаров. Наугад выбирается одна из урн. Из этой урны наугад извлекается шар. Какова вероятность того, что он окажется черным?

Решение. Событие A – извлечен черный шар.

Если было бы известно, из какой урны извлекается шар, то искомую вероятность можно было вычислить по классическому определению вероятности.

Введем предположения (гипотезы) относительно того, какая урна была выбрана для извлечения шара.

Может быть выбрана или первая урна (гипотеза H_1), или вторая (гипотеза H_2), или третья (гипотеза H_3).

Так как имеются одинаковые шансы выбрать любую из урн, то

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

Далее находим вероятности события A при каждой из гипотез:

$$P_{H_1}(A)=\frac{3}{8},\ P_{H_2}(A)=\frac{4}{8}=\frac{1}{2},\ P_{H_3}(A)=\frac{0}{8}=0.$$

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)=$$

$$=\frac{1}{3}\cdot\frac{3}{8}+\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot 0=\frac{1}{8}+\frac{1}{6}=\frac{7}{24}.$$

Пример. Магазин закупает яблоки у трех фермеров в соотношении 1:2:3. Среди яблок первого фермера примерно 80% спелых, второго – 90%, третьего – 70%. Наугад выбирается одно яблоко. Какова вероятность того, что оно окажется спелым?

Решение. Событие A – взято спелое яблоко.

Введем предположения:

H_1 – взятое яблоко выращено первым фермером,

H_2 – взятое яблоко выращено вторым фермером,

H_3 – взятое яблоко выращено третьим фермером.

$$P(H_1)=\frac{1}{6},\ P(H_2)=\frac{2}{6},\ P(H_3)=\frac{3}{6};$$

$$P_{H_1}(A)=0{,}8,\ P_{H_2}(A)=0{,}9,\ P_{H_3}(A)=0{,}7.$$

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)=$$

$$=\frac{1}{6}\cdot 0{,}8+\frac{2}{6}\cdot 0{,}9+\frac{3}{6}\cdot 0{,}7=\frac{1}{6}(0{,}8+1{,}8+2{,}1)=\frac{1}{6}\cdot 4{,}7=\frac{47}{60}.$$

Пример. Электролампы изготавливаются на трех заводах. Первый завод производит 30% общего количества электроламп, второй – 25%, а третий – остальную часть. Продукция первого завода содержит 1% бракованных электроламп, второго – 1,5%, третьего – 2%. В магазин поступает продукция всех трех заводов. Какова вероятность того, что купленная в магазине лампа оказалась бракованной?

Решение. Предположения необходимо ввести относительно того, на каком заводе была изготовлена электролампа. Зная это, мы сможем найти вероятность того, что она бракованная.

Введем обозначения для событий: A – купленная электролампа оказалась бракованной,

H_1 – лампа изготовлена первым заводом,

H_2 – лампа изготовлена вторым заводом,

H_3 – лампа изготовлена третьим заводом.

$$P(H_1)=0{,}30,\ P(H_2)=0{,}25,\ P(H_3)=0{,}45;$$

$$P_{H_1}(A)=0{,}01,\ P_{H_2}(A)=0{,}015,\ P_{H_3}(A)=0{,}02.$$

Искомую вероятность находим по формуле полной вероятности.

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)=$$
$$=0{,}30\cdot 0{,}01+0{,}25\cdot 0{,}015+0{,}45\cdot 0{,}02=0{,}01575.$$

Пример. Из урны, содержащей 3 белых и 2 черных шара, переложено 2 шара в урну, содержащую 4 белых и 4 черных шара. Найти вероятность вынуть после этого из второй урны белый шар.

Решение.

Обозначим через A событие – из второй урны вынут белый шар.

Можно выдвинуть три гипотезы:

H_1 – из первой урны во вторую переложены два белых шара;

H_2 – переложены один белый и один черный шары;

H_3 – переложены два черных шара.

$$P(H_1)=\frac{C_3^2}{C_5^2}=0{,}3,\ P(H_2)=\frac{C_3^1\cdot C_2^1}{C_5^2}=0{,}6,\ P(H_3)=\frac{C_2^2}{C_5^2}=0{,}1;$$
$$P_{H_1}(A)=0{,}6,\ P_{H_2}(A)=0{,}5,\ P_{H_3}(A)=0{,}4.$$

По формуле полной вероятности получаем:

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)=$$
$$=0{,}3\cdot 0{,}6+0{,}6\cdot 0{,}5+0{,}1\cdot 0{,}4=0{,}52.$$

Формула Байеса

Формула Байеса позволяет «переставить причину и следствие»: по известному факту события вычислить вероятность того, что оно было вызвано данной причиной.

Пусть $H_1,\ H_2, ..., H_n$ – полная группа попарно несовместных событий, A – случайное событие.

По теореме умножения

$$P(AH_i)=P(A)P_A(H_i)=P(H_i)P_{H_i}(A) \text{ для } i=1, 2, ..., n.$$

Выразив $P_A(H_i)$, получим формулу Байеса

$$P_A(H_i)=\frac{P(H_i)P_{H_i}(A)}{P(A)},\ i=1, 2, ..., n.$$

Пример. В специализированную больницу поступают в среднем 50% больных с заболеванием K, 30% – с заболеванием L, 20% – с заболеванием M. Вероятность полного излечения болезни K равна 0,7 для болезней L и M эти вероятности соответственно равны 0,8 и 0,9. Больной, поступивший в больницу, был выписан здоровым. Найти вероятность того, что этот больной страдал заболеванием K.

Решение. Введем гипотезы:

H_1 – больной страдал заболеванием K,

H_2 – больной страдал заболеванием L,

H_3 – больной страдал заболеванием M.

Тогда по условию задачи: $P(H_1)=0{,}5$, $P(H_2)=0{,}3$, $P(H_3)=0{,}2$.

Введем событие A – больной, поступивший в больницу, был выписан здоровым. По условию $P_{H_1}(A)=0{,}7$, $P_{H_2}(A)=0{,}8$, $P_{H_3}(A)=0{,}9$.

По формуле полной вероятности получаем:

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)=$$
$$=0{,}5\cdot 0{,}7+0{,}3\cdot 0{,}8+0{,}2\cdot 0{,}9=0{,}77.$$

В задаче требуется вычислить вероятность $P_A(H_1)$.

По формуле Байеса: $P_A(H_1)=\dfrac{P(H_1)P_{H_1}(A)}{P(A)}=\dfrac{0{,}5\cdot 0{,}7}{0{,}77}=\dfrac{5}{11}\approx 0{,}45$.

Пример. Пусть в урне пять шаров и все предположения о количестве белых шаров равновозможные. Из урны наудачу взят шар, он оказался белым. Какое предположение о первоначальном содержании урны наиболее вероятно?

Решение. Введем событие A – наудачу взятый шар оказался белым.

Возможны шесть предположений о первоначальном количестве шаров в урне. Пусть H_i – гипотеза, состоящая в том, что в урне i белых шаров $(i=0, 1, 2, 3, 4, 5)$.

Тогда по условию задачи имеем

$$P(H_0)=P(H_1)=P(H_2)=P(H_3)=P(H_4)=P(H_5)=\frac{1}{6}$$

и $P_{H_i}(A)=\dfrac{i}{5}$ $(i=0, 1, 2, 3, 4, 5)$.

Тогда

$$P(A)=\frac{1}{6}\cdot\frac{0}{5}+\frac{1}{6}\cdot\frac{1}{5}+\frac{1}{6}\cdot\frac{2}{5}+\frac{1}{6}\cdot\frac{3}{5}+\frac{1}{6}\cdot\frac{4}{5}+\frac{1}{6}\cdot\frac{5}{5}=\frac{1}{6}\cdot 3=\frac{1}{2}.$$

По формуле Байеса имеем $P_A(H_i)=\dfrac{P(H_i)P_{H_i}(A)}{P(A)}=\dfrac{\frac{1}{6}\cdot\frac{i}{5}}{\frac{1}{2}}=\dfrac{i}{15}$.

Таким образом, наиболее вероятной является гипотеза H_5.

Пример. Расследуются причины неудачного запуска космической ракеты, о которой можно высказать четыре предположения (гипотезы): H_1, H_2, H_3 или H_4. По данным статистики $P(H_1)=0{,}2$, $P(H_2)=0{,}4$, $P(H_3)=0{,}3$, $P(H_4)=0{,}1$. В ходе расследования обнаружено, что произошла утечка топлива (событие A). Условные вероятности события A

(по данным статистики) равны: $P_{H_1}(A)=0{,}9$, $P_{H_2}(A)=0$, $P_{H_3}(A)=0{,}2$, $P_{H_4}(A)=0{,}3$. Какая из гипотез наиболее вероятна при данных условиях?

Решение. По формуле полной вероятности получаем:

$$P(A)=P(H_1)\cdot P_{H_1}(A)+P(H_2)\cdot P_{H_2}(A)+P(H_3)\cdot P_{H_3}(A)+P(H_4)\cdot P_{H_4}(A)=$$
$$=0{,}2\cdot 0{,}9+0{,}4\cdot 0+0{,}3\cdot 0{,}2+0{,}1\cdot 0{,}3=0{,}18+0+0{,}06+0{,}03=0{,}27.$$

По формуле Байеса $P_A(H_i)=\dfrac{P(H_i)P_{H_i}(A)}{P(A)}$ $(i=1,\,2,\,3,\,4)$.

Наиболее вероятно предположение №1.

Повторные независимые испытания

Если производится несколько испытаний, причем вероятность события A в каждом испытании не зависит от исходов других испытаний, то такие испытания называют *независимыми относительно события A*.

Говорят, что имеет место «схема Бернулли».

Пусть производится три испытания, в каждом из которых событие A может наступить с вероятностью p: $P(A)=p$. Обозначим $P(\overline{A})=1-p=q$.

Найдем, например, вероятность того, что событие A наступило ровно два раза в серии из трех испытаний.

Обозначим: A_1 – событие A наступило в первом испытании, A_2 – во втором, A_3 – в третьем.

Тогда найти нужно $P\left(A_1\cdot A_2\cdot\overline{A_3}+A_1\cdot\overline{A_2}\cdot A_3+\overline{A_1}\cdot A_2\cdot A_3\right)$.

По теоремам сложения и умножения:

$$P\left(A_1\cdot A_2\cdot\overline{A_3}+A_1\cdot\overline{A_2}\cdot A_3+\overline{A_1}\cdot A_2\cdot A_3\right)=P\left(A_1\cdot A_2\cdot\overline{A_3}\right)+$$
$$+P\left(A_1\cdot\overline{A_2}\cdot A_3\right)+P\left(\overline{A_1}\cdot A_2\cdot A_3\right)=$$
$$=P(A_1)P(A_2)P\left(\overline{A_3}\right)+P(A_1)P\left(\overline{A_2}\right)P(A_3)+P\left(\overline{A_1}\right)P(A_2)P(A_3)=$$
$$=p\cdot p\cdot q+p\cdot q\cdot p+q\cdot p\cdot p=3\cdot p^2\cdot q=C_3^2\cdot p^2\cdot q^{3-2}=P_3(2).$$

Количество слагаемых равно числу сочетаний из трех по два, т.е. количеству способов выбора двух событий, которые произошли, из трех множителей A_1, A_2, A_3.

Пусть производится n независимых одинаковых испытаний, в каждом из которых событие A может наступить с одной и той же вероятностью p. Обозначим: $q=1-p$.

Тогда вероятность $P_n(k)$ того, что событие наступило k раз в этой серии испытаний можно вычислить по нескольким формулам.

1. Формула Бернулли

$$P_n(k)=C_n^k p^k q^{n-k}.$$

Вероятность того, что в n испытаниях событие наступит: а) менее k раз, б) более k раз, в) не менее k раз, г) не более k раз, д) хотя бы один раз, находят соответственно по формулам:

а) $P_n(0)+P_n(1)+...+P_n(k-1)$, б) $P_n(k+1)+P_n(k+2)+...+P_n(n)$,

в) $P_n(k)+P_n(k+1)+...+P_n(n)$, г) $P_n(0)+P_n(1)+...+P_n(k)$, д) $1-P_n(0)$.

Пример. Вероятность попадания в цель при одном выстреле равна $\frac{1}{5}$. Производят 10 выстрелов. Найти вероятность того, что попаданий было ровно 6.

Решение. Из условия задачи: $n=10$, $p=\frac{1}{5}$, $q=1-\frac{1}{5}=\frac{4}{5}$, $k=6$.

По формуле Бернулли:

$$P=P_{10}(6)=C_{10}^6\cdot\left(\frac{1}{5}\right)^6\cdot\left(\frac{4}{5}\right)^{10-6}=\frac{10!}{6!\cdot 4!}\cdot\frac{1}{16\,625}\cdot\frac{256}{625}\approx 0{,}052.$$

Пример. Монету бросают 6 раз. Найти вероятность того, что герб выпадет: а) менее двух раз; б) не менее двух раз.

Решение. Вероятность появления герба в одном испытании равна $p=\frac{1}{2}$.

а) Вероятность выпадения герба менее двух раз в шести независимых испытаниях равна:

$$P=P_6(0)+P_6(1)=C_6^0\cdot\left(\frac{1}{2}\right)^0\cdot\left(\frac{1}{2}\right)^6+C_6^1\cdot\left(\frac{1}{2}\right)^1\cdot\left(\frac{1}{2}\right)^5=\frac{7}{64}.$$

б) Вероятность появления герба в шести независимых испытаниях не менее двух раз, находим по формуле:

$$P=1-\left(P_6(0)+P_6(1)\right)=1-\frac{7}{64}=\frac{57}{64}.$$

Пример. Два равносильных шахматиста играют в шахматы. Что вероятнее: выиграть две партии из четырех или три партии из шести (ничьи во внимание не принимаются)?

Решение. Играют равносильные шахматисты, поэтому вероятность выигрыша каждого равна $p=\frac{1}{2}$. Следовательно, вероятность проигрыша q равна $\frac{1}{2}$.

Вероятность того, что две партии из четырех будут выиграны:

$$P_4(2)=C_4^2p^2q^2=\frac{4\cdot 3}{1\cdot 2}\cdot\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)^2=\frac{6}{16}.$$

Вероятность того, что будут выиграны три партии из шести:

$$P_6(3)=C_6^3p^3q^3=\frac{6\cdot 5\cdot 4}{1\cdot 2\cdot 3}\cdot\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^3=\frac{5}{16}.$$

Так как $P_4(2)>P_6(3)$, то вероятнее выиграть две партии из четырех, чем три из шести.

Пример. Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить ровно 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение. Из постановки задачи следует, что последнее купленное изделие имеет особый знак.

Следовательно, из предыдущих девяти эти знаки имели 4 изделия.

Найдем вероятность этого по формуле Бернулли:

$$P_9(4)=C_9^4(0{,}05)^4(0{,}95)^5=0{,}0006092.$$

Правило.

Для вычисления вероятности $P_n(k)$ будем применять:

1) формулу Бернулли, если число испытаний $n\le 30$,

2) формулу Пуассона, если $n>30$ и $\lambda=np<10$;

3) локальную теорему Муавра–Лапласа в остальных случаях.

2. Формула Пуассона

$$P_n(k)\approx\frac{\lambda^k e^{-\lambda}}{k!},\text{ где } \lambda=np.$$

Пример. Завод отправил на базу 5 000 доброкачественных изделий. Вероятность того, что в пути изделие повредится, равно 0,0002. Найти вероятность того, что среди доставленных на базу изделий будет ровно три негодных.

Решение. По условию,

$$n=5\,000,\ p=0{,}0002,\ k=3,\ \lambda=np=5\,000\cdot 0{,}0002=1<10.$$

По формуле Пуассона, искомая вероятность приближенно равна

$$P_{5000}(3)\approx\frac{1^3\cdot e^{-1}}{3!}=\frac{1}{6e}\approx 0{,}06.$$

Пример. В страховой компании застраховано 10 000 клиентов одного социального положения. Вероятность наступления страхового случая равна 0,0006. Первого января клиент вносит 12$. Если в течение года страхо-

вой случай наступит, то клиент получит 1 000$. Найти вероятность того, что компания получит прибыль не менее 40 000$.

Решение. 1) 12$·10 000 = 120 000$ компания получила первого января.

2) Компания может потерять: 120 000$ – 40 000$ = 80 000$.

3) Т.е. может наступить не более $\frac{80\,000\$}{1\,000\$} = 80$ страховых случаев.

4) Вероятность того, что в течение года наступит не более 80 страховых случаев.

По условию,

$$n = 10000,\ p = 0{,}0006,\ k \le 80,\ \lambda = np = 10000 \cdot 0{,}0006 = 6 < 10.$$

$$P_{10000}(0) + P_{10000}(1) + P_{10000}(2) + \ldots + P_{10000}(80) \approx$$

$$\approx \frac{6^0}{0!}e^{-6} + \frac{6^1}{1!}e^{-6} + \frac{6^2}{2!}e^{-6} + \ldots + \frac{6^{80}}{80!}e^{-6} \approx 1.$$

3. Локальная теорема Муавра–Лапласа

$$P_n(k) \approx \frac{1}{\sqrt{npq}}\varphi(x_0), \text{ где } x_0 = \frac{k-np}{\sqrt{npq}},\ \varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Значения функции $\varphi(x)$ для $0 \le x < 4$ даны в таблице. Эта функция обладает свойствами:

$$1)\ \varphi(-x) = \varphi(x) \quad \text{и} \quad 2)\ \varphi(x) \approx 0 \text{ при } x \ge 4.$$

Например, $\varphi(-0{,}25) = \varphi(0{,}25) \approx 0{,}3867$, $\varphi(1{,}03) \approx 0{,}2347$,

$\varphi(5{,}39) \approx 0$, $\varphi(-3{,}24) \approx 0{,}0021$, $\varphi(0{,}05) \approx 0{,}3984$,

$\varphi(-3{,}7) = \varphi(3{,}70) \approx 0{,}0002$.

Пример. Вероятность рождения мальчика равна 0,515. Найти вероятность того, что среди 100 новорожденных мальчиков будет ровно половина.

Решение. По условию, $n = 100$, $p = 0{,}515$, $q = 1 - 0{,}515 = 0{,}485$, $k = 50$.

Тогда $x_0 = \frac{k-np}{\sqrt{npq}} = \frac{50 - 100 \cdot 0{,}515}{\sqrt{100 \cdot 0{,}515 \cdot 0{,}485}} = \frac{50 - 51{,}5}{\sqrt{24{,}9775}} \approx -\frac{1{,}5}{4{,}998} \approx -0{,}30$,

$\varphi(x_0) = \varphi(-0{,}30) = \varphi(0{,}30) = 0{,}3814$ (по таблице).

По локальной теореме Муавра–Лапласа получаем:

$$P_{100}(50) \approx \frac{1}{\sqrt{100 \cdot 0{,}515 \cdot 0{,}485}} \cdot 0{,}3814 \approx 0{,}076.$$

Пример. Найти вероятность того, что событие A наступит ровно 80 раз в 400 испытаниях, если вероятность появления этого события в каждом испытании равна 0,2.

Решение. По условию, $n = 400$, $p = 0{,}2$, $q = 1 - 0{,}2 = 0{,}8$, $k = 80$.

Тогда $x_0 = \dfrac{80 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 0$, $\varphi(x_0) = \varphi(0) = 0,3989$ (по таблице).

По локальной теореме Муавра–Лапласа получаем:

$$P_{400}(80) \approx \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \cdot 0,3989 = \frac{1}{8} \cdot 0,3989 \approx 0,04986.$$

4. Интегральная теорема Муавра–Лапласа

Вероятность того, что событие наступит в пределах от k_1 до k_2 раз, можно вычислить, используя интегральную теорему Муавра–Лапласа:

$$P_n(k_1, k_2) \approx \Phi(x_2) - \Phi(x_1),$$

где $x_1 = \dfrac{k_1 - np}{\sqrt{npq}}$, $x_2 = \dfrac{k_2 - np}{\sqrt{npq}}$, $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа.

Значения функции $\Phi(x)$ для $0 \le x \le 4$ даны в таблице.

Эта функция обладает свойствами:

1) $\Phi(-x) = -\Phi(x)$; 2) $\Phi(x) \approx 0,5$ при $x > 4$.

Пример. Вероятность того, что деталь не прошла проверку ОТК, равна 0,2. Найти вероятность того, что среди 400 случайно отобранных деталей некачественных окажется от 70 до 100 деталей.

Решение. По условию,

$$n = 400,\ k_1 = 70,\ k_2 = 100,\ p = 0,2,\ q = 0,8.$$

Воспользуемся интегральной теоремой Муавра–Лапласа:

$$P_{400}(70,\ 100) \approx \Phi(x_2) - \Phi(x_1), \text{ где}$$

$$x_1 = \frac{k_1 - np}{\sqrt{npq}} = \frac{70 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = -1,25;\ x_2 = \frac{k_2 - np}{\sqrt{npq}} = \frac{100 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 2,5.$$

Таким образом,

$$P_{400}(70,100) \approx \Phi(2,5) - \Phi(-1,25) = \Phi(2,5) + \Phi(1,25).$$

По таблице: $\Phi(2,50) = 0,4938$, $\Phi(1,25) = 0,3944$.

Искомая вероятность $P_{400}(70,100) = 0,4938 + 0,3944 = 0,8882$.

5. Наивероятнейшее число наступлений события

Число k_0 наступления события в независимых испытаниях, в каждом из которых вероятность появления события равна p, называется *наивероятнейшим*, если вероятность того, что событие наступит в этих испытаниях k_0 раз, не меньше вероятности остальных возможных исходов испытаний.

Известно, что из вероятностей $P_n(0), P_n(1), P_n(2),...,P_n(n)$ одна или две соседние больше, чем остальные.

Должны выполняться неравенства

$$\begin{cases} P_n(k_0) \geq P_n(k_0 - 1), \\ P_n(k_0) \geq P_n(k_0 + 1). \end{cases}$$

Преобразуем первое неравенство:

$$C_n^{k_0} p^{k_0} q^{n-k_0} \geq C_n^{k_0-1} p^{k_0-1} q^{n-k_0+1},$$

$$\frac{n!}{k_0!(n-k_0)!} p \geq \frac{n!}{(k_0-1)!(n-k_0+1)!} q, \ \frac{1}{k_0} p \geq \frac{1}{(n-k_0+1)} q.$$

Т.к. все множители положительные, получим $(n-k_0+1)p \geq k_0 q$,

$$k_0 q + k_0 p \leq np + p, \ k_0 \leq np + p.$$

Преобразуем второе неравенство:

$$C_n^{k_0} p^{k_0} q^{n-k_0} \geq C_n^{k_0+1} p^{k_0+1} q^{n-k_0-1},$$

$$\frac{n!}{k_0!(n-k_0)!} q \geq \frac{n!}{(k_0+1)!(n-k_0-1)!} p, \ \frac{1}{(n-k_0)} q \geq \frac{1}{(k_0+1)} p.$$

Т.к. все множители положительные, получим $(k_0+1)q \geq (n-k_0)p$,

$$k_0 q + k_0 p \geq np - q, \ k_0 \geq np - q.$$

Наивероятнейшее число k_0 определяется из двойного неравенства

$$np - q \leq k_0 \leq np + p.$$

Замечание. Длина промежутка, которому принадлежит число k_0, равна $(np+p)-(np-q) = p+q = 1$. Если число $np-q$ целое, то k_0 принимает два значения.

Пример. При автоматической наводке орудия вероятность попадания по быстро движущейся цели равна 0,9. Найти наивероятнейшее число попаданий при 50 выстрелах.

Решение. Здесь $n = 50$, $p = 0{,}9$, $q = 0{,}1$.

Поэтому $50 \cdot 0{,}9 - 0{,}1 \leq k_0 \leq 50 \cdot 0{,}9 + 0{,}9$ или $44{,}9 \leq k_0 \leq 45{,}9$.

Следовательно, $k_0 = 45$.

Пример. Товаровед осматривает 24 образца товаров. Вероятность того, что каждый из образцов будет признан годным к продаже, равна 0,6. Найти наивероятнейшее число образцов, которые товаровед признает годными к продаже.

Решение. По условию, $n = 24$, $p = 0,6$, $q = 0,4$. Подставляя данные задачи, получим:

$$24 \cdot 0,6 - 0,4 \le k_0 \le 24 \cdot 0,6 + 0,6, \text{ или } 14 \le k_0 \le 15.$$

Наивероятнейших значений два: $k_0^{(1)} = 14$ и $k_0^{(2)} = 15$.

Пример. Чему равна вероятность p наступления события в каждом из 49 независимых испытаний, если наивероятнейшее число наступлений события в этих испытаниях равно 30?

Решение. По условию задачи, $n = 49, k_0 = 30$. Воспользуемся двойным неравенством $np - q \le k_0 \le np + p$. Подставляя данные задачи, получим систему неравенств для определения неизвестной вероятности p:

$$\begin{cases} 49p + p \ge 30, \\ 49p - (1 - p) \le 30. \end{cases}$$

Из первого неравенства: $p \ge 0,6$. Из второго неравенства: $p \le 0,62$.

Итак, искомая вероятность должна удовлетворять двойному неравенству: $0,6 \le p \le 0,62$.

Контрольные вопросы

1. Что изучает теория вероятностей?
2. Какие эксперименты рассматриваются в теории вероятностей?
3. Что называется случайным событием?
4. Какое событие называется достоверным?
5. Какое событие называется невозможным?
6. Какие события называются равновозможными?
7. Какие события называются несовместными?
8. Какие события образуют полную группу?
9. Что такое вероятность события?
10. Каким условиям должен удовлетворять эксперимент, чтобы можно было применить классическое определение вероятности?
11. Приведите формулу для нахождения классической вероятности события.
12. Приведите формулу для нахождения геометрической вероятности события.
13. Приведите формулу для нахождения статистической вероятности события.
14. Какое событие называется противоположным к событию A?
15. Как найти вероятность наступления события хотя бы один раз?
16. Какое событие называется суммой событий?
17. Приведите формулировки теоремы сложения и ее следствий.
18. Какое событие называется произведением событий?

19. Какие события называются независимыми?

20. Что такое условная вероятность события A при условии наступления события B?

21. Приведите формулировки теоремы умножения и ее следствий.

22. Какая группа событий называется гипотезами?

23. Приведите формулу полной вероятности.

24. Приведите формулу Байеса.

25. Какие испытания называются независимыми относительно события A?

26. Приведите формулу Бернулли.

27. Приведите формулу Пуассона.

28. Приведите формулу из локальной теоремы Муавра–Лапласа.

29. Как выбрать формулу для нахождения вероятности наступления события A ровно k в серии из n независимых испытаний из трех формул?

30. Приведите формулу из интегральной теоремы Муавра–Лапласа.

31. Как найти наивероятнейшее число наступлений события A?

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Основные понятия

Случайной величиной называют числовую величину, которая в результате испытания принимает одно возможное числовое значение, наперед неизвестное и зависящее от случайных обстоятельств.

Например,

1) расход электроэнергии в определенной квартире в течение данного месяца: $[0; +\infty)$;

2) количество попаданий в мишень при десяти выстрелах: 0, 1, 2, …, 10;

3) количество жилых домов, сданных в эксплуатацию с начала января 2000-го года в г. Челябинске на начало каждого месяца: 0, 1, 2, …;

4) расстояние от точки попадания до центра мишени: $[0; R]$, где R – радиус мишени.

Дискретной называют случайную величину, которая принимает отдельные, изолированные возможные значения.

Законом распределения случайной величины называют соответствие между возможными значениями случайной величины и их вероятностями.

Дискретные случайные величины

Для дискретной случайной величины закон распределения можно задать таблично (с помощью ряда распределения) или графически (с помощью многоугольника распределения).

Ряд распределения – это таблица, в первой строке которой перечисляются все возможные значения случайной величины x_k в порядке возраста-

ния, а во второй – вероятности этих значений $p_k = P(X = x_k)$ $(k = 1, 2, ..., n)$ (табл. 3).

Таблица 3

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Многоугольник распределения – это ломаная, звенья которой последовательно соединяют точки с координатами (x_k, p_k), $k = 1, 2, ..., n$ (рис. 3).

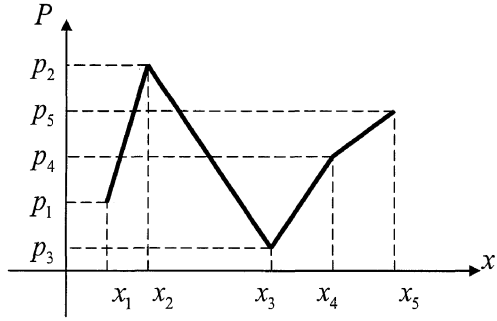


Рис. 3

Замечание. События $(X = x_1)$, $(X = x_2)$, …, $(X = x_n)$ образуют полную группу и попарно несовместны. Значит, сумма их вероятностей равна 1:

$$\sum_{k=1}^{n} P(X = x_k) = \sum_{k=1}^{n} p_k = 1.$$

Пример. В лотерее разыгрываются автомобиль стоимостью 5000 ден.ед., 4 телевизора стоимостью по 250 ден.ед. каждый и 5 телефонов стоимостью по 200 ден.ед. Всего продается 1000 билетов по 7 ден.ед. Составить закон распределения чистого выигрыша купившим один билет участником лотереи.

Решение. Чистый выигрыш получается как разность стоимости полученного приза и стоимости купленного билета. Возможные значения чистого выигрыша: 0 – 7, 200 – 7, 250 – 7 и 5000 – 7. Вероятности этих значений вычислим по классическому определению и составим ряд распределения (табл. 4).

$$P(X = -7) = \frac{990}{1\,000} = 0{,}99,\ P(X = 193) = \frac{5}{1\,000} = 0{,}005,$$

$$P(X=243)=\frac{4}{1\,000}=0{,}004,\ P(X=4\,993)=\frac{1}{1\,000}=0{,}001.$$

Таблица 4

X	–7	193	243	4 993
P	0,99	0,005	0,004	0,001

Многоугольник распределения изображен на рис. 4.

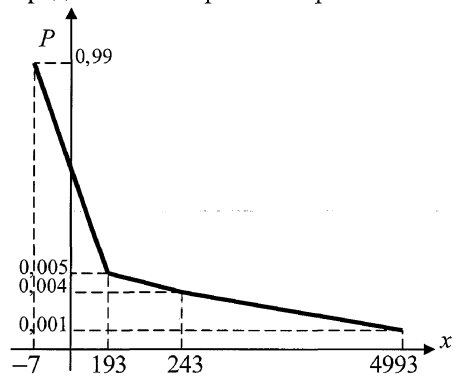


Рис. 4

Пример. Составить закон распределения числа появлений события A в трех независимых испытаниях, если вероятность появления события в каждом испытании равна 0,6.

Решение. Вероятность появления события в каждом испытании равна 0,6, следовательно, вероятность не появления события в каждом испытании $q=1-0{,}6=0{,}4$.

В трех независимых испытаниях событие A может появиться либо 3 раза, либо 2 раза, либо 1 раз, либо совсем не появиться. Таким образом, возможные значения X таковы: $x_1=0$, $x_2=1$, $x_3=2, x_4=3$. Найдем вероятности возможных значений X и составим ряд распределения (табл. 5).

$$P(X=0)=P_3(0)=C_3^0p^0q^3=(0{,}4)^3=0{,}064,$$
$$P(X=1)=P_3(1)=C_3^1p^1q^2=3\cdot 0{,}6\cdot(0{,}4)^2=0{,}288,$$
$$P(X=2)=P_3(2)=C_3^2p^2q^1=3\cdot(0{,}6)^2\cdot 0{,}4=0{,}432,$$
$$P(X=3)=P_3(3)=C_3^3p^3q^0=(0{,}6)^3=0{,}216.$$

Таблица 5

X	0	1	2	3
P	0,064	0,288	0,432	0,216

Пример. В партии из 10 деталей имеется 8 стандартных. Наудачу отобраны две детали. Составить закон распределения числа стандартных деталей среди отобранных.

Решение. Случайная величина X – число стандартных деталей среди отобранных – имеет следующие возможные значения: $x_1 = 0, x_2 = 1$, $x_3 = 2$.

Найдем вероятности возможных значений X и составим ряд распределения (табл. 6).

$$P(X=0)=\frac{C_8^0 \cdot C_2^2}{C_{10}^2}=\frac{1}{45};\ P(X=1)=\frac{C_8^1 \cdot C_2^1}{C_{10}^2}=\frac{16}{45};\ P(X=2)=\frac{C_8^2 \cdot C_2^0}{C_{10}^2}=\frac{28}{45}.$$

Таблица 6

X	0	1	2
P	$\frac{1}{45}$	$\frac{16}{45}$	$\frac{28}{45}$

Контроль: $\frac{1}{45}+\frac{16}{45}+\frac{28}{45}=1$.

Пример. Из орудия ведется стрельба по цели до первого попадания. Вероятность попадания в цель при одном выстреле равна 0,7. Всего имеется три снаряда. Составить закон распределения случайной величины X – числа израсходованных снарядов.

Решение. Пусть события A_i – попадание в цель при i-м выстреле.

Орудие израсходует один снаряд $(X=1)$, если оно попадет в цель при первом выстреле:

$$P(X=1)=P(A_1)=0,7.$$

Орудие израсходует два снаряда $(X=2)$, если при первом выстреле оно промахнется, а при втором выстреле попадет в цель:

$$P(X=2)=P(\bar{A}_1 A_2)=P(\bar{A}_1)P(A_2)=0,3\cdot 0,7=0,21.$$

Орудие израсходует три снаряда $(X=3)$, если при первом и втором выстрелах оно промахнется:

$$P(X=3)=P(\bar{A}_1\bar{A}_2)=P(\bar{A}_1)P(\bar{A}_2)=0,3\cdot 0,3=0,09.$$

Закон распределения дискретной случайной величины X – числа снарядов, израсходованных орудием – в табл. 7.

Таблица 7

X	1	2	3
P	0,7	0,21	0,09

Пример. Вероятность появления события A в одном опыте равна 0,4. Производится ряд независимых опытов, которые продолжаются до первого появления события A, после чего опыты прекращаются. Случайная величина X – число произведенных опытов. Построить ряд распределения величины X.

Решение. Возможные значения величины X: 1, 2, 3, … (теоретически они ничем не ограничены).

Для того, чтобы величина X приняла значение 1, необходимо, чтобы событие A произошло в первом же опыте; вероятность этого равна 0,4.

Для того, чтобы величина X приняла значение 2, нужно, чтобы в первом опыте событие A не появилось, а во втором – появилось; вероятность этого равна $0{,}4 \cdot 0{,}6$ и т.д.

Ряд распределения величины X имеет следующий вид (табл. 8).

Таблица 8

X	1	2	3	…	k	…
P	0,4	$0{,}4 \cdot 0{,}6$	$0{,}4 \cdot (0{,}6)^2$	…	$0{,}4 \cdot (0{,}6)^{k-1}$	…

Математические операции над дискретными случайными величинами

Пусть известен закон распределения случайной величины X (табл. 9).

Таблица 9

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Над ней можно производить следующие действия.

1) Умножение случайной величины на число.

Произведением случайной величины X на число α называется случайная величина αX, значения которой равны αx_k, $k = 1, 2, ..., n$, а соответствующие вероятности равны p_k.

2) Возведение случайной величины в натуральную степень.

Пусть $m \in \mathbf{N}$, тогда случайная величина X^m принимает значения x_k^m, $k = 1, 2, ..., n$, с вероятностями p_k.

Пример. Дан закон распределения случайной величины X (табл. 10).

Таблица 10

X	–2	1	2	4
P	0,4	0,1	0,2	0,3

Законы распределения случайных величин $3X$, $-3X$, X^3 и X^2 приведены в табл. 11–14.

Таблица 11

$3X$	–6	3	6	12
P	0,4	0,1	0,2	0,3

Таблица 12

$-3X$	–12	–6	–3	6
P	0,3	0,2	0,1	0,4

Таблица 13

X^3	–8	1	8	64
P	0,4	0,1	0,2	0,3

Таблица 14

X^2	1	4	16
P	0,1	0,6	0,3

$$P\left(X^2=1\right)=P\left(X=1\right)=0{,}1,$$

$$P\left(X^2=4\right)=P\left(X=2\right)+P\left(X=-2\right)=0{,}4+0{,}2=0{,}6,$$

$$P\left(X^2=16\right)=P\left(X=4\right)=0{,}3.$$

Пусть известны законы распределения двух случайных величин X (табл. 15) и Y (табл. 16).

Таблица 15

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Таблица 16

Y	y_1	y_2	…	y_m
P	p'_1	p'_2	…	p'_m

Случайные величины X и Y называются *независимыми*, если закон распределения одной из них не меняется от того, какие возможные значения приняла другая.

Это означает независимость событий $\left(X=x_i\right)$ и $\left(Y=y_j\right)$ при всех $i=1, 2, ..., n$, $j=1, 2, ..., m$.

Над случайными величинами X и Y можно производить следующие действия.

3) Сложение. Случайная величина $X+Y$ принимает все возможные значения x_i+y_j, где $i=1, 2, ..., n$, $j=1, 2, ..., m$, с вероятностями $P\left(X+Y=x_i+y_j\right)=p_i \cdot p'_j$.

4) Вычитание. Случайная величина $X-Y$ принимает все возможные значения x_i-y_j, где $i=1, 2, ..., n$, $j=1, 2, ..., m$, с вероятностями $P\left(X-Y=x_i-y_j\right)=p_i \cdot p'_j$.

5) Умножение. Случайная величина $X \cdot Y$ принимает все возможные значения $x_i \cdot y_j$, где $i=1, 2, ..., n$, $j=1, 2, ..., m$, с вероятностями $P\left(X \cdot Y=x_i \cdot y_j\right)=p_i \cdot p'_j$.

Пример. Даны законы распределения двух случайных величин (табл. 17, 18)

Таблица 17

X	–1	0	2
P	0,1	0,2	0,7

Таблица 18

Y	0	1
P	0,6	0,4

Запишем данные в виде таблицы (табл. 19).

Таблица 19

Y \ X	–1	0	2
0	0,06	0,12	0,42
1	0,04	0,08	0,28

Тогда законы распределения суммы и произведения (табл. 20, 21):

Таблица 20

$X+Y$	–1	0	1	2	3
P	0,06	0,16	0,08	0,42	0,28

Таблица 21

$X \cdot Y$	–1	0	2
P	0,04	0,68	0,28

Вероятность попадания случайной величины на заданный участок

При решении практических задач, связанных со случайными величинами, часто оказывается необходимым вычислить вероятность того, что случайная величина примет значение, заключенное в некоторых пределах, на-

пример, от α до β. Это событие мы будем называть «попаданием случайной величины X на участок от α до β».

Например, для случайной величины, заданной рядом распределения (табл. 22).

Таблица 22

X	–1	0	2	5	6
P	0,1	0,2	0,1	0,4	0,2

$$P(1 \le X \le 4) = P(X=2) = 0{,}1,\ P(-1 < X < 5) = P(X=0) + P(X=2) = 0{,}3,$$
$$P(-1 \le X \le 5) = P(X=-1) + P(X=0) + P(X=2) + P(X=5) = 0{,}8,$$
$$P(X<2) = P(X=-1) + P(X=0) = 0{,}3.$$

Функция распределения случайной величины

Функцией распределения случайной величины X называется функция $F(x) = P(X < x)$.

Т.е. функция распределения в точке x равна вероятности попадания X в интервал $(-\infty; x)$.

Свойства функции распределения

1) $0 \le F(x) \le 1$ для всех $x \in \mathbb{R}$.

Т.к. $F(x) = P(X < x)$, а $0 \le P(X < x) \le 1$.

2) $\lim\limits_{x \to +\infty} F(x) = 1$.

Т.к. $\lim\limits_{x \to +\infty} F(x) = \lim\limits_{x \to +\infty} P(X < x) = P(\text{достоверное событие}) = 1$.

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

Т.к. $\lim\limits_{x \to -\infty} F(x) = \lim\limits_{x \to -\infty} P(X < x) = P(\text{невозможное событие}) = 0$.

4) $F(x)$ – неубывающая.

Т.к. если $x_1 > x_2$, то $P(X < x_1) \ge P(X < x_2)$, т.е. $F(x_1) \ge F(x_2)$.

Пример. Дан ряд распределения X (табл. 23).

Таблица 23

X	–2	1	2	4
P	0,4	0,1	0,2	0,3

Составить функцию распределения X и построить ее график.

Решение. Разобьем числовую ось на части значениями случайной величины (рис. 5).

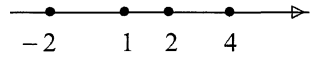


Рис. 5

1) При $x \le -2$ (рис. 6) $F(x) = P(X < x) = 0$.

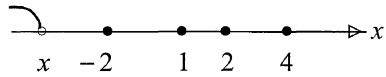


Рис. 6

2) При $-2 < x \le 1$ (рис. 7) $F(x) = P(X < x) = P(X = -2) = 0{,}4$.

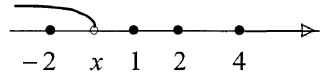


Рис. 7

3) При $1 < x \le 2$ (рис. 8)

$$F(x) = P(X < x) = P(X = -2) + P(X = 1) = 0{,}4 + 0{,}1 = 0{,}5.$$

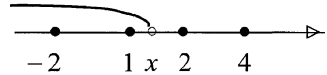


Рис. 8

4) При $2 < x \le 4$ (рис. 9)

$$F(x) = P(X < x) = P(X = -2) + P(X = 1) + P(X = 2) = 0{,}4 + 0{,}1 + 0{,}2 = 0{,}7.$$

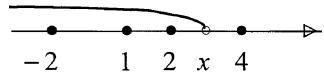


Рис. 9

5) При $x > 4$ (рис. 10)

$$F(x) = P(X < x) = P(X = -2) + P(X = 1) + P(X = 2) + P(X = 4) =$$
$$= 0{,}4 + 0{,}1 + 0{,}2 + 0{,}3 = 1.$$

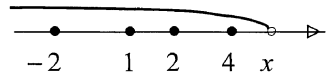


Рис. 10

Итак, $F(x)=\begin{cases} 0 \text{ при } x \le -2, \\ 0,4 \text{ при } -2 < x \le 1, \\ 0,5 \text{ при } 1 < x \le 2, \\ 0,7 \text{ при } 2 < x \le 4, \\ 1 \text{ при } x > 4. \end{cases}$

График изображен на рис. 11.

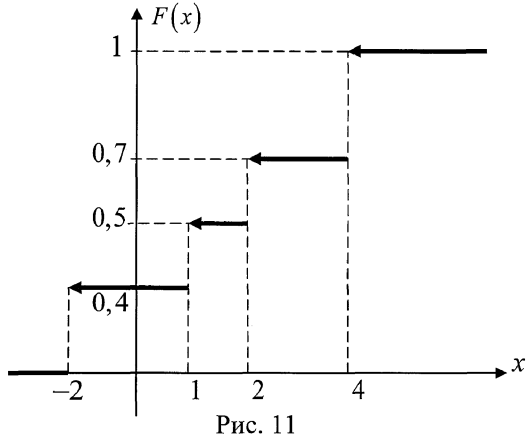


Рис. 11

Числовые характеристики дискретных случайных величин

1. Математическое ожидание $M(X)$ случайной величины X – это число, равное сумме произведений всех значений случайной величины на их вероятности.

Если дан ряд распределения X (табл. 24)

Таблица 24

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

то $M(X)=\sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + ... + x_n p_n$.

Математическое ожидание характеризует среднее, ориентировочное значение случайной величины, около которого группируются все возможные значения случайной величины.

Действительно, пусть при проведении n независимых испытаний случайная величина X может принимать m_1 раз значение x_1, m_2 раз значение x_2 …, m_n раз значение x_n. Тогда сумма всех значений равна

$$\underbrace{x_1+x_1+\ldots+x_1}_{m_1\text{ раз}}+\underbrace{x_2+x_2+\ldots+x_2}_{m_2\text{ раз}}+\ldots+\underbrace{x_n+x_n+\ldots+x_n}_{m_n\text{ раз}}=$$
$$=m_1x_1+m_2x_2+\ldots+m_nx_n.$$

А их среднее арифметическое:

$$\frac{m_1x_1+m_2x_2+\ldots+m_nx_n}{n}=\frac{m_1}{n}x_1+\frac{m_2}{n}x_2+\ldots+\frac{m_n}{n}x_n=$$
$$=p_1x_1+p_2x_2+\ldots+p_nx_n=M(X).$$

Пример. Даны законы распределения двух случайных величин (табл. 25, 26). Найти их математические ожидания.

Таблица 25

X_1	–2	1	2	4
P	0,4	0,1	0,2	0,3

$$M(X_1)=-2\cdot 0{,}4+1\cdot 0{,}1+2\cdot 0{,}2+4\cdot 0{,}3=0{,}9.$$

Таблица 26

X_2	–2	1	2	4
P	0,1	0,1	0,1	0,7

$$M(X_2)=-2\cdot 0{,}1+1\cdot 0{,}1+2\cdot 0{,}1+4\cdot 0{,}7=2{,}9.$$

Свойства математического ожидания

1) $M(C)=C$, где C – произвольное число.

Действительно, рассмотрим случайную величину с рядом распределения (табл. 27)

Таблица 27

X	C
P	1

Ее математическое ожидание равно $M(X)=C\cdot 1=C$.

2) $M(kX)=kM(X)$.

Дан ряд распределения случайной величины X (табл. 28).

Таблица 28

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Тогда случайная величина kX принимает значения kx_1, kx_2, ..., kx_n с вероятностями p_1, p_2, ..., p_n, соответственно.

Тогда $M(kX)=\sum_{i=1}^{n}(k\cdot x_i)p_i=k\cdot\sum_{i=1}^{n}x_ip_i=k\cdot M(X)$.

3) $M(X\pm Y)=M(X)\pm M(Y)$.

Даны законы распределения случайных величин X и Y (табл. 29, 30).

Таблица 29

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Таблица 30

Y	y_1	y_2	...	y_m
P	p_1'	p_2'	...	p_m'

Тогда математическое ожидание их суммы равно

$$M(X+Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}(x_i+y_j)p_ip_j'=\sum_{j=1}^{m}p_j'\sum_{i=1}^{n}x_ip_i+\sum_{i=1}^{n}p_i\sum_{j=1}^{m}y_jp_j'=$$

$$=M(X)\sum_{j=1}^{m}p_j'+M(Y)\sum_{i=1}^{n}p_i=M(X)+M(Y).$$

4) $M(X\cdot Y)=M(X)\cdot M(Y)$, если X и Y – независимые случайные величины.

5) $M(X\pm C)=M(X)\pm C$.

Пусть X распределена по закону (табл. 31).

Таблица 31

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Тогда закон распределения $X+C$ имеет следующий вид (табл. 32).

Таблица 32

$X+C$	x_1+C	x_2+C	...	x_n+C
P	p_1	p_2	...	p_n

Тогда

$$M(X+C)=\sum_{i=1}^{n}(x_i+C)p_i=\sum_{i=1}^{n}x_ip_i+\sum_{i=1}^{n}Cp_i=M(X)+C\sum_{i=1}^{n}p_i=M(X)+C.$$

6) $M(X-M(X))=0$.

Пример. Пусть $M(X)=5$, $M(Y)=-4$. Тогда

$$M(2X+3Y-8)=2M(X)+3M(Y)-8=10-12-8=-10.$$

2. Мода $Mo(X)$ случайной величины X – это ее наиболее вероятное значение.

Дан ряд распределения X (табл. 33).

Таблица 33

X	–2	1	2	4
P	0,2	0,4	0,3	0,1

Тогда $Mo(X)=1$.

Если таких значений более одного, распределение называется «полимодальным».

В общем случае мода и математическое ожидание случайной величины не совпадают. В частном случае, когда распределение является симметричным и модальным (т.е. имеет моду) и существует математическое ожидание, то оно совпадает с модой и центром симметрии распределения.

3. Медиана $Me(X)$ случайной величины X – это такое ее значение, что $P(X<Me(X))=P(X>Me(X))$.

Дан ряд распределения X (табл. 34).

Таблица 34

X	1	2	4	5
P	0,4	0,1	0,1	0,3

Тогда $Me(X)=2$, т.к. $P(X<2)=0,4=P(X>2)$.

4. Дисперсия $D(X)$ случайной величины X – это число, равное математическому ожиданию квадрата ее отклонения от математического ожидания:

$$D(X)=M(X-M(X))^2.$$

Для дискретной случайной величины дисперсия вычисляется по формуле:

$$D(X)=\sum_{i=1}^{n}(x_i-M(X))^2 p_i.$$

Дисперсия характеризует рассеяние (разброс) значений случайной величины относительно ее среднего значения.

Дисперсия имеет размерность квадрата случайной величины.

Свойства дисперсии

1) $D(C)=0$, где C – произвольное число.

2) $D(kX)=k^2D(X)$.

3) $D(X\pm Y)=D(X)+D(Y)$, если X и Y – независимые случайные величины.

4) $D(X)=M\left(X^2\right)-\left(M(X)\right)^2$, где

$$M\left(X^2\right)=\sum_{i=1}^{n}x_i^2p_i=x_1^2p_1+x_2^2p_2+...+x_n^2p_n.$$

Пример. Дан ряд распределения (табл. 35).

Таблица 35

X	–2	1	2	4
P	0,4	0,1	0,2	0,3

Тогда

$$M(X)=0{,}9,\ M\left(X^2\right)=(-2)^2\cdot 0{,}4+1^2\cdot 0{,}1+2^2\cdot 0{,}2+4^2\cdot 0{,}3=7{,}3,$$

$$D(X)=M\left(X^2\right)-\left(M(X)\right)^2=7{,}3-0{,}9^2=6{,}49.$$

Пример. Пусть $D(X)=5$, $D(Y)=4$. Тогда

$$D(2X-3Y-8)=4D(X)+9D(Y)=20+36=56.$$

5. Среднее квадратическое (стандартное) отклонение $\sigma(X)$ случайной величины X – это число, равное корню квадратному из ее дисперсии: $\sigma(X)=\sqrt{D(X)}$.

Интерпретация числовых характеристик в финансовом анализе

Пусть известно распределение доходности X некоторого актива (например, акции), т.е. известны значения доходности x_i и соответствующие им вероятности p_i за рассматриваемый промежуток времени.

Тогда математическое ожидание выражает *среднюю (прогнозируемую) доходность актива*, а дисперсия или среднее квадратическое отклонение – меру отклонения доходности от ожидаемого среднего значения, т.е. *риск* данного актива.

Основные законы распределения дискретных случайных величин

1. Биномиальный закон распределения

Дискретная случайная величина X имеет ***биномиальный закон распределения***, если она принимает значения $0, 1, 2, ..., n$ с вероятностями $P(X=k)=C_n^kp^kq^{n-k}$, где $0<p<1$, $q=1-p$, $k=0, 1, 2, ..., n$.

Биномиальный закон распределения представляет собой закон распределения числа наступлений события A в n независимых испытаниях, в каждом из которых оно может произойти с одной и той же вероятностью p.

Теорема. *Математическое ожидание случайной величины X, распределенной по биномиальному закону,* $M(X)=np$, *а ее дисперсия* $D(X)=npq$.

Пример. Студент в сессию сдает 4 экзамена. Вероятность успешной сдачи одного экзамена равна 0,7. Составить закон распределения случайной величины X – количество успешно сданных экзаменов. Найдите $M(X)$ и $D(X)$.

Решение. Вероятности возможных значений 0, 1, 2, 3, 4 находим по формуле

$$P(X=k)=C_4^k(0,7)^k(0,3)^{4-k},\ k=0,1,...,4.$$

Ряд распределения имеет следующий вид (табл. 36).

Таблица 36

X	0	1	2	3	4
P	0,0081	0,0756	0,2646	0,4116	0,2401

$$M(X)=np=4\cdot 0,7=2,8,\ D(X)=npq=4\cdot 0,7\cdot 0,3=0,84.$$

2. Геометрическое распределение

Дискретная случайная величина X имеет *геометрическое распределение*, если она принимает значения 1, 2, ..., m, ... (бесконечное, но счетное множество значений) с вероятностями $P(X=m)=pq^{m-1}$, где $0<p<1$, $q=1-p$, $m=0, 1, 2, ...$

Вероятности pq^{m-1} образуют геометрическую прогрессию с первым членом p и знаменателем q (отсюда и название «геометрическое распределение»). Случайная величина, имеющая геометрическое распределение, представляет собой число m испытаний, проведенных по схеме Бернулли, с вероятностью p наступления события в каждом испытании до первого положительного исхода.

Теорема. *Математическое ожидание случайной величины X, имеющей геометрическое распределение с параметром p равно* $M(X)=\frac{1}{p}$, *а ее дисперсия* $D(X)=\frac{q}{p^2}$.

Пример. Курсовую работу можно пересдавать неограниченное количество раз. Вероятность сдачи курсовой работы на «отлично» равна 0,1 и не зависит от номера попытки. Студент сдает работу до получения отлич-

ной оценки. Составить закон распределения случайной величины X – количество попыток. Найдите $M(X)$ и $D(X)$.

Решение. Вероятности возможных значений 1, 2, 3, … находим по формуле

$$P(X=m)=0{,}1\cdot(0{,}9)^{m-1},\ m=1,\ 2,\ \ldots.$$

Ряд распределения имеет следующий вид (табл. 37).

Таблица 37

X	1	2	3	…	m	…
P	0,1	0,09	0,081	…	$0{,}1\cdot(0{,}9)^{m-1}$	…

Тогда

$$M(X)=\frac{1}{p}=\frac{1}{0{,}1}=10,\ D(X)=\frac{q}{p^2}=\frac{0{,}9}{(0{,}1)^2}=90.$$

3. Гипергеометрическое распределение

Дискретная случайная величина X имеет *гипергеометрическое распределение*, если она принимает значения $\max(0, M+n-N), \ldots, \min(n, M)$ с вероятностями

$$P(X=m)=\frac{C_M^m\cdot C_{N-M}^{n-m}}{C_N^n}.$$

Гипергеометрическое распределение имеет случайная величина X – число объектов, обладающих данным свойством, среди n объектов, случайно извлеченных (без возврата) из совокупности N объектов, M из которых обладают этим свойством.

Теорема. *Математическое ожидание случайной величины X, имеющей гипергеометрическое распределение с параметрами* N, M, n *равно* $M(X)=n\cdot\frac{M}{N}$, *а ее дисперсия* $D(X)=n\cdot\frac{M}{N}\cdot\left(1-\frac{M}{N}\right)\cdot\frac{N-n}{N-1}$.

Пример. Проводится выборочная проверка знаний студентов второго курса. Из семи изученных дисциплин четыре гуманитарных. Составить закон распределения случайной величины X – количество гуманитарных дисциплин среди трех случайно выбранных.

Решение. Вероятности возможных значений 0, 1, 2, 3 находим по формуле

$$P(X=m)=\frac{C_4^m\cdot C_3^{n-m}}{C_7^3},\ m=0,1,2,3.$$

Тогда

$$P(X=0)=\frac{C_4^0\cdot C_3^3}{35}=\frac{1}{35},\ P(X=1)=\frac{C_4^1\cdot C_3^2}{35}=\frac{12}{35},$$

$$P(X=2)=\frac{C_4^2\cdot C_3^1}{35}=\frac{18}{35},\ P(X=3)=\frac{C_4^3\cdot C_3^0}{35}=\frac{4}{35},$$

Ряд распределения имеет следующий вид (табл. 38).

Таблица 38

X	0	1	2	3
P	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$

$$M(X)=3\cdot\frac{4}{7}=\frac{12}{7},\ D(X)=3\cdot\frac{4}{7}\cdot\left(1-\frac{4}{7}\right)\cdot\frac{7-3}{7-1}=\frac{12}{7}\cdot\frac{3}{7}\cdot\frac{4}{6}=\frac{24}{49}.$$

4. Распределение Пуассона

Дискретная случайная величина X распределена по *закону Пуассона* с параметром $\lambda>0$, если она принимает значения $0, 1, \ldots$ с вероятностями

$$P(X=m)=\frac{\lambda^m e^{-\lambda}}{m!}.$$

Распределение Пуассона имеет случайная величина X – число наступлений события в серии из n испытаний, в каждом из которых событие наступает с вероятностью p, при больших значениях n и малых значениях p. Параметр $\lambda=np$ дает среднее значение наступлений события.

Теорема. *Математическое ожидание и дисперсия случайной величины X, распределенной по закону Пуассона с параметром λ, равны значению этого параметра* $M(X)=D(X)=\lambda$.

Пример. Автоматическая телефонная станция получает в среднем за час 420 вызовов. Составить закон распределения случайной величины X – число вызовов, поступивших за данную минуту.

Решение. Т.к. вызовы независимы друг от друга, то число вызовов за определенный промежуток времени – случайная величина, распределенная по закону Пуассона. Среднее число вызовов, поступающих за минуту равно $\lambda=420\cdot\frac{1}{60}=7$. Тогда

$$P(X=1)=\frac{7^1e^{-7}}{1!}=7e^{-7}\approx 0{,}0009,\ P(X=2)=\frac{7^2e^{-7}}{2!}=\frac{49}{2}e^{-7}\approx 0{,}0222$$

и т.д. Ряд распределения имеет вид (табл. 39).

Таблица 39

X	1	2	…	m	…
P	0,0009	0,0222	…	$P(X=m)=\dfrac{7^m e^{-7}}{m!}$	…

Математическое ожидание

$$M(X)=D(X)=\lambda=7.$$

Контрольные вопросы

1. Что называется случайной величиной?
2. Что называется дискретной случайной величиной?
3. Что называется законом распределения случайной величины?
4. Как задается закон распределения дискретной случайной величины?
5. Какие действия можно производить над случайными величинами?
6. Приведите определение и свойства математического ожидания дискретной случайной величины.
7. Приведите определение и свойства дисперсии дискретной случайной величины.
8. Дайте определение случайной величины, распределенной по биномиальному закону, и приведите формулы для нахождения ее математического ожидания и дисперсии.
9. Дайте определение случайной величины, распределенной по геометрическому закону, и приведите формулы для нахождения ее математического ожидания и дисперсии.
10. Дайте определение случайной величины, распределенной по гипергеометрическом закону, и приведите формулы для нахождения ее математического ожидания и дисперсии.
11. Дайте определение случайной величины, распределенной по закону Пуассона, и приведите формулы для нахождения ее математического ожидания и дисперсии.
12. Дайте определение и приведите свойства функции распределения случайной величины.

НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Определение непрерывной случайной величины

Случайная величина называется непрерывной, если ее функция распределения есть непрерывная, кусочно-дифференцируемая функция.

Замечание. Для любой случайной величины справедлива формула

$$P(a \le X < b)=F(b)-F(a).$$

Действительно, событие $(X<b)$ можно представить в виде суммы двух несовместных событий $(a \le X < b)$ и $(X<a)$. Значит,

$$P(X<b)=P(a\le X<b)+P(X<a)$$

или

$$P(a\le X<b)=P(X<b)-P(X<a)=F(b)-F(a).$$

Утверждение 1. Если X – непрерывная случайная величина, то вероятность любого отдельно взятого ее значения равна нулю, т.е. $P(X=x_0)=0,\ x_0\in\mathbb{R}$.

Действительно,

$$P(X=x_0)=\lim_{x\to x_0}\left(F(x)-F(x_0)\right)=\lim_{x\to x_0}F(x)-\lim_{x\to x_0}F(x_0)=$$
$$=\lim_{x\to x_0}F(x)-F(x_0)=0.$$

Утверждение 2. Если X – непрерывная случайная величина, то

$$P(a\le X\le b)=P(a\le X<b)=P(a<X\le b)=P(a<X<b).$$

Например, $P(a<X\le b)=P(a<X<b)+P(X=b)=P(a<X<b)$.

Функция плотности распределения вероятностей непрерывной случайной величины

Пусть заданы $x\in\mathbb{R}$ и $\Delta x>0$. Рассмотрим вероятность попадания непрерывной случайной величины в интервал $(x;\ x+\Delta x)$:

$$P(x<X<x+\Delta x)=F(x+\Delta x)-F(x)=\Delta F(x).$$

Рассмотрим «плотность» вероятности на этом участке $\dfrac{\Delta F(x)}{\Delta x}$.

Каждому $x\in\mathbb{R}$ поставим в соответствие $\lim\limits_{\Delta x\to 0}\dfrac{\Delta F(x)}{\Delta x}$.

Получим *функцию плотности распределения вероятности*. Она обозначается $f(x)$ и равна

$$f(x)=F'(x).$$

График этой функции называется *кривой распределения*.

Из формулы $f(x)=F'(x)$ следует, что $F(x)=\int\limits_{-\infty}^{x}f(t)dt$. Тогда

$$P(a<X<b)=F(b)-F(a)=\int\limits_{-\infty}^{b}f(x)dx-\int\limits_{-\infty}^{a}f(x)dx=$$
$$=\int\limits_{-\infty}^{a}f(x)dx+\int\limits_{a}^{b}f(x)dx-\int\limits_{-\infty}^{a}f(x)dx=\int\limits_{a}^{b}f(x)dx.$$

Геометрически эта вероятность равна площади криволинейной трапеции, ограниченной графиком кривой распределения, осью Ox, прямыми $x=a$ и $x=b$.

Свойства функции плотности

1. Для всех $x\in\mathbb{R}$ $f(x)\geq 0$, т.к. эта функция является производной неубывающей функции $F(x)$.

2. $\int\limits_{-\infty}^{+\infty} f(x)dx=1$ (основное условие нормировки).

Действительно, $\int\limits_{-\infty}^{+\infty} f(x)dx=F(x)\Big|_{-\infty}^{+\infty}=\lim\limits_{x\to+\infty}F(x)-\lim\limits_{x\to-\infty}F(x)=1-0=1.$

Пример. При каком значении параметра a функция

$$f(x)=\begin{cases}0 \text{ при } x\leq 0,\\ ax^2+0{,}5 \text{ при } 0<x\leq 1,\\ 0 \text{ при } x>1\end{cases}$$

является функцией плотности распределения некоторой случайной величины.

Решение. Функция $f(x)$ должна удовлетворять основному условию нормировки:

$$\int\limits_{-\infty}^{+\infty} f(x)dx=1.$$

$$\int\limits_{-\infty}^{+\infty} f(x)dx=\int\limits_{-\infty}^{0}0dx+\int\limits_{0}^{1}\left(ax^2+0{,}5\right)dx+\int\limits_{1}^{+\infty}0dx=\left(\frac{ax^3}{3}+\frac{x}{2}\right)\Bigg|_0^1=\frac{a}{3}+\frac{1}{2}=1.$$

Отсюда $\frac{a}{3}=\frac{1}{2}$, $a=\frac{3}{2}$.

Пример. Случайная величина X задана функцией распределения

$$F(x)=\begin{cases}0 \text{ при } x\leq 0,\\ \sin x \text{ при } 0<x\leq\frac{\pi}{2},\\ 1 \text{ при } x>\frac{\pi}{2}.\end{cases}$$

Найти плотность распределения X.

Решение. Функция $F(x)$ является кусочно-дифференцируемой. Дифференцируя ее, получаем

$$f(x)=\begin{cases}0 \text{ при } x<0,\\ \cos x \text{ при } 0<x<\dfrac{\pi}{2},\\ 0 \text{ при } x>\dfrac{\pi}{2}.\end{cases}$$

Пример. Непрерывная случайная величина X задана плотностью распределения на всей числовой оси: $f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}$.

Найти ее функцию распределения и вероятность того, что X примет значение на интервале $(-1;1)$.

Решение. Искомая вероятность равна

$$P(-1<X<1)=\int_{-1}^{1} f(x)dx=\frac{1}{\pi}\int_{-1}^{1}\frac{dx}{1+x^2}=\frac{1}{\pi}\operatorname{arctg}x\Big|_{-1}^{1}=\frac{1}{\pi}\left(\frac{\pi}{4}+\frac{\pi}{4}\right)=\frac{1}{2}$$

или

$$P(-1<X<1)=F(1)-F(-1).$$

Найдем функцию распределения по формуле

$$F(x)=\int_{-\infty}^{x} f(t)dt:$$

$$F(x)=\frac{1}{\pi}\int_{-\infty}^{x}\frac{dt}{1+t^2}=\operatorname{arctg}t\Big|_{-\infty}^{x}=\frac{1}{\pi}\operatorname{arctg}x-\frac{1}{\pi}\lim_{x\to-\infty}\operatorname{arctg}t=\frac{1}{\pi}\operatorname{arctg}x+\frac{1}{2}.$$

Тогда $P(-1<X<1)=F(1)-F(-1)=\left(\frac{1}{\pi}\operatorname{arctg}1+\frac{1}{2}\right)-\left(\frac{1}{\pi}\operatorname{arctg}(-1)+\frac{1}{2}\right)=$

$$=\left(\frac{1}{4}+\frac{1}{2}\right)-\left(-\frac{1}{4}+\frac{1}{2}\right)=\frac{1}{2}.$$

Числовые характеристики непрерывных случайных величин

1) *Математическое ожидание* $M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx$.

2) *Дисперсия* $D(X)=M\left(X^2\right)-\left(M(X)\right)^2$, где $M\left(X^2\right)=\int_{-\infty}^{+\infty} x^2 f(x)dx$.

3) Для непрерывной случайной величины ***мода*** – такое значение случайной величины, при которой плотность распределения достигает максимума.

4) Геометрически *медиана* – абсцисса точки, в которой площадь, ограниченная кривой распределения делится пополам.

Пример. Найдите числовые характеристики случайной величины, заданной функцией плотности

$$f(x)=\begin{cases}0 \text{ при } x\le 0,\\ 1{,}5x^2+0{,}5 \text{ при } 0<x\le 1,\\ 0 \text{ при } x>1.\end{cases}$$

Решение.

1) $M(X)=\int\limits_{-\infty}^{+\infty} x\cdot f(x)dx=\int\limits_{-\infty}^{0} x\cdot 0\cdot dx+\int\limits_{0}^{1} x\left(1{,}5x^2+0{,}5\right)dx+\int\limits_{1}^{+\infty} x\cdot 0\cdot dx=$

$$=\left(\frac{3x^4}{2\cdot 4}+\frac{x^2}{2\cdot 2}\right)\Bigg|_0^1=\frac{3}{8}+\frac{1}{4}=\frac{5}{8}.$$

2) $M\left(X^2\right)=\int\limits_{-\infty}^{+\infty} x^2 f(x)dx=\int\limits_{-\infty}^{0} x^2\cdot 0\cdot dx+\int\limits_{0}^{1} x^2\left(1{,}5x^2+0{,}5\right)dx+$

$$+\int\limits_{1}^{+\infty} x^2\cdot 0\cdot dx=\left(\frac{3x^5}{2\cdot 5}+\frac{x^3}{2\cdot 3}\right)\Bigg|_0^1=\frac{3}{10}+\frac{1}{6}=\frac{14}{30}=\frac{7}{15}.$$

$$D(X)=M\left(X^2\right)-\left(M(X)\right)^2=\frac{7}{15}-\left(\frac{5}{8}\right)^2=\frac{7}{15}-\frac{25}{64}=\frac{73}{960}.$$

3) $Mo(X)=1$.

4) Обозначим $m=Me(X)$.

Должно выполняться $\int\limits_0^m f(x)dx=\frac{1}{2}$, т.е.

$$\int\limits_0^m\left(1{,}5x^2+0{,}5\right)dx=\left(\frac{1}{2}x^3+\frac{1}{2}x\right)\Bigg|_0^m==\frac{1}{2}m^3+\frac{1}{2}m=\frac{1}{2}.$$

Отсюда получаем $m\approx 0{,}69$.

Основные законы распределения непрерывных случайных величин

1. Равномерный закон распределения

Непрерывная случайная величина X имеет *равномерный закон распределения* на отрезке $[a;b]$, если ее плотность вероятности $f(x)$ постоянна на этом отрезке и равна нулю вне этого отрезка, т.е.

$$f(x)=\begin{cases}0 \text{ при } x<a,\\ C \text{ при } a\le x\le b,\\ 0 \text{ при } x>b.\end{cases}$$

Определим значение константы C, исходя из свойства плотности вероятности: $\int\limits_{-\infty}^{+\infty} f(x)dx=1$.

$$\int\limits_{-\infty}^{+\infty} f(x)dx=\int\limits_{-\infty}^{a} 0\cdot dx+\int\limits_{a}^{b} C\cdot dx+\int\limits_{b}^{+\infty} 0\cdot dx=C\cdot x\Big|_{a}^{b}=C(b-a)=1.$$

Отсюда $C=\dfrac{1}{b-a}$.

Таким образом, непрерывная случайная величина X имеет *равномерный закон распределения*, на отрезке $[a, b]$, если ее плотность вероятности имеет вид

$$f(x)=\begin{cases}0 \text{ при } x<a,\\ \dfrac{1}{b-a} \text{ при } a\le x\le b,\\ 0 \text{ при } x>b.\end{cases}$$

Кривая распределения $f(x)$ приведена на рис. 12.

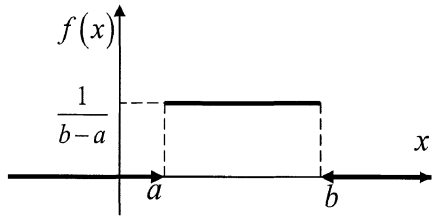


Рис. 12

Теорема. *Функция распределения случайной величины X, распределенной по равномерному закону, есть* $F(x)=\begin{cases} 0 \text{ при } x<a, \\ \dfrac{x-a}{b-a} \text{ при } a\le x\le b, \\ 1 \text{ при } x>b, \end{cases}$ *ее математическое ожидание* $M(X)=\dfrac{a+b}{2}$, *а ее дисперсия* $D(X)=\dfrac{(b-a)^2}{12}$.

Доказательство. При $x\le a$ $F(x)=\int\limits_{-\infty}^{x} 0\cdot dx=0$.

При $a<x\le b$ $F(x)=\int\limits_{-\infty}^{a} 0\cdot dx+\int\limits_{a}^{x}\dfrac{1}{b-a}dx=\dfrac{1}{b-a}x\Bigg|_{a}^{b}=\dfrac{x-a}{b-a}$.

При $x>b$ $F(x)=\int\limits_{-\infty}^{a} 0\cdot dx+\int\limits_{a}^{b}\dfrac{1}{b-a}dx+\int\limits_{b}^{+\infty} 0\cdot dx=1$.

$$M(X)=\int\limits_{-\infty}^{+\infty} x\cdot f(x)dx=\int\limits_{-\infty}^{a} x\cdot 0dx+\int\limits_{a}^{b} x\frac{1}{b-a}dx+\int\limits_{b}^{+\infty} x\cdot 0dx=\frac{1}{b-a}\frac{x^2}{2}\Bigg|_{a}^{b}=$$

$$=\frac{1}{b-a}\cdot\frac{b^2-a^2}{2}=\frac{a+b}{2}.$$

$$M\left(X^2\right)=\int\limits_{-\infty}^{+\infty} x^2\cdot f(x)dx=\int\limits_{a}^{b} x^2\cdot\frac{1}{b-a}dx=\frac{1}{b-a}\cdot\frac{x^3}{3}\Bigg|_{a}^{b}=$$

$$=\frac{1}{b-a}\frac{b^3-a^3}{3}=\frac{a^2+ab+b^2}{3}.$$

Тогда $D(X)=\dfrac{a^2+ab+b^2}{3}-\left(\dfrac{a+b}{2}\right)^2=\dfrac{(b-a)^2}{12}$.

График функции распределения $F(x)$ равномерно распределенной случайной величины X представлен на рис. 13.

Вероятность попадания равномерно распределенной случайной величины в интервал $(\alpha;\beta)$, если $a\le\alpha<\beta\le b$, равна

$$P(\alpha<X<\beta)=F(\beta)-F(\alpha)=\frac{\beta-a}{b-a}-\frac{\alpha-a}{b-a}=\frac{\beta-\alpha}{b-a}.$$

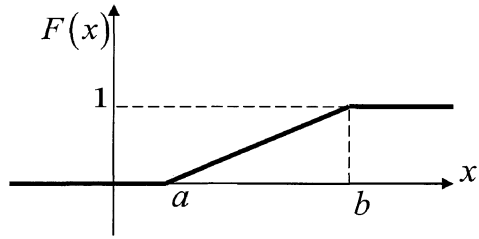


Рис. 13

Пример. Случайная величина распределена равномерно на отрезке $[-1;2]$.

Тогда

$$P(0<X<1)=\frac{1-0}{2-(-1)}=\frac{1}{3},\ P(-1<X<2)=\frac{2-(-1)}{2-(-1)}=1,$$

$$P(-2<X<1)=P(-1<X<1)=\frac{1-(-1)}{2-(-1)}=\frac{2}{3},$$

$$P(0<X<3)=P(0<X<1)=\frac{1-0}{2-(-1)}=\frac{1}{3}.$$

Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчетов (например, ошибка округления числа до целого распределена равномерно на отрезке [–0,5; 0,5]), в ряде задач массового обслуживания, при статистическом моделировании наблюдений, подчиненных заданному распределению.

Так, случайная величина X, распределенная равномерно на отрезке [0; 1], называемая «случайным числом от 0 до 1», служит основой для получения случайных величин с любым законом распределения.

Пример. Автобус ходит регулярно с интервалом 10 минут. Пассажир приходит на остановку в случайный момент времени. Найти математическое ожидание и среднее квадратическое отклонение случайной величины X – времени ожидания автобуса. Какова вероятность того, что ждать пассажиру придется не больше двух минут?

Решение. Случайная величина X– время ожидания автобуса на временно́м (в минутах) отрезке [0; 10] имеет равномерный закон распределения $f(x)=\frac{1}{10}=0{,}1$. Поэтому, среднее время ожидания автобуса $M(X)=\frac{0+10}{2}=5$ мин. Дисперсия времени ожидания автобуса $D(X)=\frac{(10-0)^2}{12}=\frac{100}{12}\approx 8{,}34$. Отсюда $\sigma(X)=\sqrt{D(X)}=\sqrt{8{,}34}\approx 2{,}89$.

Вероятность события A – пассажиру придется ждать не больше двух минут $P(A)=P(X\le 2)=P(0\le X\le 2)=\int\limits_0^2 0{,}1\cdot dx=0{,}2$.

2. Показательный (экспоненциальный) закон распределения

Непрерывная случайная величина X имеет *показательный (экспоненциальный) закон распределения* с параметром λ, если ее плотность вероятности $f(x)$ имеет вид:

$$f(x)=\begin{cases}0 \text{ при } x<0,\\ \lambda e^{-\lambda x} \text{ при } x\ge 0.\end{cases}$$

Параметр λ должен быть положительным, т.к. функция плотности удовлетворяет свойству $f(x)\ge 0$ и $e^{-\lambda x}\ge 0$.

Проверим, что эта функция удовлетворяет основному условию нормировки:

$$\int\limits_{-\infty}^{+\infty} f(x)dx=\int\limits_{-\infty}^{0} 0\cdot dx+\int\limits_{0}^{+\infty}\lambda e^{-\lambda x}dx=-e^{-\lambda x}\Big|_0^{+\infty}=-\lim_{x\to+\infty}e^{-\lambda x}+1=0+1=1.$$

Кривая распределения $f(x)$ приведена на рис. 14.

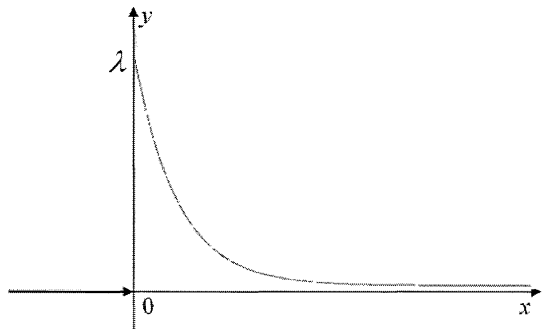


Рис. 14

Теорема. *Функция распределения случайной величины X, распределенной по показательному закону, есть*

$$F(x)=\begin{cases}0 \text{ при } x<0,\\ 1-e^{-\lambda x} \text{ при } x\ge 0,\end{cases}$$

ее математическое ожидание $M(X)=\dfrac{1}{\lambda}$, *а ее дисперсия* $D(X)=\dfrac{1}{\lambda^2}$.

Доказательство. При $x<0$ $F(x)=\int\limits_{-\infty}^{x} 0\cdot dt=0$.

При $x\geq 0$ $F(x)=\int\limits_{-\infty}^{0} 0dt+\int\limits_{0}^{x}\lambda e^{-\lambda t}dt=-e^{-\lambda t}\Big|_{0}^{x}=-e^{-\lambda x}+1$.

$$M(X)=\int\limits_{0}^{+\infty}\lambda xe^{-\lambda x}dx=\begin{bmatrix} u=x & dv=\lambda e^{-\lambda x}dx \\ du=dx & v=-e^{-\lambda x}\end{bmatrix}=-xe^{-\lambda x}\Big|_{0}^{+\infty}+\int\limits_{0}^{+\infty}e^{-\lambda x}dx=$$

$$=-\lim_{x\to+\infty}xe^{-\lambda x}-0-\frac{1}{\lambda}e^{-\lambda x}\Big|_{0}^{+\infty}=-\lim_{x\to+\infty}\frac{x}{e^{\lambda x}}-\frac{1}{\lambda}\lim_{x\to+\infty}e^{-\lambda x}+\frac{1}{\lambda}=$$

$$=-\lim_{x\to+\infty}\frac{1}{\lambda e^{\lambda x}}-0+\frac{1}{\lambda}=\frac{1}{\lambda}.$$

Отсюда следует, что $\int\limits_{0}^{+\infty}xe^{-\lambda x}dx=\frac{1}{\lambda^2}$.

$$M\left(X^2\right)=\int\limits_{0}^{+\infty}\lambda x^2e^{-\lambda x}dx=\begin{bmatrix} u=x^2 & dv=\lambda e^{-\lambda x}dx \\ du=2xdx & v=-e^{-\lambda x}\end{bmatrix}=$$

$$=-x^2e^{-\lambda x}\Big|_{0}^{+\infty}+\int\limits_{0}^{+\infty}2xe^{-\lambda x}dx=\frac{2}{\lambda^2};$$

$$D(X)=M\left(X^2\right)-\left(M(X)\right)^2=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2},\ \sigma(X)=\frac{1}{\lambda}.$$

Следствие. *Для случайной величины, распределенной по показательному закону, математическое ожидание равно среднему квадратическому отклонению:* $M(X)=\sigma(X)$.

Вероятность попадания показательно распределенной случайной величины в интервал $(\alpha;\beta)$, если $\alpha\geq 0$, равна

$$P(\alpha<X<\beta)=F(\beta)-F(\alpha)=\left(1-e^{-\lambda\beta}\right)-\left(1-e^{-\lambda\alpha}\right)=e^{-\lambda\alpha}-e^{-\lambda\beta}.$$

Пример. Установлено, что время ремонта телевизоров есть случайная величина X, распределенная по показательному закону. Определить вероятность того, что на ремонт телевизора потребуется не менее 20 дней, если среднее время ремонта телевизоров составляет 15 дней. Найти плотность

вероятности, функцию распределения и среднее квадратическое отклонение случайной величины X.

Решение. Параметр $\lambda = \frac{1}{M(X)} = \frac{1}{15}$. Вероятность того, что на ремонт телевизора потребуется не менее 20 дней равна

$$P(X \geq 20) = 1 - P(X < 20) = 1 - F(20) = 1 - \left(1 - e^{-\frac{20}{15}}\right) = e^{-\frac{4}{3}} \approx 0{,}26.$$

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ \frac{1}{15} e^{-\frac{1}{15}x} \text{ при } x \geq 0; \end{cases} \qquad F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 1 - e^{-\frac{1}{15}x} \text{ при } x \geq 0; \end{cases}$$

$$\sigma(X) = 15.$$

3. Нормальный закон распределения

Непрерывная случайная величина X имеет *нормальный закон распределения* с параметрами a и σ, если ее плотность вероятности $f(x)$ имеет вид

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

Параметр σ должен быть положительным, т.к. $f(x) \geq 0$ при всех значениях x.

Область определения этой функции – вся числовая ось.

Исследуем функцию средствами дифференциального исчисления.

$$f'(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \cdot \left(-\frac{(x-a)^2}{2\sigma^2}\right)' = -\frac{2(x-a)}{2\sigma^2\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} =$$

$$= -\frac{x-a}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Эта производная равна нулю в точке $x = a$. Причем слева от этой точки производная положительна, а справа – отрицательна. Значит, в самой точке функция имеет максимум.

$$f''(x) = \left(-\frac{x-a}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}\right)' = -\frac{1}{\sigma^3\sqrt{2\pi}} \left((x-a) e^{-\frac{(x-a)^2}{2\sigma^2}}\right)' =$$

$$=-\frac{1}{\sigma^3\sqrt{2\pi}}\left(e^{-\frac{(x-a)^2}{2\sigma^2}}-(x-a)\cdot\frac{2(x-a)}{2\sigma^2}\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}\right)=$$

$$=-\frac{1}{\sigma^3\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\left(1-\frac{(x-a)^2}{\sigma^2}\right).$$

Эта производная равна нулю в точках $x=a\pm\sigma$. В этих точках график функции имеет перегибы.

Итак, есть максимум в точке $\left(a;\frac{1}{\sigma\sqrt{2\pi}}\right)$ и перегибы в точках $\left(a\pm\sigma;\frac{1}{\sigma\sqrt{2\pi e}}\right)$.

Кривая распределения нормального закона $f(x)$ (нормальная кривая или кривая Гаусса) приведена на рис. 15.

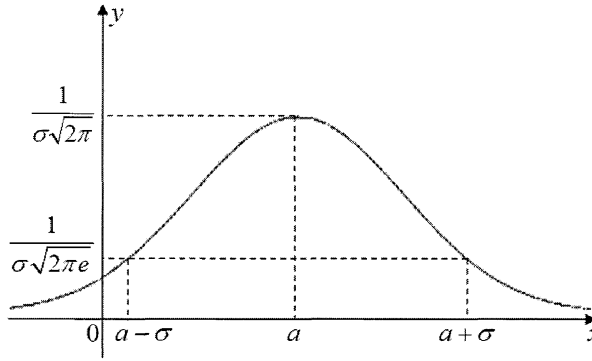


Рис. 15

Теорема. *Математическое ожидание случайной величины X, распределенной по нормальному закону, равно параметру a этого закона, а ее дисперсия – квадрату параметра* σ, *т.е.* $M(X)=a,\ D(X)=\sigma^2$.

Доказательство.

$$M(X)=\int_{-\infty}^{+\infty}xf(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}xe^{-\frac{(x-a)^2}{2\sigma^2}}dx.$$

Сделаем в этом интеграле замену $t=\frac{x-a}{\sigma\sqrt{2}}$, $x=\sigma t\sqrt{2}+a$. Тогда $dx=\sigma\sqrt{2}dt$.

$$M(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(\sigma t\sqrt{2}+a\right)e^{-t^2}\sigma\sqrt{2}dt=\frac{\sigma\sqrt{2}}{\sigma\sqrt{2\pi}}\left(\sigma\sqrt{2}\int_{-\infty}^{+\infty}te^{-t^2}dt+a\int_{-\infty}^{+\infty}e^{-t^2}dt\right)=$$

$$=\frac{a}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-t^2}dt.$$

Интеграл

$$\int_{-\infty}^{+\infty}e^{-x^2}dx=\sqrt{\pi}$$

встречается в теории вероятностей и в задачах математической физики. С.Д. Пуассон предложил прием для вычисления этого интеграла. Впервые этот интеграл был вычислен Л. Эйлером, поэтому называется интегралом Эйлера–Пуассона.

Таким образом, $M(X)=\frac{a}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-t^2}dt=a$.

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx.$$

Сделаем в этом интеграле замену $t=\frac{x-a}{\sigma\sqrt{2}}$, $x=\sigma t\sqrt{2}+a$. Тогда $dx=\sigma\sqrt{2}dt$.

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\cdot 2\sigma^3\sqrt{2}\int_{-\infty}^{+\infty}t^2e^{-t^2}dt=\frac{4\sigma^2}{\sqrt{\pi}}\int_{0}^{+\infty}t^2e^{-t^2}dt=\begin{bmatrix} u=t & dv=te^{-t^2}dt \\ du=dt & v=-\frac{1}{2}e^{-t^2}\end{bmatrix}=$$

$$=\frac{4\sigma^2}{\sqrt{\pi}}\left(-\frac{1}{2}te^{-t^2}\Big|_{0}^{+\infty}+\frac{1}{2}\int_{0}^{+\infty}e^{-t^2}dt\right)=\frac{4\sigma^2}{\sqrt{\pi}}\cdot\frac{1}{2}\cdot\frac{\sqrt{\pi}}{2}=\sigma^2.$$

Таким образом, $D(X)=\sigma^2$ и $\sigma(X)=\sigma$.

Рассмотрим зависимость положения и формы кривой Гаусса от значений параметров нормального распределения.

Если $\sigma=const$ и меняется параметр a, то кривая распределения будет смещаться вдоль оси абсцисс, не меняя формы (рис. 16).

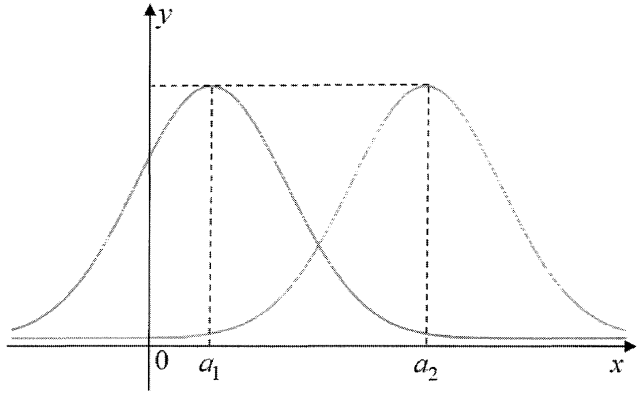


Рис. 16

Если $a = const$ и меняется параметр σ, то меняется ордината точки максимума кривой.

При увеличении σ кривая распределения становится более плоской, растягиваясь вдоль оси абсцисс; при уменьшении σ кривая вытягивается вверх, одновременно сжимаясь с боков (рис. 17). Здесь $\sigma_1 > \sigma_2$.

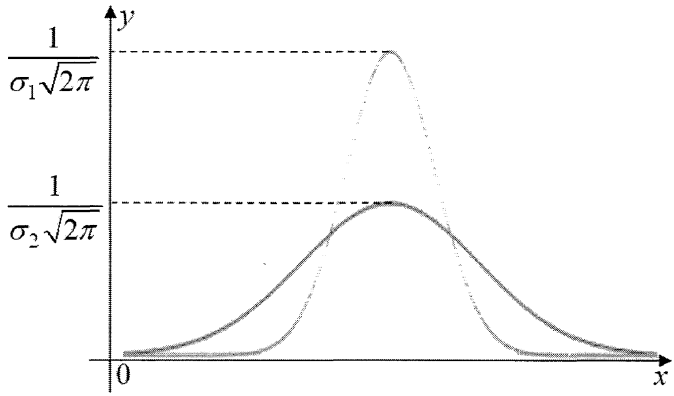


Рис. 17

Функция распределения $F(x)$ случайной величины X, распределенной по нормальному закону имеет вид

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right).$$

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

Сделаем в этом интеграле замену $z=\frac{t-a}{\sigma}$, $t=\sigma z+a$. Тогда $dt=\sigma dz$.

Если $t=x$, то $z=\frac{x-a}{\sigma}$.

Тогда $F(x)=\frac{\sigma}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}}dz=\frac{1}{\sqrt{2\pi}}\left(\int\limits_{-\infty}^{0} e^{-\frac{z^2}{2}}dz+\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}}dz\right)=$

$$=\frac{1}{\sqrt{2\pi}}\left(\sqrt{\frac{\pi}{2}}+\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}}dz\right)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right).$$

Нормальный закон распределения случайной величины с параметрами $a=0$ и $\sigma=1$ называется *стандартным* или *нормированным*, а соответствующая нормальная кривая – стандартной или нормированной.

Его функция плотности имеет вид

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

Вероятность попадания случайной величины в интервал $(\alpha;\beta)$, равна

$$P(\alpha<X<\beta)=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right).$$

Т.к. $F(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right)$, то

$$P(\alpha<X<\beta)=F(\beta)-F(\alpha)=\left(\frac{1}{2}+\Phi\left(\frac{\beta-a}{\sigma}\right)\right)-\left(\frac{1}{2}+\Phi\left(\frac{\alpha-a}{\sigma}\right)\right)=$$

$$=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right).$$

Следствие. Вероятность попадания в симметричный относительно математического ожидания интервал:

$$P(|X-a|<\varepsilon)=P(a-\varepsilon<X<a+\varepsilon)=$$

$$=\Phi\left(\frac{a+\varepsilon-a}{\sigma}\right)-\Phi\left(\frac{a-\varepsilon-a}{\sigma}\right)=2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

В частности,

$$P(|X-a|<\sigma)=2\Phi\left(\frac{\sigma}{\sigma}\right)=2\Phi(1)\approx 0,6826,$$

$$P(|X-a|<2\sigma)=2\Phi\left(\frac{2\sigma}{\sigma}\right)=2\Phi(2)\approx 0,9544,$$

$$P(|X-a|<3\sigma)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)\approx 2\cdot 0,49865\approx 1.$$

Правило трех сигм. Если случайная величина X распределена нормально (с параметрами a и σ), то практически достоверно, что абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения, т.е. $P(|X-a|<3\sigma)\approx 1$.

Другими словами, если случайная величина X имеет нормальный закон распределения с параметрами a и σ, то практически достоверно, что ее значения заключены в интервале $(a-3\sigma;\ a+3\sigma)$.

Из последнего равенства можно сделать вывод о том, что нарушение «правила трех сигм», т.е. отклонение нормально распределенной случайной величины X больше, чем на 3σ (по абсолютной величине), является событием практически невозможным, так как его вероятность достаточно мала.

Пример. Станок-автомат изготавливает шарики. Шарик считается годным, если отклонение X диаметра шарика от проектного размера по абсолютной величине не превосходит 0,7 мм. Опытным путем установлено, что $\sigma=0,4$ мм. Считая, что X является нормально распределенной случайной величиной, определить, сколько процентов годных шариков изготавливает автомат.

Решение. Отклонение X диаметра шарика от проектного размера – нормально распределенная случайная величина с параметрами $a=0$ и $\sigma=0,4$.

Каждый изготовленный станком-автоматом шарик будет считаться годным, если ошибка X его диаметра будет удовлетворять неравенству $|X|\le 0,7$, т.е. она будет отклоняться от среднего значения не более, чем на 0,7 мм.

Найдем вероятность этого события.

$$P(|X|\le 0,7)=2\Phi\left(\frac{0,7}{0,4}\right)=2\Phi(1,75)\approx 2\cdot 0,4599\approx 0,92.$$

Итак, вероятность того, что каждый изготовленный станком-автоматом шарик окажется годным, составляет 0,92. Это значит, что этот станок изготовляет в среднем 92% годных шариков.

Центральная предельная теорема

Это группа теорем, посвященных установлению условий, при которых возникает нормальный закон распределения.

Теорема Ляпунова. *Пусть* X_1, X_2, ..., X_n *– независимые случайные величины с математическими ожиданиями* $M(X_k)=a_k$ *и дисперсиями* $D(X_k)=\sigma_k^2$, k = 1, 2,..., n. *Кроме того,* $M\left(\left|X_k-a_k\right|^3\right)=m_k$ *и*

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} m_k}{\left(\sum_{k=1}^{n}\sigma_k^2\right)}=0.$$

Тогда закон распределения суммы $X_1+X_2+...+X_n$, при $n\to\infty$ неограниченно приближается к нормальному с математическим ожиданием $\sum_{k=1}^{n} a_k$ и дисперсией $\sum_{k=1}^{n}\sigma_k^2$.

Теорема Ляпунова объясняет широкое распространение нормального закона распределения и поясняет механизм его образования. Теорема позволяет утверждать, что всегда, когда случайная величина образуется в результате сложения большого числа независимых случайных величин, дисперсии которых малы по сравнению с дисперсией суммы, закон распределения этой случайной величины оказывается практически нормальным законом.

Поскольку случайные величины всегда порождаются бесконечным количеством причин и, чаще всего ни одна из них не имеет дисперсии, сравнимой с дисперсией самой случайной величины, то большинство встречающихся в практике случайных величин подчинено нормальному закону распределения.

При стрельбе из орудия под влиянием очень большого числа причин случайного характера происходит рассеяние снарядов на некоторой площади. Случайные воздействия на траекторию снаряда можно считать независимыми. Каждая причина вызывает лишь незначительное изменение траектории по сравнению с суммарным изменением под воздействием всех причин. Поэтому следует ожидать, что отклонение места разрыва снаряда от цели будет случайной величиной, распределенной по нормальному закону.

Вероятность отклонения относительной частоты события от его вероятности

Индикатор события – это случайная величина, принимающая два значения: 0 с вероятностью p и 1 с вероятностью $q = 1 - p$.

Рассмотрим частоту появления события в схеме Бернулли, т.е. если производится n независимых одинаковых испытаний, в каждом из которых событие A может наступить с одной и той же вероятностью p.

Она может быть представлена через индикатор события I_i, в виде

$$\frac{k}{n} = \frac{I_1 + I_2 + \ldots + I_n}{n} = \frac{I_1}{n} + \frac{I_2}{n} + \ldots + \frac{I_n}{n},$$

т.е. является суммой большого числа независимых одинаково распределенных случайных величин $\frac{I_i}{n}$, каждая из которых имеет дисперсию $D\left(\frac{I_i}{n}\right) = \frac{pq}{n} < \infty$. По центральной предельной теореме частота события $\frac{k}{n}$ имеет закон, близкий к нормальному закону распределения, с параметрами

$$M\left(\frac{k}{n}\right) = M\left(\frac{I_1 + I_2 + \ldots + I_n}{n}\right) = \frac{M\left(I_1 + I_2 + \ldots + I_n\right)}{n} = \frac{np}{n} = p,$$

$$D\left(\frac{k}{n}\right) = D\left(\frac{I_1 + I_2 + \ldots + I_n}{n}\right) = \frac{D\left(I_1 + I_2 + \ldots + I_n\right)}{n} = \frac{npq}{n^2} = \frac{pq}{n}.$$

Итак, частота события $\frac{k}{n}$ имеет закон распределения, близкий к нормальному с параметрами $M\left(\frac{k}{n}\right) = p$, $D\left(\frac{k}{n}\right) = \frac{pq}{n}$.

Используем формулу $P\left(|X - a| < \varepsilon\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$ при $X = \frac{k}{n}$:

$$P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 2\Phi\left(\frac{\varepsilon}{\sqrt{\frac{pq}{n}}}\right) = 2\Phi\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right).$$

Итак, вероятность того, что это отклонение относительной частоты события от его вероятности не превышает по модулю заданного числа ε находится по формуле:

$$P\left(\left|\frac{k}{n} - p\right| \le \varepsilon\right) \approx 2\Phi\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right). \qquad (*)$$

Пример. Вероятность получения нестандартной детали $p = 0{,}1$. Найти вероятность того, что среди случайно взятых 200 деталей относительная

частота появления нестандартной детали отклонится от вероятности p по абсолютной величине не более чем на 0,03.

Решение. В данном случае $n=200$, $q=1-p=0{,}9$, $\varepsilon=0{,}03$. По формуле (*) имеем

$$P\left(\left|\frac{k}{200}-0{,}1\right|\le 0{,}03\right)\approx 2\Phi\left(0{,}03\sqrt{\frac{200}{0{,}1\cdot 0{,}9}}\right)=2\Phi\left(0{,}03\sqrt{\frac{200}{0{,}1\cdot 0{,}9}}\right)=$$
$$=2\Phi(1{,}41)\approx 0{,}8414.$$

Смысл полученного результата состоит в том, что при достаточно большом числе проб, каждая из которых содержит 200 случайно выбранных деталей, в 84% случаев отклонение относительной частоты от постоянной вероятности p по абсолютной величине не превысит 0,03.

Следствие формулы (*). $\left|\frac{k}{n}-p\right|\le\varepsilon\Leftrightarrow|k-np|\le n\varepsilon$. Обозначим $\delta=n\varepsilon$.

Т.к. $\frac{\delta}{n}\sqrt{\frac{n}{pq}}=\frac{\delta}{\sqrt{npq}}$, то формула (*) пример вид

$$P\left(|k-np|\le\delta\right)\approx 2\Phi\left(\frac{\delta}{\sqrt{npq}}\right).$$

Неравенство Маркова

Если среди значений случайной величины X нет отрицательных, то вероятность того, что она примет какое-нибудь значение, превосходящее положительное число A, удовлетворяет условию:

$$P(X>A)\le\frac{M(X)}{A} \text{ или } P(X\le A)\ge 1-\frac{M(X)}{A}.$$

Пусть дискретная случайная величина X принимает значения x_1, x_2, …, x_n (в порядке возрастания). Все $x_i\ge 0$. И пусть k значений x_1, x_2, …, x_k будут не больше числа A, а остальные x_{k+1}, …, x_n будут больше этого числа. Математическое ожидание этой случайной величины равно

$$M(X)=x_1p_1+x_2p_2+\ldots+x_np_n,$$

где p_1, p_2, …, p_n – вероятности соответствующих значений.

Отбросим первые k неотрицательных слагаемых. Тогда

$$x_{k+1}p_{k+1}+\ldots+x_np_n\le M(X).$$

Заменим в этом неравенстве значения x_{k+1}, …, x_n меньшим числом A:

$$Ap_{k+1}+\ldots+Ap_n=A(p_{k+1}+\ldots+p_n)\le M(X)$$

или

$$p_{k+1}+\ldots+p_n\le\frac{M(X)}{A}.$$

Сумма слева в этом неравенстве – это вероятность того, что $X > A$. Получаем неравенство Маркова

$$P(X > A) \le \frac{M(X)}{A}.$$

Случайные величины могут иметь различные распределения при одинаковых математических ожиданиях. Однако для них неравенство Маркова даст одинаковую оценку вероятности того или иного результата испытания.

Этот недостаток связан с общностью неравенства Маркова: добиться лучшей оценки сразу для всех случайных величин невозможно.

Пример. Сумма всех вкладов в некотором банке составляет 2 000 000 ден. ед., а вероятность того, что случайно взятый вклад не превысит 10 000 ден. ед., равна 0,8. Оценить число вкладчиков банка.

Решение. Пусть n – число вкладчиков, а случайная величина X описывает размер случайно выбранного вклада. Тогда средний размер вклада $M(X) = \frac{2000000}{n}$ ден. ед., и по неравенству Маркова

$$P(X \le 10000) \ge 1 - \frac{200}{n}.$$

По условию $P(X \le 10000) = 0{,}8$, откуда $0{,}8 \ge 1 - \frac{200}{n}$ и, значит, $n \le 1000$ человек.

Пример. Среднее количество вызовов, поступающих на коммутатор завода в течение часа, равно 300. Оценить вероятность того, что в течение следующего часа число вызовов на коммутатор: а) превысит 400; б) будет не более 500.

Решение.

а) По условию $M(X) = 300$. По формуле $P(X > A) \le \frac{M(X)}{A}$ получим $P(X > 400) \le \frac{300}{400}$, т.е. вероятность того, что число вызовов превысит 400, не более 0,75.

б) По формуле $P(X \le A) \ge 1 - \frac{M(X)}{A}$: $P(X \le 500) \ge 1 - \frac{300}{500} = 0{,}4$, т.е. вероятность того, что число вызовов не более 500, будет не менее 0,4.

Неравенство Чебышёва

Вероятность того, что отклонение случайной величины от ее математического ожидания превзойдет по абсолютной величине положительное число ε, удовлетворяет условию

$$P\left(\left|X-M(X)\right|>\varepsilon\right)\le\frac{D(X)}{\varepsilon^2}.$$

Действительно, применим неравенство Маркова к случайной величине $\left(X-M(X)\right)^2$ и числу $A=\varepsilon^2$:

$$P\left(\left(X-M(X)\right)^2>\varepsilon^2\right)\le\frac{M\left(X-M(X)\right)^2}{\varepsilon^2}$$

или

$$P\left(\left|X-M(X)\right|>\varepsilon\right)\le\frac{D(X)}{\varepsilon^2}.$$

Следствие. $P\left(\left|X-M(X)\right|\le\varepsilon\right)\ge1-\frac{D(X)}{\varepsilon^2}$.

Пример. С помощью неравенства Чебышёва можно оценить вероятность $P\left(\left|X-M(X)\right|\le3\sigma\right)$ для произвольной случайной величины с конечным математическим ожиданием и конечной дисперсией.

Решение.

$$P\left(\left|X-M(X)\right|\le3\sigma\right)\ge1-\frac{\sigma^2}{9\sigma^2}$$

или

$$P\left(\left|X-M(X)\right|\le3\sigma\right)\ge0{,}8889.$$

Замечание. Для нормального распределения эта вероятность равна 0,9973, для равномерного 1, для показательного 0,9827.

Пример. Для новогоднего праздника Петя должен сделать гирлянду из 400 электрических лампочек. Он решает включить их параллельно. Лампочки оказались очень низкого качества – вероятность того, что какая-либо из них погаснет во время праздника, составляет 0,5. С помощью неравенства Чебышёва оценить вероятность того, что число горящих лампочек будет заключено между 100 и 300.

Решение.

$$P(100\le X\le300)=P\left(\left|X-200\right|\le100\right)\ge1-\frac{400\cdot0{,}5\cdot0{,}5}{100^2},$$

$$P\left(\left|X-200\right|\le100\right)\ge1-0{,}001,$$

$$P\left(\left|X-200\right|\le100\right)\ge0{,}999.$$

Пример. Вероятность выхода с автомата стандартной детали равна 0,96. Оценить с помощью неравенства Чебышёва вероятность того, что число бракованных деталей среди 2000 деталей находится в границах от 60

до 100 (включительно). Уточнить вероятность того же события с помощью интегральной теоремы Муавра–Лапласа.

Решение. По условию вероятность того, что деталь бракованная, равна $p = 1 - 0{,}96 = 0{,}04$ Случайная величина X – число бракованных деталей – имеет биномиальный закон распределения, а границы 60 и 100 симметричны относительно математического ожидания $M(X) = np = 2000 \cdot 0{,}04 = 80$.

Неравенство $60 \le k \le 100$ можно записать в виде $|k - 80| \le 20$.

Тогда $P(|k - 80| \le 20) \ge 1 - \dfrac{2000 \cdot 0{,}04 \cdot 0{,}96}{20^2} = 0{,}808$.

По следствию из интегральной теоремы Муавра–Лапласа

$$P(|k - 80| \le 20) \approx 2\Phi\left(\frac{20}{\sqrt{76{,}8}}\right) = 2\Phi(2{,}28) = 0{,}979.$$

Полученный результат не противоречит оценке, найденной с помощью неравенства Чебышёва. Различие результатов объясняется тем, что неравенство Чебышёва дает лишь нижнюю границу оценки вероятности искомого события для любой случайной величины, а интегральная теорема Муавра–Лапласа дает достаточно точное значение самой вероятности, т.к. она применима лишь для случайной величины, имеющий определенный (биномиальный) закон распределения.

Пример. По статистическим данным в среднем 87% новорожденных доживает до 50 лет.

1. Найти вероятность того, что из 1000 новорожденных доля доживших до 50 лет будет: а) заключена в пределах от 0,9 до 0,95;

б) будет отличаться от вероятности этого события не более, чем на 0,04 (по абсолютной величине).

2. При каком числе новорожденных с надежностью 0,95 доля доживших до 50 лет будет заключена в границах от 0,86 до 0,88?

3. С помощью неравенства Чебышёва оценить вероятность того, что из 1000 новорожденных доля доживших до 50 лет будет отличаться от вероятности этого события не более, чем на 0,04 (по абсолютной величине).

Решение. 1. а) Вероятность p того, что новорожденный доживет до 50 лет, равна 0,87. Т.к. число $n = 1000$ велико, то используем интегральную теорему Муавра–Лапласа:

$$P_n(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right),$$

$$P_n\left(\frac{k_1}{n} \le \frac{k}{n} \le \frac{k_2}{n}\right) = P_n\left(p_1 \le \frac{k}{n} \le p_2\right) \approx \Phi\left(\frac{np_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{np_1 - np}{\sqrt{npq}}\right) =$$

$$=\Phi\left(\frac{p_2-p}{\sqrt{pq/n}}\right)-\Phi\left(\frac{p_1-p}{\sqrt{pq/n}}\right),$$

$$P_{1000}\left(0{,}9\le\frac{k}{1000}\le 0{,}95\right)\approx\Phi\left(\frac{0{,}95-0{,}87}{\sqrt{0{,}87\cdot 0{,}13/1000}}\right)-\Phi\left(\frac{0{,}9-0{,}87}{\sqrt{0{,}87\cdot 0{,}13/1000}}\right)\approx$$

$$\approx\Phi(7{,}52)-\Phi(2{,}82)\approx 0{,}0024.$$

б) По формуле $P\left(\left|\frac{k}{n}-p\right|\le\varepsilon\right)\approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$:

$$P\left(\left|\frac{k}{1000}-0{,}87\right|\le 0{,}04\right)\approx 2\Phi\left(0{,}04\sqrt{\frac{1000}{0{,}87\cdot 0{,}13}}\right)\approx 2\Phi(3{,}76)\approx 0{,}9998.$$

$$\left|\frac{k}{1000}-0{,}87\right|\le 0{,}04\Leftrightarrow 0{,}83\le\frac{k}{1000}\le 0{,}91.$$

Т.е. практически достоверно, что от 830 до 910 новорожденных из 1000 доживут до 50 лет.

2. По условию $P_n\left(0{,}86\le\frac{k}{n}\le 0{,}88\right)=0{,}95$, или

$$P_n\left(-0{,}01\le\frac{k}{n}-0{,}87\le 0{,}01\right)=P_n\left(\left|\frac{k}{n}-0{,}87\right|\le 0{,}01\right)\approx$$

$$\approx 2\Phi\left(0{,}01\sqrt{\frac{n}{0{,}87\cdot 0{,}13}}\right)=0{,}95.$$

По таблице $0{,}01\sqrt{\frac{n}{0{,}87\cdot 0{,}13}}=1{,}96$. Отсюда $n=4345$.

3. $P\left(\left|\frac{k}{1000}-0{,}87\right|\le 0{,}04\right)\ge 1-\frac{0{,}87\cdot 0{,}13}{1000\cdot 0{,}04^2}=0{,}929$.

Закон больших чисел

Закон больших чисел – это

1) принцип, согласно которому частота финансовых потерь определенного вида может быть предсказана с высокой точностью тогда, когда есть большое количество данных о потерях аналогичных видов;

2) закон, согласно которому общая закономерность не может быть выявлена на основе частных проявлений, случаев, а только на основании большого количества наблюдений, исключающих случайные факторы.

(Райзберг Б.А., Лозовский Л.Ш., Стародубцева Е.Б. Современный экономический словарь).

В практических исследованиях очень важно знать, в каких случаях можно гарантировать, что вероятность события будет или достаточно мала, или как угодно близка к единице.

Под законом больших чисел понимается совокупность предложений, в которых утверждается, что с вероятностью, как угодно близкой к единице, произойдет событие, зависящее от очень большого, неограниченно увеличивающегося числа случайных факторов, каждый из которых оказывает на него лишь незначительное влияние.

Отдельные, единичные явления, которые мы наблюдаем в природе и в общественной жизни, часто проявляются как случайные (например, регистрируемый смертный случай, пол родившегося ребенка, температура воздуха и др.) вследствие того, что на такие явления действует много факторов, не связанных с существом возникновения или развития явления.

Предсказать суммарное действие их на наблюдаемое явление нельзя, и они различно проявляются в единичных явлениях. По результатам одного явления нельзя ничего сказать о закономерностях, присущих многим таким явлениям.

Давно было замечено, что средняя арифметическая числовых характеристик некоторых признаков (относительные частоты появления события, результатов измерений и т. д.) при большом числе повторений опыта подвержена очень незначительным колебаниям. В средней проявляется закономерность, присущая существу явлений, в ней взаимно погашается влияние отдельных факторов, которые делали случайными результаты единичных наблюдений.

Теоретически объяснить такое поведение средней можно с помощью *закона больших чисел*. Если будут выполнены некоторые условия относительно случайных величин, то устойчивость средней арифметической будет практически достоверным событием.

Эти условия и составляют наиболее важное содержание закона больших чисел.

Теорема Чебышёва. *Если дисперсии n независимых случайных величин X_1, X_2, ..., X_n ограничены одной и той же постоянной, то при неограниченном увеличении числа n средняя арифметическая этих случайных величин сходится по вероятности к средней арифметической их математических ожиданий a_1, a_2, ..., a_n:*

$$P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{a_1+a_2+\ldots+a_n}{n}\right|<\varepsilon\right)\to 1 \text{ при } n\to\infty.$$

Замечание. Сходимость по вероятности не означает, что неравенство $\left|\frac{X_1+X_2+...+X_n}{n}-\frac{a_1+a_2+...+a_n}{n}\right|<\varepsilon$ будет выполняться при всех n, даже начиная с некоторого значения. Возможно, что в отдельных случаях неравенство выполняться не будет. Но при увеличении n оно будет практически достоверно.

Следствие. Если дисперсии n независимых случайных величин X_1, X_2, ..., X_n ограничены одной и той же постоянной, а математические ожидания у всех случайных величин одинаковы и равны a, то

$$P\left(\left|\frac{X_1+X_2+...+X_n}{n}-a\right|<\varepsilon\right)\to 1 \text{ при } n\to\infty.$$

Например, страховой компании необходимо установить размер страхового взноса, который должен уплачивать страхователь. Рассматривая убытки страхователя как величину случайную, и имея статистику страховых случаев, можно определить средние убытки, которые на основании теоремы Чебышёва с большой степенью уверенности можно считать величиной почти не случайной.

Или, например, надо измерить некоторую величину, истинное значение которой равно a. Проводят n независимых измерений. Результат каждого измерения – случайная величина. Если систематические ошибки отсутствуют, то средняя арифметическая результатов измерений сходится по вероятности к истинному значению.

Теорема Бернулли. *Если проводится n независимых испытаний и вероятность случайного события A* $P(A)=p$ *в каждом испытании, то относительная частота* $\frac{m}{n}$ *появления события A (m – число появлений A) при большом n приближенно равна вероятности p:* $\frac{m}{n}\approx p$ *или*

$$P\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)\to 1 \text{ при } n\to\infty. \qquad (*)$$

Эта теорема не утверждает, что соотношение (*) достоверно, однако, если n достаточно велико, то вероятность его выполнения близка к 1 (например, 0,98 или 0,999), что практически достоверно.

Если мы собираемся провести эксперимент, состоящий из достаточно большого числа n испытаний, то мы можем быть практически уверены, что соотношение (*) будет выполнено.

Теорема Бернулли дает теоретическое обоснование замены неизвестной вероятности события его частотой, или статистической вероятностью, по-

лученной в n повторных независимых испытаниях, проводимых при одном и том же комплексе условий.

В целях экспериментальной проверки закона больших чисел в разное время были произведены следующие опыты.

1. Опыт Бюффона. Монета брошена 4040 раз. Герб выпал 2048 раз. Частота его выпадения оказалась равной 0,50694.

2. Опыт Пирсона. Монета брошена 12 000 и 24 000 раз. Частота выпадения герба в первом случае оказалась равной 0,5016, во втором – 0,5005.

3. Опыт Вестергаарда. Из урны, в которой было поровну белых и черных шаров, получено при 10 000 извлечений (с возвратом очередного вынутого шара в урну) 5011 белых и 4989 черных шаров. Частота белых шаров составила 0,50110, а черных – 0,49890.

4. Опыт В.И. Романовского. Четыре монеты брошены 21160 раз. Частоты различных комбинаций выпадения герба и решки распределились следующим образом (табл. 40).

Таблица 40

Комбинации числа выпадений герба и решки	Количество выпадений	Частоты	
		Эмпирические	Теоретические
4 и 0	1181	0,05858	0,0625
3 и 1	4909	0,24350	0,2500
2 и 2	7583	0,37614	0,3750
1 и 3	5085	0,25224	0,2500
1 и 4	1402	0,06954	0,0625
Итого	20160	1,0000	1,0000

Результаты экспериментальных проверок закона больших чисел убеждают нас в большой близости опытных частот к вероятностям.

Системы двух случайных величин

Часто результат опыта описывается не одной случайной величиной, а несколькими.

Двумерной называют случайную величину $(X; Y)$, возможные значения которой есть пары чисел $(x; y)$. Составляющие X и Y рассматриваемые одновременно, образуют систему двух случайных величин.

Дискретной называют двумерную случайную величину, составляющие которой дискретны. *Непрерывной* называют двумерную случайную величину, составляющие которой непрерывны.

Законом распределения дискретной двумерной случайной величины называют перечень возможных значений этой величины, т.е. пар чисел $(x_i; y_j)$ и их вероятностей

$$p(x_i, y_j) = P((X = x_i)(Y = y_j)) \ (i = 1, 2, ..., n; j = 1, 2, ..., m).$$

Обычно закон распределения задают в виде таблицы с двойным входом.

Пример. Найти законы распределения составляющих двумерной случайной величины, заданной законом распределения (табл. 41):

Таблица 41

Y	X		
	x_1	x_2	x_3
y_1	0,10	0,30	0,20
y_2	0,06	0,18	0,16

Решение. Сложив вероятности по столбцам, получим вероятности возможных значений X: $P(x_1) = 0{,}16$, $P(x_2) = 0{,}48$, $P(x_3) = 0{,}36$. Напишем закон распределения составляющей X (табл. 42).

Таблица 42

X	x_1	x_2	x_3
P	0,16	0,48	0,36

Сложив вероятности по строкам, получим вероятности возможных значений Y: $P(y_1) = 0{,}6, P(y_2) = 0{,}4$. Напишем закон распределения составляющей Y (табл. 43):

Таблица 43

Y	y_1	y_2
P	0,6	0,4

Пример. В двух ящиках находятся по шесть шаров; в первом ящике один шар – с №1, 2 шара – с №2, 3 шара – с №3; во втором ящике: 2 шара – с №1, 3 шара – с №2, 1 шар – с №3. Пусть X – номер шара, вынутого из первого ящика, Y – номер шара, вынутого из второго ящика. Из каждого ящика вынули по шару. Составить закон распределения системы случайных величин $(X; Y)$.

Решение. Вероятность появления шаров №1 из первого и из второго ящиков, равна: $\frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18}$.

Вероятность появления шара №1 из первого ящика и шара №2 из второго ящика равна $\frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$.

Аналогичным образом находим вероятности остальных сочетаний шаров (табл. 44).

Таблица 44

Y	X		
	1	2	3
1	1/18	1/9	1/6
2	1/12	1/6	1/4
3	1/36	1/18	1/12

Пример. Из коробки, в которой находятся 2 белых и 3 черных шара, вынимают два шара. Пусть X – число вынутых белых шаров, а случайная величина Y определяется следующим образом:

$$Y = \begin{cases} 0, \text{ если вынут хотя бы один черный шар,} \\ 1, \text{ в остальных случаях.} \end{cases}$$

Составить закон распределения случайной величины $(X; Y)$.

Решение. Поскольку $p_{11} = P(X = 0; Y = 0)$ вероятность того, что белых шаров не вынуто (вынуто два черных шара), то $p_{11} = \dfrac{C_2^0 C_3^2}{C_5^2} = \dfrac{3}{10}$.

$p_{12} = P(X = 1; Y = 0)$ – вероятность того, что вынут один белый шар (и, значит, один черный), тогда $p_{12} = \dfrac{C_2^1 C_3^1}{C_5^2} = \dfrac{6}{10}$.

$p_{13} = P(X = 2; Y = 0) = 0$.

$p_{21} = P(X = 0; Y = 1) = 0$ – вероятность того, что белых шаров не вынуто (и, значит, вынуто два черных шара).

$p_{22} = P(X = 1; Y = 1) = 0$ – вероятность того, что вынут один белый шар (и, значит, один черный).

$p_{23} = P(X = 2; Y = 1)$ – вероятность того, что вынуто два белых шара (и, значит, ни одного черного), тогда $p_{23} = \dfrac{C_2^2 C_3^0}{C_5^2} = \dfrac{1}{10}$.

Таким образом, ряд распределения двумерной случайной величины $(X; Y)$ имеет вид (табл. 45):

Таблица 45

Y	X		
	0	1	2
0	0,3	0,6	0
1	0	0	0,1

Пусть составляющие X и Y дискретны и имеют соответственно следующие возможные значения: $x_1, x_2, ..., x_n$; $y_1, y_2, ..., y_m$.

Условным распределением составляющей X при $Y = y_j$ называют совокупность условных вероятностей $p(x_1|y_j)$, $p(x_2|y_j)$, ..., $p(x_n|y_j)$, где $p(x_i|y_j) = P_{(Y=y_j)}(X = x_i)$, $i = 1, 2, ..., n$.

Аналогично определяется условное распределение Y.

Условные вероятности составляющих X и Y вычисляют по формулам

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}, \quad p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}.$$

Эти формулы следуют из теоремы умножения вероятностей: $P_A(B) = \dfrac{P(AB)}{P(A)}$.

Пример. Дискретная двумерная случайная величина задана табл. 46.

Таблица 46

Y	X	
	3	6
10	0,25	0,10
14	0,15	0,05
18	0,32	0,13

Найти: а) условный закон распределения X при условии, что $Y = 10$; б) условный закон распределения Y при условии, что $X = 6$.

Решение.

а) Найдем условный закон распределения X при условии, что Y приняла значение $y_1 = 10$:

$$p(x_1|y_1) = \frac{p(x_1, y_1)}{p(y_1)} = \frac{0,25}{0,35} = \frac{5}{7}; \quad p(x_2|y_1) = \frac{p(x_2, y_1)}{p(y_1)} = \frac{0,10}{0,35} = \frac{2}{7}.$$

Напишем искомый условный закон распределения X (табл. 47).

Таблица 47

X	3	6
$p(X\|y_1)$	$\frac{5}{7}$	$\frac{2}{7}$

б) Аналогично найдем условный закон распределения Y при условии, что X приняла значение $x_2 = 6$:

$$p(y_1|x_2) = \frac{p(x_2, y_1)}{p(x_2)} = \frac{0,10}{0,28} = \frac{5}{14}, \quad p(y_2|x_2) = \frac{p(x_2, y_2)}{p(x_2)} = \frac{0,05}{0,28} = \frac{5}{28},$$

$$p(y_3|x_2)=\frac{p(x_2,y_3)}{p(x_2)}=\frac{0{,}13}{0{,}28}=\frac{13}{28}.$$

Получаем закон распределения (табл. 48).

Таблица 48

Y	10	14	18
$p(Y\|x_2)$	$\frac{5}{14}$	$\frac{5}{28}$	$\frac{13}{28}$

Функция распределения системы двух случайных величин

Пусть дана двумерная случайная величина (X, Y). Функцией распределения называется функция, равная $F(x, y)=P(X<x, Y<y)$.

В точке $M(x, y)$ функция распределения равна вероятности попадания значения случайной величины в фигуру, изображенную на рис. 18.

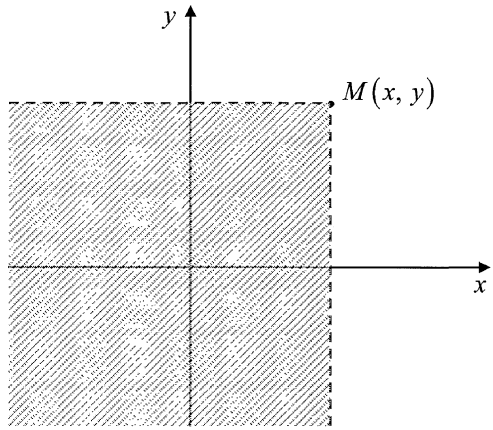


Рис. 18

Свойства функции распределения

1. $0\le F(x, y)\le 1$ при всех $M(x, y)\in R^2$.
2. Функция является неубывающей по каждому аргументу.
3. $\lim\limits_{x\to-\infty} F(x, y)=\lim\limits_{y\to-\infty} F(x, y)=0$.
4. $\lim\limits_{y\to+\infty} F(x, y)=P(X<x, Y<+\infty)=P(X<x)=F_1(x)$,

$$\lim_{x\to+\infty} F(x, y)=P(X<+\infty, Y<y)=P(Y<y)=F_2(y).$$

5. $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(x, y)=1$.

6. Вероятность попадания двумерной случайной величины в прямоугольник, изображенный на рис. 19, равна:

$$P(x_1 \le X < x_2,\ y_1 \le Y < y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)\ .$$

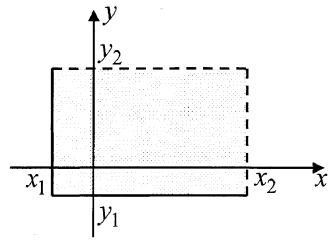


Рис. 19

Пример. Функция распределения двумерной случайной величины

$$F(x, y) = \sin x \cdot \sin y \text{ при } 0 \le x \le \frac{\pi}{2},\ 0 \le y \le \frac{\pi}{2}.$$

Найти вероятность попадания случайной точки (X, Y) в прямоугольник, ограниченный прямыми $x=0, x=\frac{\pi}{4}, y=\frac{\pi}{6}, y=\frac{\pi}{3}$.

Решение. Используем формулу:

$$P(x_1 < X < x_2,\ y_1 < Y < y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

Положив в этой формуле $x_1 = 0, x_2 = \frac{\pi}{4}, y_1 = \frac{\pi}{6}, y_2 = \frac{\pi}{3}$ получим:

$$P = \sin\frac{\pi}{4}\cdot\sin\frac{\pi}{3} - \sin 0\cdot\sin\frac{\pi}{3} - \sin\frac{\pi}{4}\cdot\sin\frac{\pi}{6} + \sin 0\cdot\sin\frac{\pi}{6} = \frac{\sqrt{6}-\sqrt{2}}{4} \approx 0{,}26\ .$$

Плотность вероятности двумерной непрерывной случайной величины

Функцией плотности вероятности двумерной случайной величины называется функция

$$f(x, y) = F''_{xy}(x, y).$$

Свойства функции плотности

1. $f(x, y) \ge 0$ на всей плоскости; **2.** $\iint\limits_{R^2} f(x, y)\,dxdy = 1$.

Пример. Найти плотность совместного распределения $f(x, y)$ системы случайных величин (X, Y), если задана функция распределения:

$$F(x, y) = \begin{cases} \left(1 - e^{-4x}\right) \cdot \left(1 - e^{-2y}\right), & x \geq 0, y \geq 0, \\ 0, & x < 0 \text{ или } y < 0. \end{cases}$$

Решение. Используем формулу $f(x, y) = \dfrac{\partial^2 F}{\partial x \partial y}$. Найдем частные производные: $\dfrac{\partial F}{\partial x} = 4e^{-4x} \cdot \left(1 - e^{-2y}\right)$; $\dfrac{\partial^2 F}{\partial x \partial y} = 8e^{-4x} \cdot e^{-2y}$.

Итак, искомая двумерная плотность вероятности

$$f(x, y) = \begin{cases} 8e^{-4x} e^{-2y}, & x \geq 0, y \geq 0, \\ 0, & x < 0 \text{ или } y < 0. \end{cases}$$

Законы распределения составляющих X и Y двумерной случайной величины (X, Y) можно найти по формулам:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy \text{ для случайной величины } X \text{ и}$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx \text{ для случайной величины } Y.$$

Условные законы распределения составляющих:

$$f_y(x) = \frac{f(x, y)}{f_2(y)} \text{ для случайной величины } X \text{ при условии } Y = y,$$

$$f_x(y) = \frac{f(x, y)}{f_1(x)} \text{ для случайной величины } Y \text{ при условии } X = x.$$

Зависимость случайных величин

Пусть дана двумерная случайная величина (X, Y), распределение которой известно. Тогда можно найти математические ожидания $M(X) = a_X$, $M(Y) = a_Y$ и дисперсии $D(X) = \sigma_X^2$, $D(Y) = \sigma_Y^2$ одномерных составляющих X и Y.

Ковариацией (или *корреляционным моментом*) K_{XY} случайных величин X и Y называется математическое ожидание произведения отклонений этих величин от своих математических ожиданий, т.е.

$$K_{XY} = M\left[(X - a_X)(Y - a_Y)\right].$$

Свойства ковариации

1) Если X и Y независимы, то $K_{XY}=0$.

2) $K_{XY}=M(XY)-a_X a_Y$.

3) $|K_{XY}|\le\sigma_X\sigma_Y$.

Замечание. Для дискретной случайной величины (X,Y):

$$M(XY)=\sum_{i=1}^{n}\sum_{j=1}^{m}x_i y_j p(x_i,y_j).$$

Коэффициентом корреляции двух случайных величин называется отношение их ковариации к произведению средних квадратических отклонений, т.е. $\rho_{XY}=\dfrac{K_{XY}}{\sigma_X\sigma_Y}$.

Свойства коэффициента корреляции

1) $-1\le\rho_{XY}\le 1$.

2) Если X и Y независимы, то $\rho_{XY}=0$.

3) Если коэффициент корреляции двух случайных величин равен по абсолютной величине 1, то между случайными величинами существует линейная функциональная зависимость. Причем при $\rho_{XY}=1$: $Y=a_Y+\dfrac{\sigma_Y}{\sigma_X}(X-a_X)$, при $\rho_{XY}=-1$: $Y=a_Y-\dfrac{\sigma_Y}{\sigma_X}(X-a_X)$.

Пример. Закон распределения дискретной двумерной случайной величины задан табл. 49. Найти коэффициент корреляции ρ_{XY}.

Таблица 49

X	Y	
	0	1
0	$\frac{1}{2}$	$\frac{1}{4}$
1	$\frac{1}{8}$	$\frac{1}{8}$

Решение. Найдем законы распределения составляющих (табл. 50, 51).

Таблица 50

X	0	1
P	$\frac{3}{4}$	$\frac{1}{4}$

Таблица 51

Y	0	1
P	$\frac{5}{8}$	$\frac{3}{8}$

Математические ожидания компонент: $M(X)=\dfrac{1}{4}$, $M(Y)=\dfrac{3}{8}$.

$$M\left(X^2\right)=\frac{1}{4},\ M\left(Y^2\right)=\frac{3}{8}.$$

Дисперсии компонент: $D(X)=\frac{3}{16},\ D(Y)=\frac{15}{64}$.

$$M(XY)=\frac{1}{8},\ K_{XY}=\frac{1}{8}-\frac{1}{4}\cdot\frac{3}{8}=\frac{1}{32}.$$

Средние квадратические отклонения компонент:

$$\sigma(X)=\sqrt{\frac{3}{16}}=\frac{\sqrt{3}}{4},\ \sigma(Y)=\sqrt{\frac{15}{64}}=\frac{\sqrt{15}}{8}.$$

Коэффициент корреляции: $\rho_{XY}=\frac{K_{XY}}{\sigma_X\sigma_Y}=\frac{\frac{1}{32}}{\frac{\sqrt{3}}{4}\cdot\frac{\sqrt{15}}{8}}=\frac{1}{\sqrt{5}}\approx 0{,}15$.

Зависимость между двумя случайными величинами называется *вероятностной* (*стохастической*), если каждому значению одной из них соответствует определенное распределение другой.

В случае вероятностной зависимости между случайными величинами нельзя, зная значение одной из них, точно определить значение другой. Можно лишь указать распределение другой случайной величины.

Например, вероятностная зависимость существует между:

1) числом отказов оборудования и затратами на его профилактический ремонт;

2) ростом и весом человека;

3) затратами времени студента на выполнение домашнего задания по математике и оценкой за экзамен.

Независимость X и Y можно проверить по одному из условий:

1) $F(x,y)=F_1(x)F_2(y)$;

2) $f(x,y)=f_1(x)f_2(y)$;

3) $f_y(x)=f_1(x)$ и $f_x(y)=f_2(y)$.

Пусть найден закон распределения Y при условии, что $X=x$. Можно вычислить математическое ожидание построенной случайной величины (условное математическое ожидание) $M_x(Y)$.

Оно является функцией от x, т.е. $M_x(Y)=\varphi(x)$.

Эта функция называется *функцией регрессии* Y на X.

Аналогично, функция регрессии X на Y – это $M_y(X)=\varphi(y)$.

Графики этих функций называются *линиями регрессии*.

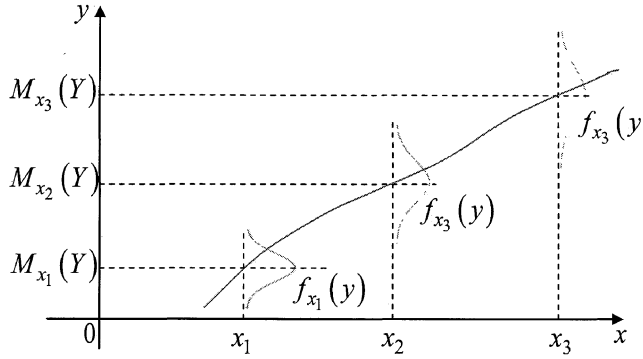


Рис. 20

На рис. 20 зависимость случайных величин проявляется в том, что с изменением x меняется распределение Y и $M_x(Y)$.

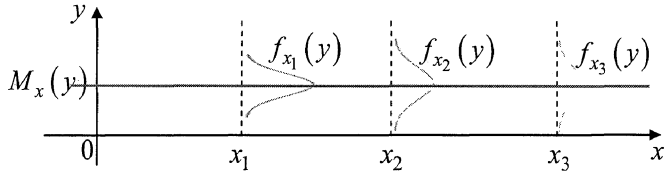


Рис. 21

На рис. 21 зависимость проявляется в изменении условных дисперсий, а $M_x(Y) = const$.

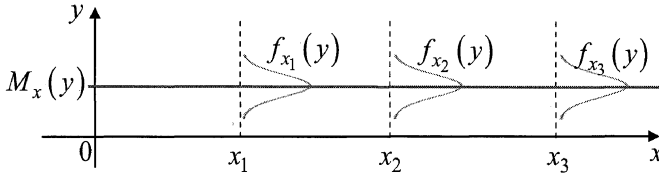


Рис. 22

Случайные величины независимы (рис. 22).

Если случайные величины независимы, то линии регрессии параллельны координатным осям.

Пример. Найти условное математическое ожидание составляющей Y при $X = x_1 = 1$ для дискретной двумерной случайной величины, заданной табл. 52:

Таблица 52

Y	X			
	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,30	0,10	0,03	0,07

Решение. $p(x_1) = P(X = x_1) = 0{,}15 + 0{,}30 = 0{,}45$,

$$p(y_1 \mid x_1) = \frac{p(x_1, y_1)}{p(x_1)} = \frac{0{,}15}{0{,}45} = \frac{1}{3},\ p(y_2 \mid x_1) = \frac{p(x_1, y_2)}{p(x_1)} = \frac{0{,}30}{0{,}45} = \frac{2}{3},$$

$$M(Y \mid X = x_1) = \sum_{j=1}^{2} y_j \cdot p(y_2 \mid x_1) = \frac{3}{3} + \frac{12}{3} = 5.$$

Контрольные вопросы

1. Какая случайная величина называется непрерывной?
2. Приведите определение и свойства функции плотности распределения непрерывной случайной величины.
3. Приведите формулы для нахождения математического ожидания и дисперсии непрерывной случайной величины.
4. Дайте определение случайной величины, распределенной по равномерному закону, запишите формулы и постройте графики функции распределения и функции плотности распределения, приведите формулы для нахождения ее математического ожидания, дисперсии и вероятности попадания в заданный интервал.
5. Дайте определение случайной величины, распределенной по показательному закону, запишите формулы и постройте графики функции распределения и функции плотности распределения, приведите формулы для нахождения ее математического ожидания, дисперсии и вероятности попадания в заданный интервал.
6. Дайте определение случайной величины, распределенной по нормальному закону, запишите формулы и постройте графики функции распределения и функции плотности распределения, приведите формулы для нахождения ее математического ожидания, дисперсии и вероятности попадания в заданный интервал.
7. Сформулируйте правило трех сигм для нормально распределенной случайной величины.
8. Запишите неравенство Маркова.
9. Запишите неравенство Чебышёва.

10. Что называется законом больших чисел?

11. Как задается система двух случайных величин?

12. Как проверить независимость двух случайных величин?

13. Приведите определение и свойства ковариации и коэффициента корреляции.

ФУНКЦИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Понятие функции случайной величины

Рассмотрим задачу о законе распределения функции одного случайного аргумента $Y=\varphi(X)$.

Пусть X – дискретная случайная величина, имеющая ряд распределения (табл. 53).

Таблица 53

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Тогда $Y=\varphi(X)$ также дискретная случайная величина с возможными значениями $y_1=\varphi(x_1)$, $y_2=\varphi(x_2)$, …, $y_n=\varphi(x_n)$. Если все значения y_1, y_2, …, y_n различны, то для каждого $k=1, 2, ..., n$ события $(X=x_k)$ и $Y=y_k=\varphi(x_k)$ тождественны.

Следовательно, $P(Y=y_k)=P(X=x_k)=p_k$.

Если среди чисел $y_1=\varphi(x_1)$, $y_2=\varphi(x_2)$, …, $y_n=\varphi(x_n)$ есть одинаковые, то каждой группе одинаковых значений $y_k=\varphi(x_k)$ нужно отвести в таблице один столбец и соответствующие вероятности сложить.

Для непрерывных случайных величин, зная плотность распределения $f(x)$ случайной величины X, можно найти плотность распределения $g(y)$ случайной величины $Y=\varphi(X)$.

Рассмотрим два случая.

I. Предположим, что функция $y=\varphi(x)$ является монотонно возрастающей, непрерывной и дифференцируемой на интервале $(a; b)$, на котором лежат все возможные значения величины X.

Тогда обратная функция $x=\psi(y)$ существует, при этом являясь также монотонно возрастающей, непрерывной и дифференцируемой. В этом случае получаем $g(y)=f(\psi(y))\cdot|\psi'(y)|$.

Пример. Случайная величина X распределена с плотностью $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Найти закон распределения случайной величины $Y=X^3$

Решение. Т.к. функция $y=x^3$ монотонна на промежутке $(-\infty;+\infty)$, то можно применить формулу $g(y)=f(\psi(y))\cdot|\psi'(y)|$.

Обратная функция по отношению к функции $\varphi(x)=x^3$ есть $\psi(y)=\sqrt[3]{y}$, ее производная $\psi'(y)=\frac{1}{3\sqrt[3]{y^2}}$. Следовательно,

$$g(y)=\frac{1}{3\sqrt{2\pi}\sqrt[3]{y^2}}e^{-\frac{\sqrt[3]{y^2}}{2}}.$$

II. Пусть функция $y=\varphi(x)$ такова, что обратная функция $x=\psi(y)$ неоднозначна, т.е. одному значению y соответствует несколько значений аргумента x, которые обозначим $x_1=\psi_1(y)$, $x_2=\psi_2(y)$, ..., $x_n=\psi_n(y)$, где n – число участков, на которых функция $y=\varphi(x)$ изменяется монотонно. Тогда

$$g(y)=\sum_{k=1}^{n}f(\psi_k(y))\cdot|\psi'_k(y)|.$$

Пример. Случайная величина X распределена с плотностью $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Найти закон распределения случайной величины $Y=X^2$.

Решение. Обратная функция неоднозначна. Одному значению y соответствует два значения x: $x_1=\psi_1(y)=\sqrt{y}$, $x_2=\psi_2(y)=-\sqrt{y}$.

Применяя формулу $g(y)=f(\psi_1(y))\cdot|\psi'_1(y)|+f(\psi_2(y))\cdot|\psi'_2(y)|$, получаем:

$$g(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(-\sqrt{y})^2}{2}}\cdot\left|-\frac{1}{2\sqrt{y}}\right|+\frac{1}{\sqrt{2\pi}}e^{-\frac{(\sqrt{y})^2}{2}}\cdot\left|\frac{1}{2\sqrt{y}}\right|=\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}.$$

Математическое ожидание функции случайной величины

На практике часто встречаются случаи, когда нет необходимости определять закон распределения функции случайных величин, а достаточно только указать его числовые характеристики.

Пусть случайная величина Y является функцией случайного аргумента X с заданным законом распределения $Y=\varphi(X)$.

Требуется, не находя закона распределения величины Y, определить ее математическое ожидание $M(Y)=M(\varphi(X))$.

Пусть X – дискретная случайная величина, имеющая ряд распределения (табл. 54).

Таблица 54

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Тогда математическое ожидание случайной величины Y можно определить по формуле

$$M(\varphi(X))=\sum_{i=1}^{n}\varphi(x_i)p_i.$$

Эта формула не содержит в явном виде закон распределения самой функции $\varphi(X)$, а содержит только закон распределения аргумента X. Таким образом, для определения математического ожидания функции $Y=\varphi(X)$ достаточно знать закон распределения аргумента X.

Для непрерывной случайной величины математическое ожидание вычисляется по формуле

$$M(\varphi(X))=\int_{-\infty}^{+\infty}\varphi(x)f(x)dx,$$

где $f(x)$ – плотность распределения вероятностей случайной величины.

Дисперсия функции случайной величины

По определению дисперсии имеем $D(Y)=M\left((Y-M(Y))^2\right)$. Тогда

$$D(\varphi(X))=M\left((\varphi(X)-M(\varphi(X)))^2\right) \text{ или}$$

$$D(\varphi(X))=M\left((\varphi(X))^2\right)-\left(M(\varphi(X))\right)^2.$$

Для дискретной случайной величины

$$M\left((\varphi(X))^2\right)=\sum_{i=1}^{n}(\varphi(x_i))^2 p_i.$$

Для непрерывной случайной величины

$$M\left(\left(\varphi(X)\right)^2\right)=\int_{-\infty}^{+\infty}\varphi^2(x)f(x)dx.$$

Пример. Непрерывная случайная величина X задана плотностью распределения $f(x)=\sin x$ в интервале $\left(0;\frac{\pi}{2}\right)$. Найти математическое ожидание и дисперсию случайной величины $Y=\varphi(X)=X^2$.

Решение.

$$M(Y)=\int_{-\infty}^{+\infty}\varphi(x)f(x)dx=\int_0^{\frac{\pi}{2}}x^2\sin xdx=\begin{bmatrix} u=x^2 & du=2xdx \\ dv=\sin xdx & v=-\cos x\end{bmatrix}=$$

$$=-x^2\cos x\Big|_0^{\frac{\pi}{2}}+\int_0^{\frac{\pi}{2}}2x\cos xdx=\begin{bmatrix} u=2x & du=2dx \\ dv=\cos xdx & v=\sin x\end{bmatrix}=$$

$$=2x\sin x\Big|_0^{\frac{\pi}{2}}-\int_0^{\frac{\pi}{2}}2\sin xdx=\pi+2\cos x\Big|_0^{\frac{\pi}{2}}=\pi-2;$$

$$M\left(Y^2\right)=\int_{-\infty}^{+\infty}\varphi^2(x)f(x)dx=\int_0^{\frac{\pi}{2}}x^4\sin xdx=\ldots=$$

$$\frac{\pi^3}{2}-12\int_0^{\frac{\pi}{2}}x^2\sin xdx=\frac{\pi^3}{2}-12(\pi-2);$$

$$D(Y)=M\left(Y^2\right)-\left(M(Y)\right)^2=\frac{\pi^3}{2}-12(\pi-2)-(\pi-2)^2=$$

$$=\frac{\pi^3}{2}-12\pi+24-\pi^2+4\pi-4=\frac{\pi^3}{2}-\pi^2-8\pi+20.$$

Контрольные вопросы

1. Как задать функцию от дискретной случайной величины?
2. Как задать функцию от непрерывной случайной величины?
3. Приведите формулы для нахождения математического ожидания и дисперсии функции от случайной величины.

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Основные понятия

В практике статистических наблюдений различают два вида наблюдений: *сплошное*, когда изучают все объекты (элементы, наблюдения) совокупности, и несплошное, *выборочное*, когда изучается часть объектов. Примером сплошного наблюдения является перепись населения, охватывающая все население страны. Выборочными наблюдениями являются, например, проводимые социологические исследования.

Вся, подлежащая изучению совокупность объектов (наблюдений) называется *генеральной совокупностью*.

Понятие генеральной совокупности аналогично понятию случайной величины, т.к. обусловлена определенным комплексом условий.

Та часть объектов, которая отобрана для непосредственного изучения из генеральной совокупности, называется *выборочной совокупностью*, или *выборкой*.

Число объектов в генеральной совокупности называется *объемом генеральной совокупности*, число объектов в выборке – *объемом выборки*.

Сущность выборочного метода состоит в том, что по некоторой части генеральной совокупности (по выборке) делаются выводы о свойствах генеральной совокупности в целом.

Преимущества выборочного метода

1. Позволяет существенно экономить затраты ресурсов (материальных, трудовых, временных).

2. Является единственно возможным в случае бесконечной генеральной совокупности или в случае, когда исследование связано с уничтожением наблюдаемых объектов (например, исследование долговечности электрических лампочек, предельных режимов работы приборов и т.д.).

3. При тех же затратах ресурсов дает возможность проведения углубленного исследования за счет расширения программы исследования.

4. Позволяет снизить ошибки регистрации, т.е. расхождения между истинным и зарегистрированным значениями признака.

Основной *недостаток* выборочного метода – ошибки исследования, называемые *ошибками репрезентативности* (*представительства*).

Однако ошибки, возникающие при выборочном методе исследования в связи с изучением только части объектов могут быть заранее оценены и с помощью правильной организации выборки сведены к минимуму.

Чтобы по данным выборки иметь возможность судить о генеральной совокупности, она должна быть отобрана *случайно*. Случайность отбора элементов в выборку достигается соблюдением принципа равной возмож-

ности всем элементам генеральной совокупности быть отобранным в выборку.

На практике это достигается тем, что извлечение элементов в выборку проводится с помощью случайных чисел, имеющихся в специальных таблицах или вырабатываемых ЭВМ с помощью датчика случайных чисел.

Выборка называется *репрезентативной* (*представительной*), если она достаточно хорошо воспроизводит генеральную совокупность.

Существует *два способа образования выборки*:

– *повторный отбор*, когда каждый элемент, случайно отобранный и обследованный, возвращается в общую совокупность и может быть повторно отобран;

– *бесповторный отбор*, когда отобранный элемент не возвращается в общую совокупность.

Математическая теория выборочного метода основывается на анализе так называемой *собственно-случайной выборки*, в которую элементы отбираются случайным образом из всей генеральной совокупности. Такие выборки и будем рассматривать далее.

Обозначим:

x_i – значение признака (случайной величины X);

N и n – объемы генеральной и выборочной совокупностей;

N_i и n_i – число элементов генеральной и выборочной совокупностей со значением признака x_i $\left(N_1 + N_2 + ... + N_k = N,\ n_1 + n_2 + ... + n_k = n\right)$;

M и m – число элементов генеральной и выборочной совокупностей, обладающих данным признаком.

Средние арифметические распределения признака в генеральной и выборочной совокупностях называются соответственно *генеральной* и *выборочной средними*, а дисперсии этих распределений – *генеральной* и *выборочной дисперсиями*.

Отношение числа элементов генеральной и выборочной совокупностей, обладающих некоторым признаком, к их объемам, называются *генеральной* и *выборочной долями*, соответственно.

В табл. 55 приведены формулы для вычисления параметров генеральной и выборочной совокупностей.

Таблица 55

Наименование характеристики	Генеральная совокупность	Выборка
Средняя	$\bar{x}_0 = \frac{1}{N}\sum_{i=1}^{m} x_i N_i$	$\bar{x} = \frac{1}{n}\sum_{i=1}^{m} x_i n_i$
Дисперсия	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{m}\left(x_i - \bar{x}_0\right)^2 N_i$	$D = \frac{1}{n}\sum_{i=1}^{m} n_i\left(x_i - \bar{x}\right)^2$
Доля	$p = \frac{M}{N}$	$w = \frac{m}{n}$

Первоначальная обработка выборки

Для анализа и последующей обработки экспериментальных данных составляются *вариационные ряды* – последовательности наблюдаемых значений, записанных в возрастающем порядке.

Обозначим через $w_i = \frac{n_i}{n}$ относительные частоты значений признака.

Тогда данные можно записать в виде дискретного вариационного ряда (табл. 56 или 57).

Таблица 56

x_i	x_1	x_2	…	x_k
n_i	n_1	n_2	…	n_k

Таблица 57

x_i	x_1	x_2	…	x_k
w_i	w_1	w_2	…	w_k

Перечень элементов выборки и соответствующих им частот (или относительных частот) называют *статистическим распределением выборки*.

На практике большие выборки из непрерывных распределений подвергаются *группировке*.

Группировка производится следующим образом:

1) находятся $x_{\min}$ и $x_{\max}$ (минимальный и максимальный элементы выборки);

2) находится размах выборки $x_{\max} - x_{\min}$;

3) промежуток $\left[x_{\min}; x_{\max}\right]$ разбивается на m интервалов. Число частичных интервалов рекомендуется определить по формуле

$$m = 1 + 3{,}322 \cdot \lg n;$$

4) для каждого частичного интервала находится n_i $(i = 1, 2, ..., n)$ – число элементов, попавших в i-й интервал, длина каждого интервала равна

$$h = \frac{x_{\max} - x_{\min}}{m};$$

5) составляется таблица – интервальный вариационный ряд (табл. 58).

Таблица 58

$[x_1; x_2]$	n_1
$(x_2; x_3]$	n_2
...	...
$(x_{i-1}; x_i]$	n_i
...	...
$(x_{m-1}; x_m]$	n_m

Для наглядности строят график статистического распределения – гистограмму. *Гистограммой частот (относительных частот)* называют ступенчатую фигуру (рис. 23), состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной h, а высоты равны отношению $\frac{n_i}{h}$ (плотность частоты) или $\frac{w_i}{h}$ (плотность относительной частоты).

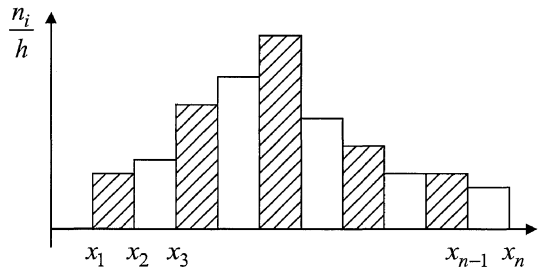


Рис. 23

Гистограмма является первым приближением плотности распределения $f(x)$ количественного признака X генеральной совокупности.

Примечание. Дискретный вариационный ряд получается из интервального, если за новые значения признака X взять середины интервалов.

Статистические оценки параметров распределения

Основной задачей выборочного метода является оценка параметров генеральной совокупности по данным выборки.

Теоретическую основу применимости выборочного метода составляет *закон больших чисел*, согласно которому при неограниченном увеличении объема выборки практически достоверно, что случайные выборочные характеристики как угодно близко приближаются (сходятся по вероятности) к определенным параметрам генеральной совокупности.

Если из теоретических соображений удалось установить, какое именно распределение имеет признак, то возникает задача оценки параметров, которыми определяется это распределение.

Например, если известно, что изучаемый признак генеральной совокупности распределен нормально, то необходимо оценить математическое ожидание a и среднее квадратичное отклонение σ.

Пусть требуется оценить неизвестный параметр Θ (читается «тэтта»). Исследовать все элементы генеральной совокупности не представляется возможным. Обычно в распоряжении исследователя имеются лишь данные выборки, состоящей из значений $x_1, x_2, ..., x_n$.

Эти значения можно рассматривать как частные значения случайных величин $X_1, X_2, ..., X_n$.

Оценкой Θ_n параметра Θ называется функция результатов наблюдений над случайной величиной X, с помощью которой судят о значении параметра Θ:

$$\Theta_n = \Theta_n\left(X_1, X_2, ..., X_n\right).$$

Оценка является также случайной величиной, зависящей от закона распределения случайной величины X и числа n.

Существует множество функций $\Theta_n = \Theta_n\left(X_1, X_2, ..., X_n\right)$, которые можно предложить в качестве оценки параметра Θ. О качестве (точности) оценки можно судить не по отдельным значениям оценки параметра, а лишь по распределению ее значений в большом количестве испытаний.

Основными свойствами оценок являются несмещенность, состоятельность и эффективность.

Несмещенность оценки означает отсутствие систематических погрешностей в наблюдаемых данных, для этого ее математическое ожидание должно быть равно оцениваемому параметру.

Состоятельность оценки заключается в том, что с ростом числа наблюдений дисперсия стремится к нулю.

Для исследуемого параметра оценка *эффективна*, если имеет минимальную дисперсию среди всех возможных оценок, построенных по данной выборке.

Выборочное среднее $\overline{x} = \frac{1}{n}\sum_{i=1}^{m} n_i x_i$ является несмещенной и состоятельной оценкой математического ожидания генеральной совокупности.

Выборочная дисперсия $D = \frac{1}{n}\sum_{i=1}^{m} n_i \left(x_i - \overline{x}\right)^2$ – это смещенная оценка генеральной дисперсии; поэтому в качестве оценки генеральной дисперсии берут *исправленную выборочную дисперсию* $s^2 = \frac{n}{n-1} D$.

Для вычисления дисперсии удобнее применять формулу

$$D = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^2 - \overline{x}^2 .$$

Для оценки среднего квадратического отклонения генеральной совокупности применяют *выборочное среднее квадратическое отклонение* $\sigma = \sqrt{D}$ или *исправленное выборочное среднее квадратическое отклонение* $s = \sqrt{\frac{n}{n-1} \cdot D}$.

Оценка параметров бывает *точечной* и *интервальной*:

– *точечной*, если определяется одним числом;

– *интервальной*, если по данным выборки строится числовой интервал, который с заданной вероятностью накрывает неизвестное значение оцениваемого параметра.

Границы этого интервала являются случайными величинами.

Доверительным интервалом называется интервал $\left(\tilde{\Theta}_n^{(1)}; \tilde{\Theta}_n^{(2)}\right)$, который с заданной надежностью (доверительной вероятностью) γ покрывает неизвестный параметр Θ.

Величина доверительного интервала уменьшается с ростом объема выборки и увеличивается с приближением надежности оценки γ к единице.

Часто доверительный интервал выбирается симметричным относительно точечной оценки параметра $\left(\tilde{\Theta}_n - \delta; \tilde{\Theta}_n + \delta\right)$, где $\tilde{\Theta}_n$ – точечная оценка параметра, δ – *точность оценки* или *предельная ошибка выборки*.

Предельная ошибка выборки возникает вследствие того, что исследуется не вся совокупность, а лишь ее часть (выборка), отобранная случайно.

Доверительные интервалы для генеральной средней и генеральной доли

Рассмотрим случай больших выборок (порядка сотен наблюдений).

Для генеральной средней:

$$\bar{x} - \delta \le \bar{x}_0 \le \bar{x} + \delta,$$

где $\delta = t\sigma_{\bar{x}}$, $\Phi(t) = \frac{\gamma}{2}$, $\sigma_{\bar{x}}$ – средняя квадратическая ошибка выборки.

Для генеральной доли:

$$w - \Delta \le p \le w + \Delta,$$

где $\Delta = t\sigma_w$, $\Phi(t) = \frac{\gamma}{2}$, σ_w – средняя квадратическая ошибка выборки.

Формулы приведены в сводной таблице (табл. 59).

Таблица 59

Оцениваемый параметр	Формулы для нахождения средних квадратических ошибок выборки	
	Повторная выборка	Бесповторная выборка
Средняя	$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} \approx \sqrt{\frac{D}{n}}$	$\sigma_{\bar{x}} \approx \sqrt{\frac{D}{n}\left(1 - \frac{n}{N}\right)}$
Доля	$\sigma_w = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{w(1-w)}{n}}$	$\sigma_w \approx \sqrt{\frac{w(1-w)}{n}\left(1 - \frac{n}{N}\right)}$

Т.к. генеральные доля p и дисперсия σ^2 неизвестны, то в формулах в таблице они заменены оценками по выборке w и D, соответственно. При достаточно большом объеме выборки практически достоверно, что $w \approx p$ и $D \approx \sigma^2$.

Пример. При обследовании выработки 1000 рабочих цеха в отчетном году по сравнению с предыдущим по схеме собственно-случайной выборки было отобрано 100 рабочих. Получены следующие данные (табл. 60). Найти границы, в которых с вероятностью 0,9544 заключена средняя выработка рабочих цеха для случаев повторной и бесповторной выборок.

Таблица 60

Выработка в отчетном году в процентах к предыдущему	94–100	100–106	106–112	112–118	118–124	124–130	130–136	136–142
Количество рабочих	3	7	11	20	28	19	10	2

Решение. Имеем $N = 1000$, $n = 100$.

Приведем необходимые вычисления в таблице (табл. 61).

Таблица 61

Интервалы	Количество рабочих, n_i	Середина интервала, x_i	$x_i n_i$	$x_i^2 n_i$
94–100	3	97	291	28 227
100–106	7	103	721	74 263
106–112	11	109	1 199	130 691
112–118	20	115	2 300	264 500
118–124	28	121	3 388	409 948
124–130	19	127	2 413	306 451
130–136	10	133	1 330	176 890
136–142	2	139	278	38 642
Сумма	100		11 920	1 429 612

Тогда

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{m} n_i x_i = \frac{11920}{100} = 119{,}2,$$

$$D = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^2 - \overline{x}^2 = \frac{1429612}{100} - 119{,}2^2 = 87{,}48.$$

Средняя квадратическая ошибка выборки для средней:

– для повторной выборки $\sigma_{\overline{x}} \approx \sqrt{\frac{D}{n}} = \sqrt{\frac{87{,}48}{100}} \approx 0{,}935$;

– для бесповторной выборки

$$\sigma_{\overline{x}} \approx \sqrt{\frac{D}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{87{,}48}{100}\left(1 - \frac{100}{1000}\right)} \approx 0{,}887.$$

Доверительная вероятность $\gamma = 0{,}9544$, $\Phi(t) = 0{,}4772$, $t = 2$.

Предельные ошибки выборки:

– для повторной выборки $\delta = 2 \cdot 0{,}935 \approx 1{,}87$;

– для бесповторной выборки $\delta = 2 \cdot 0{,}887 \approx 1{,}77$.

Доверительные интервалы:

– для повторной выборки $(119{,}2 - 1{,}87; 119{,}2 + 1{,}87)$ или $(117{,}33; 121{,}07)$;

– для бесповторной выборки $(119{,}2 - 1{,}77; 119{,}2 + 1{,}77)$ или $(117{,}43; 120{,}97)$.

Доверительные интервалы для генеральной средней при неизвестном σ

Пусть признак X генеральной совокупности распределен нормально.

Тогда доверительный интервал для генеральной средней имеет вид:

$$\overline{x} - \frac{t_\gamma s}{\sqrt{n}} \le \overline{x}_0 \le \overline{x} + \frac{t_\gamma s}{\sqrt{n}},$$

где t_γ находится по таблице.

Пример. Признак X генеральной совокупности распределен нормально. По выборке объема $n = 16$ найдены выборочная средняя $\overline{x} = 20{,}2$ и исправленное среднее квадратическое отклонение $s = 0{,}8$. Оцените неизвестное математическое ожидание при помощи доверительного интервала с надежностью $\gamma = 0{,}95$.

Решение. Найдем $t_\gamma = 2{,}13$ из таблицы: $\frac{t_\gamma s}{\sqrt{n}} = \frac{2{,}13 \cdot 0{,}8}{4} = 0{,}426$. Доверительный интервал $19{,}774 \le \overline{x}_0 \le 20{,}626$.

Доверительный интервал для генеральной дисперсии

Пусть распределение признака X в генеральной совокупности является нормальным.

Тогда доверительный интервал имеет вид

$$s(1-q) \le \sigma \le s(1+q),$$

где значение q находится из таблицы.

Если $q > 1$, то неравенство принимает вид

$$0 \le \sigma \le s(1+q).$$

Эмпирическая функция распределения

Пусть известно статистическое распределение частот количественного признака X.

Введем обозначения:

n_x – число наблюдений, при которых наблюдалось значение признака, меньшее x;

n – общее число наблюдений (объем выборки).

Относительная частота события $(X < x)$ равна $\frac{n_x}{n}$.

Если x будет изменяться, то будет изменяться и относительная частота, то есть относительная частота $\frac{n_x}{n}$ есть функция от x.

Эмпирической функцией распределения называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $(X < x)$.

Итак, по определению $F^*(x) = \frac{n_x}{n}$.

Из определения функции $F^*(x)$ вытекают следующие ее свойства:

1) значения эмпирической функции принадлежат отрезку $[0;1]$ $\left(0 \le F^*(x) \le 1\right)$;

2) $F^*(x)$ – неубывающая функция;

3) если x_1 – наименьший элемент выборки, то $F^*(x) = 0$, при $x < x_1$;

если x_k – наибольший элемент выборки, то $F^*(x) = 1$ при $x > x_k$.

Эмпирическая функция распределения выборки служит для оценки теоретической функции распределения генеральной совокупности.

Объем выборки

Для проведения выборочного наблюдения важно правильно установить объем выборки n, т.к. он определяет необходимые для обработки выборки временные и стоимостные затраты.

Пусть задана надежность (доверительная вероятность) оценки γ и точность (предельная ошибка выборки) δ.

Формулы для объема выборки выводятся из формул для предельной ошибки выборки при оценке генеральной средней (табл. 62). Например, для повторной выборки $\delta = t\sqrt{\frac{\sigma^2}{n}}$. Значит, $n = \frac{t^2\sigma^2}{\delta^2}$, где $\Phi(t) = \frac{\gamma}{2}$.

Таблица 62

Оцениваемый параметр	Объем выборки	
	Повторная выборка	Бесповторная выборка
Генеральная средняя	$n = \frac{t^2\sigma^2}{\delta^2}$	$n' = \frac{Nt^2\sigma^2}{t^2\sigma^2 + N\delta^2}$
Генеральная доля	$n = \frac{t^2pq}{\Delta^2}$	$n' = \frac{Nt^2pq}{t^2pq + N\Delta^2}$

Т.к. точные значения σ^2 и p неизвестны, в качестве этих характеристик используются выборочные данные, т.е. полагают $\sigma^2 \approx D$ и $p \approx w$.

Если никаких сведений о значениях σ^2 и p нет, то организуют специальную пробную выборку небольшого объема, находят требуемые оценки и определяют объем выборки.

При оценке генеральной доли вместо проведения исследования пробной выборки можно взять pq равным его наибольшему значению 0,25. Но при этом найденное значение объема выборки будет больше минимально необходимого для заданных точности и надежности оценок.

Пример. По данным предыдущего примера определить объем выборки, при котором с вероятностью 0,9973 отклонение средней выработки рабочих в выборке от средней выработки всех рабочих цеха не превзойдет по абсолютной величине 1%.

Решение. В качестве неизвестного значения σ^2 для определения объема выборки берем его оценку $D = 87{,}48$, найденную ранее.

Учитывая, что $\Phi(t) = \dfrac{0{,}9973}{2} = 0{,}49865$, находим $t = 3$.

Объем повторной выборки $n = \dfrac{t^2\sigma^2}{\delta^2} = \dfrac{3^2 \cdot 87{,}48}{1} = 787$, бесповторной

$$n' = \frac{1000 \cdot 3^2 \cdot 87{,}48}{3^2 \cdot 87{,}48 + 1000 \cdot 1^2} = 440{,}5 \approx 441.$$

Статистическая проверка гипотезы о законе распределения

Выдвижение и проверка гипотезы о виде закона распределения генеральной совокупности обычно является завершающим этапом обработки выборки.

Статистической называют гипотезу о виде неизвестного распределения или о параметрах известных распределений.

Нулевой (основной) называют выдвинутую гипотезу H_0.

Выдвинутая гипотеза может быть правильной или неправильной, поэтому и возникает необходимость ее проверки. *Критерием согласия* называют критерий проверки гипотезы о предполагаемом законе распределения случайной величины.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Основной принцип проверки статистических гипотез можно сформулировать так: если наблюдаемое значение критерия принадлежит критической области – гипотезу отвергают, если наблюдаемое значение принадлежит области принятия гипотезы – гипотезу принимают.

Критическими точками называют точки, отделяющие критическую область от области принятия гипотезы.

В теории статистических гипотез различают два вида ошибок:

ошибка 1 рода – это отклонение правильной гипотезы,

ошибка 2 рода – принятие неверной гипотезы.

Проверка статистической гипотезы осуществляется с помощью критерия (случайной величины, закон распределения которой установлен для данной задачи). При этом задается вероятность α совершить ошибку 1 рода, называемая *уровнем значимости*.

По уровню значимости формируется критическая область, если в нее попадает наблюдаемое значение критерия, гипотеза отвергается.

Иначе, гипотеза принимается, т.е. отклонения выборочных данных считаются случайными.

Для проверки гипотезы о нормальном распределении генеральной совокупности применяется *критерий Пирсона* (или χ^2). Он основывается на сравнении эмпирических (полученных опытным путем) и теоретических (вычисленных в предположении нормального распределения) частот.

В качестве проверки нулевой гипотезы принимают случайную величину $\chi^2_{набл.} = \sum_{i=1}^{k} \frac{\left(n_i - n_i^0\right)^2}{n_i^0}$, где n_i – эмпирические частоты; n_i^0 – теоретические частоты.

Алгоритм применения критерия Пирсона

1. Выдвигают нулевую гипотезу о нормальном законе распределения случайной величины X и находят его параметры $\bar{x}, s$.

2. Определяют теоретические частоты n_i^0, соответствующие эмпирическим частотам. Если среди эмпирических частот имеются малочисленные, то их необходимо объединить с соседними. Число интервалов после объединения должно быть не менее 4.

Если случайная величина X непрерывна, то $n_i^0 = n_i p_i = n\left(\Phi\left(\frac{x_i - \bar{x}}{s}\right) - \Phi\left(\frac{x_{i-1} - \bar{x}}{s}\right)\right)$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа, значения которой находятся по таблице.

3. Вычисляют величину $\chi^2_{набл.}$.

4. Определяют число степеней свободы $k = l - r - 1$, где l – число интервалов после объединения, r – число параметров в предполагаемом законе распределения. В нормальном законе распределения $r = 2$, следовательно, $k = l - 3$.

5. Находят уровень значимости $\alpha = 1 - \gamma$, где γ – доверительная вероятность. При $\gamma = 0{,}95$ $\alpha = 1 - 0{,}95 = 0{,}05$.

6. По таблице находят значение $\chi^2_{теор.}$ при заданных α и k, которое является критической точкой $\chi^2_{кр}(\alpha, k)$.

7. Если $\chi^2_{набл.} \le \chi^2_{кр}$ – нет оснований отвергнуть нулевую гипотезу, она не противоречит опытным данным. Если $\chi^2_{набл.} > \chi^2_{кр}$ – нулевую гипотезу отвергают.

Контрольные вопросы

1. В чем заключается сущность выборочного метода, и какие у него преимущества и недостатки по сравнению со сплошным наблюдением?
2. Что такое генеральная и выборочная совокупности?
3. Как построить вариационный ряд и интервальный вариационный ряд по данной выборке?
4. Какие бывают оценки параметров распределения генеральной совокупности?
5. Приведите формулы для точечных оценок генеральной средней и генеральной дисперсии.
6. Приведите формулы для интервальных оценок генеральной средней и генеральной дисперсии.
7. Что такое статистическая гипотеза?
8. Приведите алгоритм критерия Пирсона.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гмурман, В.Е. Теория вероятностей и математическая статистика: учебник для вузов / В.Е. Гмурман. – М.: Юрайт, 2016. – 479 с.

2. Кремер, Н.Ш. Теория вероятностей и математическая статистика: учебник и практикум для вузов по экон. направлениям и специальностям. – М.: Юрайт, 2015. – 514 с.

3. Лютикас, В.С. Школьнику о теории вероятностей: учеб. пособие по факультативному курсу для учащихся 8–10 классов / С.В. Лютикас. – М.: Просвещение, 1983. – 127 с.

ОГЛАВЛЕНИЕ

Учебное издание

Корытова Марина Александровна,
Шунайлова Светлана Александровна

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Техн. редактор *А.В. Миних*

Издательский центр Южно-Уральского государственного университета

Подписано в печать 19.06.2019.Формат 60×84 1/16. Печать цифровая.
Усл. печ. л. 6,28. Тираж 30 экз. Заказ 207/524.

Отпечатано в типографии Издательского центра ЮУрГУ.
454080, г. Челябинск, проспект Ленина, 76.