

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное
образовательное учреждение высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

*Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета*

Составитель Г.Е. Шевелёв

Издательство
Томского политехнического университета
2019

УДК 519.2
ББК 22.17я73
Т24

Т24

Теория вероятностей и математическая статистика:
учебное пособие / сост. Г.Е. Шевелев; Томский
политехнический университет. – Томск: Изд-во
Томского политехнического университета, 2019. – 114 с.
В авторской редакции

Учебное пособие состоит из семи глав, написано в соответствии с программой курса «Теория вероятностей и математическая статистика» для вузов. В нем рассмотрены основные теоретические понятия и определения теории вероятностей и математической статистики; приведены примеры решения задач.

Пособие подготовлено в отделении информационных технологий ИШИТР ТПУ и предназначено для студентов ИШИТР, обучающихся по направлениям 09.03.01, 09.03.02 и 09.03.04

УДК 519.2
ББК 22.17я73

Рецензенты

Доктор физико-математических наук
Профессор ТГАСУ
Б.М. Шумилов

Кандидат технических наук
доцент ТУСУР
А.А. Шелестов

© Составление. ФГАОУВО НИ ТПУ, 2019
© Шевелев Г.Е., составление, 2019
© Обложка. Издательство Томского
политехнического университета, 2019

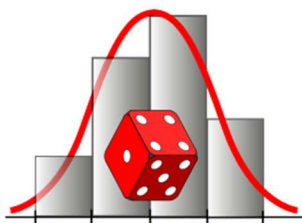
Оглавление

I. ТЕОРИЯ ВЕРОЯТНОСТЕЙ	5
ВВЕДЕНИЕ.....	5
1. Случайные события	6
Предмет теории вероятностей.....	6
Случайные события, их классификация.....	8
Действия над событиями	9
Алгебра событий. Теоретико-множественная трактовка	11
Свойство статистической устойчивости относительной частоты события.....	12
Статистическое определение вероятности.....	13
Классическое определение вероятности	14
Геометрическое определение вероятности	15
Элементы комбинаторики	17
Аксиоматическое определение вероятности	21
Свойства вероятностей	22
Условные вероятности	23
Вероятность произведения событий. Независимость событий	24
Вероятность суммы событий.....	25
Формула полной вероятности	26
Формула Байеса	27
Независимые испытания. Схема Бернулли. Формула Бернулли	29
Предельные теоремы в схеме Бернулли.....	30
2. Случайные величины	33
Понятие случайной величины. Закон распределения случайной величины.....	33
Закон распределения дискретной случайной величины. Многоугольник распределения...34	
Функция распределения и ее свойства. Функция распределения дискретной случайной величины.....	35
Плотность распределения и ее свойства	37
Числовые характеристики случайных величин.....	39
Основные законы распределения случайных величин	44
3. Системы случайных величин.....	55
Понятие о системе случайных величин и законе ее распределения	55
Функция распределения двумерной случайной величины и ее свойства.....	55
Плотность распределения вероятностей двумерной случайной величины и ее свойства ...58	
Условные законы распределения.....	59
Числовые характеристики двумерной случайной величины	61
Двумерное нормальное распределение	64
Общий случай n -мерного нормального распределения.....	66
Распределение функций нормальных случайных величин	67
4. Предельные теоремы теории вероятностей	73
Неравенство Чебышева	73
Закон больших чисел в форме Чебышева	74
Теорема Бернулли.....	76
Центральная предельная теорема	77

Доска Гальтона	79
II. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА	81
ВВЕДЕНИЕ.....	81
5. <i>Выборки и их характеристики</i>	82
Предмет математической статистики.....	82
Генеральная и выборочная совокупности.....	83
Статистическое распределение выборки. Эмпирическая функция распределения.....	84
Числовые характеристики статистического распределения	89
6. <i>Элементы теории оценок</i>	90
Оценка неизвестных параметров	90
Методы нахождения точечных оценок	92
Интервальное оценивание параметров.....	96
Доверительные интервалы для параметров нормального распределения.....	96
7. <i>Проверка статистических гипотез</i>.....	101
Статистическая гипотеза. Статистический критерий	101
Проверка гипотезы о равенстве центров распределения нормальных генеральных совокупностей при неизвестном σ	105
Проверка гипотез о законе распределения.....	106
ПРИЛОЖЕНИЯ	111

I. ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Введение



Теория вероятностей возникла в середине XVII века. Первые работы по теории вероятностей, принадлежащие французским учёным Б. Паскалю и П. Ферма и голландский учёному Х. Гюйгенсу, появились в связи с подсчётом различных вероятностей в азартных играх. Крупный успех теории вероятностей связан с именем швейцарского математика Я. Бернулли, установившего закон больших чисел для схемы независимых испытаний с двумя исходами.

Второй период истории теории вероятностей (XVIII в. и начало XIX в.) связан с именами А. Муавра (Англия), П. Лапласа (Франция), К. Гаусса (Германия) и С. Пуассона (Франция). Это – период, когда теория вероятностей уже находит ряд весьма актуальных применений в теории ошибок наблюдений, развившейся в связи с потребностями геодезии и астрономии, в теории стрельбы.

Третий период истории теории вероятностей, (вторая половина XIX в.) связан в основном с именами русских математиков П. Л. Чебышева, А. М. Ляпунова и А. А. Маркова (старшего). Теория вероятностей развивалась в России и раньше (в XVIII в. Ряд трудов по теории вероятности был написан работавший в России Л. Эйлером, Н. Бернулли и Д. Бернулли; во второй период развития теории вероятностей следует отметить работы М. В. Остроградского по вопросам теории вероятностей, связанным с математической статистикой, и В. Я. Буняковского по применению теории вероятностей к страховому делу, статистике и демографии).

Теория вероятностей есть математическая наука, выясняющая закономерности, которые возникают при взаимодействии большого числа случайных факторов. Наиболее распространённая в настоящее время логическая схема построения основ теории вероятностей разработана в 1933 советским математиком А. Н. Колмогоровым.

В Западной Европе во 2-й половине XIX в. получили большое развитие работы по математической статистике (в Бельгии – А. Кетле, в Англии – Ф. Гальтон) и статистической физике (в Австрии – Л. Больцман), которые наряду с основными теоретическими работами Чебышева, Ляпунова и Маркова создали основу для существенного расширения проблематики теории вероятностей в четвёртого (современном) периоде её развития. В этот период при очень большом усилении работы по теории вероятностей за рубежом современная наука продолжает занимать значительное, а в ряде направлений и ведущее положение.

1. Случайные события

Предмет теории вероятностей

При исследовании различных физических и технических задач часто приходится встречаться с особыми явлениями, которые принято называть случайными. *Случайное явление* – это такое явление, которое при неоднократном воспроизведении одного и того же опыта протекает каждый раз несколько по-иному.

Примеры случайных явлений.

1. Производится стрельба из орудия, установленного под заданным углом к горизонту. Фактическая траектория каждого отдельного снаряда неизбежно несколько отклоняется от теоретической за счет совокупного влияния многих факторов. Среди этих факторов можно, например, назвать: ошибки изготовления снаряда, отклонение веса заряда от номинала, неоднородность структуры заряда, метеорологические условия и т.д.

2. Одно и то же тело несколько раз взвешивается на аналитических весах; результаты повторных взвешиваний несколько отличаются друг от друга. Эти различия обусловлены влиянием многих второстепенных факторов, сопровождающих операцию взвешивания, таких, как положение тела на чашке весов, случайные вибрации аппаратуры, ошибки отсчета показаний приборов и т.д.

3. Самолет совершает полет на заданной высоте; теоретически он летит горизонтально, равномерно и прямолинейно. Фактически полет сопровождается отклонениями центра массы самолета от теоретической траектории и колебаниями самолета около центра массы. Эти отклонения и колебания являются случайными и связаны с турбулентностью атмосферы; от раза к разу они не повторяются.

Во всех приведенных примерах основные условия опыта, определяющие в общих и грубых чертах его протекание, сохраняются неизменными; второстепенные – меняются от опыта к опыту и вносят случайные различия в их результаты.

Случайные отклонения неизбежно сопутствуют любому закономерному явлению. Тем не менее, в ряде практических задач этими случайными элементами можно пренебречь, рассматривая вместо реального явления его упрощенную схему, «модель», и предполагая, что в данных условиях опыта явление протекает вполне определенным образом. При этом из бесчисленного множества факторов, влияющих на данное явление, выделяются самые главные, основные, решающие; влиянием остальных, второстепенных факторов просто пренебрегают. При пользовании этой схемой для решения любой задачи прежде всего выделяется основной круг учитываемых условий и выясняется, на какие параметры задачи они влияют; затем применяется тот или иной математический аппарат. По мере развития науки число учитываемых факторов становится все больше; явление исследуется подробнее; научный прогноз становится точнее.

Однако для решения ряда вопросов описанная схема – классическая схема так называемых «точных наук» – оказывается плохо приспособленной. Существуют

такие задачи, где интересующий нас исход опыта зависит от столь большого числа факторов, что практически невозможно зарегистрировать и учесть все эти факторы.

Очевидно, должна существовать принципиальная разница в методах учета основных факторов, определяющих в главных чертах течение явления, и второстепенных факторов, влияющих на течение явления в качестве «погрешностей» или «возмущений». Элемент неопределенности, сложности, присущий случайным явлениям, требует создания специальных методов для изучения этих явлений.

Такие методы и разрабатываются в теории вероятностей. Её предметом являются специфические закономерности, наблюдаемые в случайных явлениях. Практика показывает, что, наблюдая в совокупности массы однородных случайных явлений, мы обычно обнаруживаем в них вполне определенные закономерности, своего рода устойчивости, свойственные именно массовым случайным явлениям.

Например, если много раз подряд бросить монету, частота появления герба (отношение числа появившихся гербов к общему числу бросаний) постепенно стабилизируется, приближаясь к вполне определенному числу, именно, к $\frac{1}{2}$. Такое же свойство «устойчивости частот» обнаруживается и при многократном повторении любого другого опыта, исход которого представляется заранее неопределенным, случайным.

Подобные специфические, так называемые «статистические», закономерности наблюдаются всегда, когда мы имеем дело с массой однородных случайных явлений. Эти отдельные особенности в массе как бы взаимно погашаются, нивелируются, и средний результат массы случайных явлений оказывается практически уже не случайным. Именно эта многократно подтвержденная опытом устойчивость массовых случайных явлений и служит базой для применения вероятностных методов исследования.

Методы теории вероятностей приспособлены только для исследования *массовых случайных явлений*; они не дают возможности предсказать исход отдельного случайного явления, но дают возможность предсказать средний суммарный результат массы однородных случайных явлений. Вероятностный, или статистический, метод в науке не противопоставляет себя классическому, обычному методу точных наук, а является его дополнением, позволяющим глубже анализировать явление с учетом присущих ему элементов случайности.

В настоящее время нет почти ни одной естественной науки, в которой, так или иначе, не применялись бы вероятностные методы. Целые разделы современной физики (в частности, ядерная физика) базируются на методах теории вероятностей. Все шире применяются вероятностные методы в современной электротехнике, радиотехнике, метеорологии и астрономии, теории автоматического регулирования. Обширно поле применения находит теория вероятностей в разнообразных областях военной техники.

Теория вероятностей – математическая наука, изучающая закономерности, присущие *массовым* случайным явлениям (их можно наблюдать практически неограниченное число раз в одинаковых условиях).

Предметом теории вероятностей являются математические модели случайных явлений. При этом под *случайным явлением* понимают явление, предсказать исход, которого невозможно.

Цель теории вероятностей – осуществление прогноза в области случайных явлений, влияние на ход этих явлений, контроль их, ограничение сферы действия случайности.

Случайные события, их классификация

Случайное событие – это любой факт, который в результате испытания может произойти или не произойти. Испытание – это эксперимент, выполнение определенного комплекса условий, в которых наблюдается то или иное явление, фиксируется тот или иной результат. При этом рассматриваются только такие эксперименты, которые можно повторять, хотя бы теоретически, при неизменном комплексе условий произвольное число раз.

Случайным событием (или просто *событием*) называется любой исход опыта, который может произойти или не произойти.

События обозначаются, как правило, заглавными буквами латинского алфавита: A, B, C, \dots

Пример 1.1. Опыт: бросание игральной кости; событие A – выпадение 5 очков, событие B – выпадение четного числа очков, событие C – выпадение 7 очков, событие D – выпадение целого числа очков, событие E – выпадение не менее 3-х очков,

Непосредственные исходы опыта называются *элементарными событиями* и обозначаются через w . Элементарные события (их называют также «элементами», «точками», «случаями») рассматриваются как неразложимые и взаимоисключающие исходы $w_1, w_2, w_3 \dots$ этого опыта.

Множество всех элементарных событий называется *пространством элементарных событий (ПЭС)* или *пространством исходов*, обозначается через Ω .

Рассмотрим **пример 1.1**. Здесь 6 элементарных событий $w_1, w_2, w_3, w_4, w_5, w_6$. Событие w_i означает, что в результате бросания кости выпало i очков, $i = 1, 2, 3, 4, 5, 6$. Пространство элементарных событий таково: $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6\}$ или $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Событие называется *достоверным*, если оно обязательно наступит в результате данного опыта, обозначается через Ω .

Событие называется *невозможным*, если оно заведомо не произойдет в

результате проведения опыта, обозначается через \emptyset .

В примере 1.1 события A и B – случайные, событие C – невозможное, событие D – достоверное.

Два события называются *несовместными*, если появление одного из них исключает появление другого события в одном и том же опыте, т. е. не смогут произойти вместе в одном опыте. В противном случае события называются *совместными*.

Так, в примере 1.1 события A и B – несовместные, A и E – совместные. События A_1, A_2, \dots, A_n называются *попарно-несовместными*, если любые два из них несовместны.

Несколько событий образуют *полную группу*, если они попарно несовместны и в результате каждого опыта происходит одно и только одно из них.

В примере 1.1 события w_1-w_6 образуют полную группу, w_1-w_5 – нет.

Несколько событий в данном опыте называются *равновозможными*, если ни одно из них не является объективно более возможным, чем другие, т. е. все события имеют равные «шансы».

В примере 1.1 элементарные события $w_1, w_2, w_3, w_4, w_5, w_6$ равновозможны. Выпадение герба (A) или решки (B) при бросании монеты – равновозможные события, если, конечно, монета имеет симметричную форму, не погнута,

Действия над событиями

Введем основные операции над событиями; они полностью соответствуют основным операциям над множествами.

Суммой событий A и B называется событие $C = A + B$, состоящее в наступлении хотя бы одного из них (т. е. или A , или B , или A и B вместе).

Произведением событий A и B называется событие $C = A \cdot B$, состоящее в совместном наступлении этих событий (т. е. и A и B одновременно).

Разностью событий A и B называется событие $C = A - B$, происходящее тогда и только тогда, когда происходит событие A , но не происходит событие B .

Противоположным событию A называется событие \bar{A} , которое происходит тогда и только тогда, когда не происходит событие A (\bar{A} означает, что событие A не наступило).

Событие A влечет событие B (или A является частным случаем B), если из того, что происходит событие A , следует, что происходит событие B записывают $A \subseteq B$.

Если $A \subseteq B$ и $B \subseteq A$, то события A и B называются *равными*; записывают $A = B$.

Так, в примере 1.1 (п. 1.2) $B = \{2,4,6\}$, $E = \{3,4,5,6\}$, $A = \{5\}$, $D = \{1,2,3,4,5,6\}$. Тогда: $B + E = \{2,3,4,5,6\}$, $B \cdot E = \{4,6\}$, $B - E = \{2\}$, $\bar{A} = \{1,2,3,4,6\}$, $B \subseteq D$, $D = \Omega = \{1, 2,3,4,5, 6\}$.

События и действия над ними можно наглядно иллюстрировать с помощью *диаграмм Эйлера-Венна*: достоверное событие Ω изображается прямоугольником; элементарные случайные события – точками прямоугольника; случайное событие – областью внутри него.

Действия над событиями можно изобразить так, как показано на рис. 1-5.

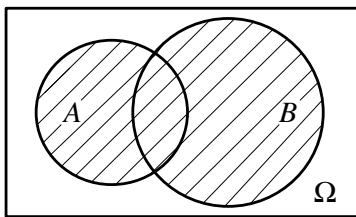


Рис. 1. $A+B$

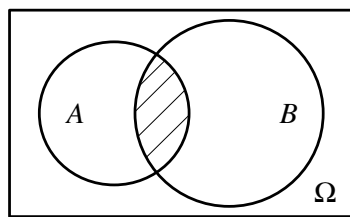


Рис. 2. $A \cdot B$

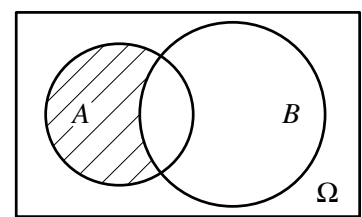


Рис. 3. $A - B$

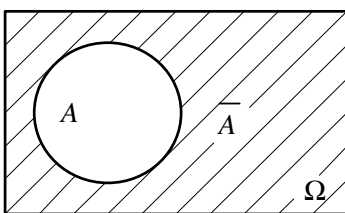


Рис. 4. \bar{A}

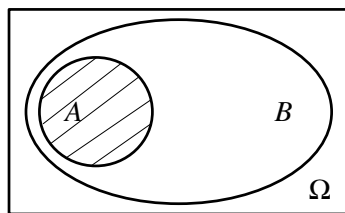


Рис. 5. $A \subseteq B$

Операции над событиями обладают следующими свойствами:

- $A + B = B + A$, $A \cdot B = B \cdot A$ (переместительное);
- $(A+B) \cdot C = A \cdot C + B \cdot C$, $A \cdot B + C = (A+C) \cdot (B+C)$ (распределительное);
- $(A+B)+C = A+(B+C)$, $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ (сочетательное);
- $A + A = A$, $A \cdot A = A$;

- $A + \Omega = \Omega$, $A \cdot \Omega = A$;
- $A + \bar{A} = \Omega$, $A \cdot \bar{A} = \emptyset$;
- $\bar{\emptyset} = \Omega$, $\bar{\Omega} = \emptyset$, $\bar{\bar{A}} = A$;
- $A - B = A \cdot \bar{B}$;
- $\overline{A+B} = \bar{A} \cdot \bar{B}$; и $\overline{A \cdot B} = \bar{A} + \bar{B}$ – законы де Моргана.

В их справедливости можно убедиться с помощью диаграмм Эйлера-Венна.

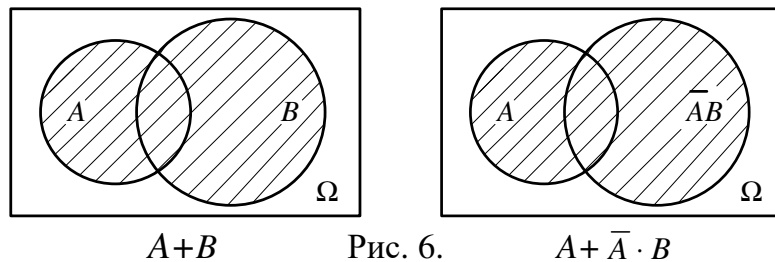
Пример 1.2. Доказать формулу: $A + B = A + \bar{A} \cdot B$

Доказательство. Используя некоторые из вышеприведенных правил, получим:

$$A + B = (A + B) \cdot \Omega = A \cdot \Omega + B \cdot \Omega = A \cdot \Omega + B \cdot (A + \bar{A}) = A \cdot \Omega + (A + \bar{A}) \cdot B = A \cdot \Omega + A \cdot B + \bar{A} \cdot B = (\Omega + B) \cdot A + \bar{A} \cdot B = A + \bar{A} \cdot B$$

Таким образом, сумму любых двух событий можно представить в виде суммы двух несовместных событий.

Геометрическое доказательство представлено на рис. 6.



$A+B$

Рис. 6.

$A + \bar{A} \cdot B$

Алгебра событий. Теоретико-множественная трактовка



А.Н. Колмогоров (1903-1987)

В 1900 г. на II Международном конгрессе математиков в Париже Д. Гильберт представил список 23 кардинальных проблем математики еще не решенных к тому времени (сейчас решены 16 этих проблем). Проблема аксиоматизации вероятностей включена Д Гильбертом в формулировку 6-ой проблемы «Математическое изложение основ физики». А.Н. Колмогоров под влиянием идей нескольких теорий (множеств, меры, интегрирования, функций) сформулировал простую систему аксиом, которая позволила описать уже существующие к тому времени классические разделы теории вероятностей, дать толчок развитию ее новых разделов и стала общепринятой в современной теории вероятностей.

Определим теперь основные понятия теории вероятностей, следуя теоретико-множественному подходу, разработанному академиком Колмогоровым в 1933 г.

Пусть производится некоторый опыт со случайным исходом.

Множество $\Omega = \{w\}$ всех возможных взаимоисключающих исходов данного опыта называется *пространством элементарных событий*, а сами исходы w – *элементарными событиями* (или «элементами», «точками»).

Случайным событием A (или просто *событием* A) называется любое

подмножество множества Ω , если Ω конечно или счетно (т. е. элементы этого множества можно пронумеровать с помощью множества натуральных чисел):
 $A \subseteq \Omega$.

Элементарные события, входящие в подмножество A пространства Ω , называются благоприятствующими событию A .

Множество Ω называется достоверным событием. Ему благоприятствует любое элементарное событие; в результате опыта оно обязательно произойдет.

Пустое множество \emptyset называется невозможным событием; в результате опыта оно произойти не может.

Над событиями можно проводить все операции, выполнимые для множеств.

Сумма (или объединение) двух событий $A \in \Omega$ и $B \in \Omega$ (обозначается $A+B$ или $A \cup B$) – это множество, которое содержит элементы, принадлежащие хотя бы одному из событий A и B .

Произведение двух событий $A \in \Omega$ и $B \in \Omega$ (обозначается AB или $A \cap B$) – это множество, которое содержит элементы общие для событий A и B .

Разность событий $A \in \Omega$ и $B \in \Omega$ (обозначается $A-B$ или $A \setminus B$) – это множество, которое содержит элементы события A , не принадлежащие событию B .

Противоположное событию $A \in \Omega$ событие $\bar{A} = \Omega \setminus A$ называется *дополнением* множества A .

Событие A влечет событие B (обозначается $A \subseteq B$), если каждый элемент события A содержится в B .

По определению $\emptyset \subseteq A$ для любого A .

События A и B называются *несовместными*, если их произведение есть невозможное событие, т.е. $A \cdot B = \emptyset$.

Несколько событий A_1, A_2, \dots, A_n образуют полную группу несовместных событий, если их сумма представляет все пространство элементарных событий, а сами события несовместны, т.е. $\sum_{i=1}^n A_i = \Omega$ и $A_i \cdot A_j = \emptyset$ ($i \neq j$).

Полную группу, например, образуют события A и \bar{A} .

В случае несчетного пространства Ω в качестве событий рассматриваются не все подмножества Ω , а лишь некоторый класс этих подмножеств, называемый алгеброй множеств (событий) S . При расширении операций сложения и умножения на случай счетного множества алгебра множеств S называется σ -алгеброй.

Свойство статистической устойчивости относительной частоты события

Пусть в n повторяющихся опытах некоторое событие A наступило n_A раз. Число n_A называется *частотой* события A , а отношение

$$\frac{n_A}{n} = P^*(A) \quad (1.1)$$

называется *относительной частотой* (или *частотью*) события A в рассматриваемой серии опытов.

Относительная частота события обладает следующими свойствами:

1. Частость любого события заключена между нулем и единицей, т. е.

$$0 \leq P^*(A) \leq 1.$$

2. Частость невозможного события равна нулю, т. е.

$$P^*(\emptyset) = 0.$$

3. Частость достоверного события равна 1, т. е.

$$P^*(\Omega) = 1.$$

4. Частость суммы двух несовместных событий равна сумме частоты этих событий, т. е. если $A \cdot B = \emptyset$, то

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

Частость обладает еще одним фундаментальным свойством, называемым *свойством статистической устойчивости*: с увеличением числа опытов (т. е. n) она принимает значения, близкие к некоторому постоянному числу (говорят: частость стабилизируется, приближаясь к некоторому числу, частость колеблется около некоторого числа).

Бросание монеты: A – выпадение герба

Ж. Бюффон: $n=4040$ $n_A=2048$ $P^*(A)=0,5069$

К. Пирсон: $n=24000$ $n_A=12012$ $P^*(A)=0,5005$

Отметим, что теория вероятностей изучает только те массовые случайные явления с неопределенным исходом, для которых предполагается наличие устойчивости относительной частоты.

Статистическое определение вероятности

Для математического изучения случайного события необходимо ввести какую-либо количественную оценку события. Понятно, что одни события имеют больше шансов («более вероятны») наступить, чем другие. Такой оценкой является вероятность события, т. е. число, выражающее степень возможности его появления в рассматриваемом опыте. Математических определений вероятности существует несколько, все они дополняют и обобщают друг друга

Рассмотрим опыт, который можно повторять любое число раз (говорят: «проводятся повторные испытания»), в котором наблюдается некоторое событие A .

Статистической вероятностью события A называется число, около которого колеблется относительная частота события A при достаточно большом числе испытаний (опытов).

Вероятность события A обозначается символом $P(A)$. Согласно данному определению

$$P(A) \cong \frac{n_A}{n} = P^*(A) \quad (1.2)$$

Математическим обоснованием близости относительной частоты $P^*(A)$ и вероятности $P(A)$ некоторого события A служит теорема Я. Бернулли.

Вероятности $P(A)$ приписываются свойства 1-4 относительной частоты $P^*(A)$.

Недостатком статистического определения является неоднозначность статистической вероятности; так в примере с бросанием монеты в качестве вероятности можно принять не только число 0,5, но и 0,49 или 0,51 и т.д. Для надежного определения вероятности нужно проделать большое число испытаний (опытов), что не всегда просто (или дешево).

Классическое определение вероятности

Пусть проводится опыт с n исходами, которые можно представить в виде *полной группы несовместных равновозможных событий*. Такие исходы называются *случаями, шансами, элементарными событиями*, опыт – классическим. Про такой опыт говорят, что он сводится к *схеме случаев*.

Случай w , который приводит к наступлению события A , называется *благоприятным* (или – *благоприятствующим*) ему, т.е. случай w влечет событие A : $w \subseteq A$.

Вероятностью события A называется отношения числа m случаев, благоприятствующих этому событию, к общему числу n случаев, т.е.

$$p = P(A) = \frac{m}{n} \quad (1.3)$$

Из классического определения вероятности (1.3) вытекают следующие свойства:

1. Вероятность любого события заключена между нулем и единицей, т. е.
 $0 \leq P(A) \leq 1$.

2. Вероятность невозможного события равна нулю, т. е.
 $P(\emptyset) = 0$.

3. Вероятность достоверного события равна 1, т. е.
 $P(\Omega) = 1$.

4. Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий, т. е. если $A \cdot B = \emptyset$, то

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

Пример 1.3. В урне находятся 12 белых и 8 черных шаров. Какова вероятность того, что наудачу вынутый шар будет белым?

Решение. Пусть A – событие, состоящее в том, что вынут белый шар. Ясно, что $n = 12 + 8 = 20$ – число всех равновозможных случаев (исходов опыта). Число случаев, благоприятствующих событию A , равно 12, т. е. $m = 12$. Следовательно, по

формуле (1.3) имеем: $P(A)=12/20$, т.е. $P(A) = 0,6$.

Геометрическое определение вероятности

Недостатком классического определения вероятности является предположение, что число элементарных исходов конечно, хотя на практике часто встречаются испытания с бесконечным числом всевозможных исходов. Для преодоления этого недостатка вводятся геометрические вероятности – вероятности попадания точки в область (отрезок, часть плоскости, часть объема, часть отрезка времени и так далее).

Рассмотрим на плоскости некоторую область Ω , имеющую площадь S_Ω , и внутри области Ω область D с площадью S_D (см. рис. 7).

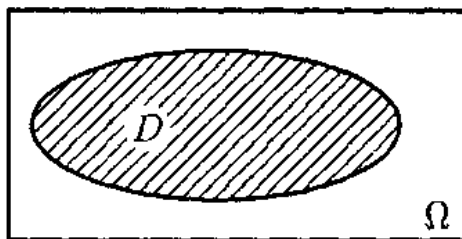


Рис. 7

В области Ω случайно выбирается точка X . Этот выбор можно интерпретировать как *бросание точки X в область Ω* . При этом попадание точки в область Ω – достоверное событие, в D – случайное. Предполагается, что все точки области Ω равноправны (все элементарные события равновозможны), т. е. что брошенная точка может попасть в любую точку области Ω и вероятность попасть в область D пропорциональна площади этой области и не зависит от ее расположения и формы. Пусть событие $A = \{X \in D\}$, т. е. брошенная точка попадет в область D .

Геометрической вероятностью события A называется отношение площади области D к площади области Ω , т. е.

$$P(A) = \frac{S_D}{S_\Omega} \quad (1.4)$$

Геометрическая вероятность обладает всеми свойствами, присущими классическому (и другим) определению.

Пример 1.4. На окружности взята точка A . На окружность «бросают» точку B . Какова вероятность того, что длина хорды AB будет меньше радиуса окружности?

Решение. Пусть C – длина окружности. Чтобы хорда AB была меньше радиуса окружности r , точка B должна попасть на дугу $B_1 A B_2$, длина которой равна

$$\frac{1}{3}C \quad (\text{см. рис. 8}). \text{ Тогда } p = \frac{\frac{1}{3}C}{C} = \frac{1}{3}.$$

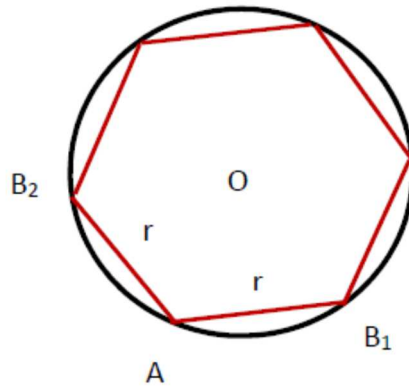


Рис. 8

Пример 1.5. (Задача о встрече.) Два человека договорились о встрече между 9 и 10 часами утра. Пришедший первым ждет второго в течение 15 мин, после чего уходит (если не встретились). Найти вероятность того, что встреча состоится, если каждый наудачу выбирает момент своего прихода.

Решение. Пусть x – время прихода первого, а y – второго. Возможные значения x и y : $0 \leq x \leq 60$, $0 \leq y \leq 60$ (в качестве единиц масштаба возьмем минуты), которые на плоскости Oxy определяют квадрат со стороной, равной 60. Точки этого квадрата изображают время встречающихся (см. рис. 9).

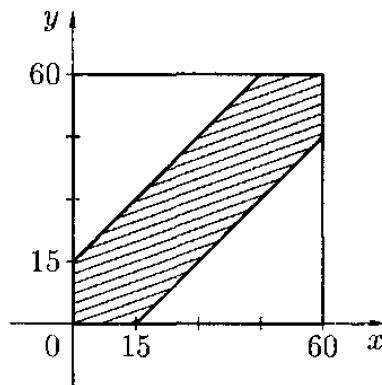


Рис. 9

Тогда $\Omega = \{(x, y) : 0 \leq x \leq 60, 0 \leq y \leq 60\}$; все исходы Ω равновозможны, т.к. лица приходят наудачу. Событие A – лица встретятся – произойдет, если разность между моментами их прихода будет не более 15 минут (по модулю), т.е. $A = \{(x, y) : |y - x| \leq 15\}$. Неравенство $|y - x| \leq 15$, т.е. $x - 15 \leq y \leq x + 15$ определяет область, заштрихованную на рис. 8, т.е. точки полосы есть исходы, благоприятствующие встрече. Искомая вероятность равна:

$$P(A) = \frac{60^2 - 2 \cdot \frac{1}{2} \cdot 45 \cdot 45}{60^2} = \frac{7}{16} \approx 0.44$$

Элементы комбинаторики

Согласно классическому определению подсчет вероятности события A сводится к подсчету числа благоприятствующих ему исходов. Делают это обычно комбинаторными методами.

Комбинаторика – раздел математики, в котором изучаются задачи выбора элементов из заданного множества и расположения их в группы по заданным правилам, в частности задачи о подсчете числа комбинаций (выборки), получаемых из элементов заданного конечного множества. В каждой из них требуется подсчитать число возможных вариантов осуществления некоторого действия, ответить на вопрос «сколькими способами?».

Многие комбинаторные задачи могут быть решены с помощью следующих двух важных правил, называемых соответственно правилами умножения и сложения.

Правило умножения

Если элемент x_1 строки (x_1, x_2, \dots, x_k) можно выбрать n_1 способами и после каждого такого выбора x_1 элемент x_2 можно выбрать n_2 – способами, и после выбора x_1 и x_2 элемент x_3 можно выбрать n_3 способами и т.д., наконец, x_k независимо от выбора всех предыдущих элементов можно выбрать n_k способами. Тогда количество возможностей (комбинаций) образования строки (x_1, x_2, \dots, x_k) равно:
$$N = n_1 \cdot n_2 \cdot n_3 \cdot \dots \cdot n_k .$$

Пример 1.6. Обед в университетской столовой состоит из трех блюд. Первое блюдо в меню может быть выбрано 5 способами, второе блюдо – 4, а третье блюдо – 3. Сколько дней студент может съесть новый обед, если любая комбинация блюд возможна, и один обед от другого должен отличаться хотя бы одним блюдом?

Решение. При решении данной задачи применим правило произведения (комбинаторика), и учтем, что строка состоит из трех элементов. Первое блюдо (первый элемент строки) можно выбрать пятью различными способами, второе – четырьмя различными способами независимо от выбора первого. Таким образом, первые два блюда можно выбрать 5·4 различными комбинациями. Учитывая выбор третьего блюда, окончательно получим:

$$N = 5 \cdot 4 \cdot 3 = 60 .$$

Правило суммы

Пусть множество E_1 содержит n_1 элемент, множество E_2 – n_2 элементов, ..., и множество E_k – n_k элементов. И если эти множества попарно не пересекаются (нет одинаковых элементов), то число элементов в их объединении равно сумме чисел элементов, содержащихся в каждом из этих множеств: $N = n_1 + n_2 + n_3 + \dots + n_k$.

Пример 1.7. В студенческой группе 14 девушек и 6 юношей. Сколькими способами можно выбрать, для выполнения различных заданий, двух студентов одного пола?

Решение. По правилу умножения двух девушек можно выбрать $14 \cdot 13 = 182$ способами, а двух юношей – $6 \cdot 5 = 30$ способами. Следует выбрать двух студентов одного пола: двух студенток или двух юношей. Согласно правилу сложения таких способов выбора, будет

$$N = 182 + 30 = 212.$$

Решение вероятностных (и не только их) задач часто облегчается, если использовать комбинаторные формулы. Каждая из них определяет число всевозможных исходов в некотором опыте (эксперименте), состоящем в выборе наудачу g элементов из n различных элементов рассматриваемого множества.

Существуют две схемы выбора g элементов ($0 < g \leq n$) из исходного множества: без возвращения (без повторений) и с возвращением (с повторением). В первом случае выбранные элементы не возвращаются обратно; можно отобрать сразу все g элементов или последовательно отбирать их по одному. Во второй схеме выбор осуществляется поэлементно с обязательным возвращением отобранного элемента на каждом шаге.

Схема выбора без возвращений

Правило перестановки

Пусть $E^{(n)} = \{x_1, x_2, \dots, x_n\}$ – произвольное (неупорядоченное) n -элементное множество. Рассмотрим различные комбинации его упорядочивания. Получаемые при этом упорядоченные множества отличаются друг от друга только порядком следования входящих в них элементов, и называются перестановками из n -элементов. Число всех таких перестановок обозначается символом P_n и находится по формуле

$$P_n = n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n. \quad (0! = 1! = 1).$$

Пример 1.8. Пятеро гостей случайным образом рассаживаются за столом. Сколькими способами можно их рассадить так, чтобы хотя бы 2 гостя поменялись местами (изменился порядок)?

Решение. При решении данной задачи, учитывая, что за столом всегда сидят все 5 гостей, применим правило перестановки:

$$N = P_5 = 5! = 120.$$

Правило размещения

Различные упорядоченные m – элементные подмножества данного неупорядоченного множества $E^{(n)}$ ($m < n$) называются размещениями из n элементов по m . Число таких размещений обозначается A_n^m и вычисляется по формуле

$$A_n^m = \frac{n!}{(n-m)!} = n(n-1)\dots(n-m+1).$$

Пример 1.9. Десять участников финала разыгрывают одну золотую, одну серебряную и одну бронзовую медали. Сколькими способами эти награды могут быть распределены между спортсменами?

Решение. Согласно условию данной задачи награды, получают только три финалиста из десяти, а ценность медалей различна, т.е. порядок призеров имеет значение. Тогда для определения числа комбинаций призеров применим правило размещений:

$$A_{10}^3 = \frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720.$$

Правило сочетания

Различные неупорядоченные m – элементные подмножества множества $E^{(n)}$ ($m < n$) называются сочетаниями из n элементов по m . Число всех таких сочетаний обозначается символом C_n^m и определяется по формуле

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Можно доказать справедливость следующих формул:

$$C_n^m = C_n^{n-m}, \quad (m \leq n);$$

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n;$$

$$C_n^m = C_{n-1}^m + C_{n-1}^{m-1}, \quad (1 < m < n).$$

Пример 1.10. В полуфинальном забеге участвуют десять спортсменов. Три спортсмена, показавшие лучший результат, попадают в финал. Сколько существует различных троек финалистов?

Решение. По условию задачи в финал войдут только три спортсмена из десяти, причем место в призовой тройке не имеет значения. Тогда для определения числа комбинаций призеров применим правило сочетаний:

$$C_{10}^3 = \frac{10!}{7!3!} = 120.$$

Примечания:

Размещения из n -элементов по m представляют собой такие m – элементные выборки из неупорядоченного множества $E^{(n)}$, которые отличаются друг от друга либо самими элементами (хотя бы одним), либо порядком их расположения.

Сочетания же из n -элементов по m представляет собой m – элементные выборки, отличающиеся только самими элементами.

Схема выбора с возвращением

Правило размещения

Любая строка длиной m , составленная из элементов множества $E^{(n)}$, причем элементы в строке могут повторяться, называется размещением с повторением из n -элементов по m . Число всех размещений с повторениями обозначается символом \bar{A}_n^m и вычисляется по формуле:

$$\bar{A}_n^m = n^m.$$

Пример 1.11. Для автомобильных номеров используются 10 цифр и 28 букв. Каждый номер состоит из 3 букв и 4 цифр. Какое максимальное число машин может получить номера при такой системе нумерации?

Решение. Сначала осуществим выбор 4 цифр. Каждый такой комплект цифр представляет собой четырехэлементную выборку из 10 – элементного массива цифр, т.е. является размещением с повторениями из 10 элементов по 4. Следовательно, общее число таких элементов равно 10^4 . Исключим из выборки номер 00-00, если он недопустим. Аналогично выбор трех букв из 28 осуществляется 28^3 числом способов. Так как номер каждой машины есть упорядоченная «пара», состоящая из комплекта цифр и комплекта букв, то по правилу произведения число всех номеров будет равно:

$$N = (10^4 - 1) \times 28^3 = 219\,498\,048.$$

Правило сочетания

Рассмотрим сочетания из n элементов по m и предположим, что в комбинации возможны повторения. В этом случае выбор элементов комбинации осуществляется не только по одному разу из n элементов, но и еще до $(m - 1)$ раза одного из этих элементов. В этом случае общее число элементов, из которых осуществляется комбинация, следует увеличить до $(n + m - 1)$ элементов. Следовательно, число сочетаний из n элементов по m с повторениями определяется по формуле

$$\bar{C}_n^m = C_{n+m-1}^m.$$

Пример 1.12. В цветочном киоске продается 10 наименований цветов. Покупатель желает приобрести букет из 5 цветов. Сколько существует комбинаций таких букетов?

Решение. Очевидно, что цветы одного наименования могут повторяться в букете, и так как порядок цветов в букете не имеет значения, то здесь применима формула числа сочетаний с повторениями:

$$\bar{C}_{10}^5 = C_{10+5-1}^5 = C_{14}^5 = \frac{14!}{5! \cdot 9!} = 2002.$$

Правило перестановки

Рассмотрим перестановки, содержащие одинаковые элементы. Например, в перестановках из n элементов имеются k различных элементов ($k < n$). При этом первый элемент встречается n_1 раз. Это означает, что общее число перестановок должно быть уменьшено в $n_1!$ раз, так как взаимные перестановки одного и того же элемента равнозначны. Аналогично происходит и с остальными элементами, которые могут встречаться n_2, n_3, \dots, n_k раз, причем $n_1 + n_2 + \dots + n_k = n$. Поэтому общее число перестановок с повторениями подсчитывается по формуле

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

Пример 1.13. Имеется шестизначная кодовая комбинация, состоящая из трех цифр 1, 3, 5, в которой цифра 1 встречается один раз, цифра 3 – два раза и цифра 5 – три раза. Сколько существует комбинаций таких наборов?

Решение. В данном случае имеют место перестановки с повторениями. Их число будет равно

$$P_6(1;2;3) = \frac{6!}{1! 2! 3!} = 60.$$

Аксиоматическое определение вероятности

До начала XX в. теория вероятностей представляла собой еще не сложившуюся науку, в которой основные понятия были недостаточно четко определены. Эта нечеткость приводила нередко к парадоксальным выводам (наиболее известные из них – парадоксы Бертрана).

Парадокс Бертрана заключается в следующем: рассмотрим равносторонний треугольник, вписанный в окружность. Наудачу выбирается хорда окружности. Какова вероятность того, что выбранная хорда длиннее стороны треугольника?

Классическое решение проблемы зависит от метода, которым случайно выбрана хорда. Тогда и только тогда, когда метод случайного выбора задан, проблема имеет четко определённое решение. Метод отбора не уникален, поэтому не может быть единственного решения.

Возникающие противоречия подрывали имевшиеся представления о вероятности. Было объявлено, что понятие о вероятности пригодно только для конечного числа элементарных исходов (то, что сейчас называется элементарной теорией вероятностей), а для бесконечного числа исходов оно не пригодно. Однако такой подход не мог устраивать практиков, ведь на практике чаще приходилось сталкиваться именно с бесконечным числом исходов. Развитие естествознания в конце XIX – начале XX вв. предъявило к вопросам обоснования наук повышенные требования. Однако первые работы по аксиоматическому обоснованию теории вероятностей не дали удовлетворительного решения задачи. Наиболее перспективным оказался путь, предложенный французским математиком Э. Борелем, который подметил глубокую аналогию между вероятностью и мерой, одним из наиболее важных понятий современной теории функций.

Мера множества – неотрицательная величина, интуитивно интерпретируемая как размер (объем) множества. Собственно, мера — это некоторая числовая функция, ставящая в соответствие каждому множеству (из некоторого семейства множеств) некоторое неотрицательное число.

Определение вероятности как меры, позволило связать теорию вероятностей с метрической теорией функций и теорией множеств. На этой базе А.Н. Колмогорову (1933 г.) удалось построить логически совершенное здание современной теории вероятностей.

Пусть Ω – множество всех возможных исходов некоторого опыта (эксперимента), S – алгебра событий. Совокупность S подмножеств множества Ω называется *алгеброй* (*σ -алгеброй*), если выполнены следующие условия:

1. S содержит невозможное и достоверное события.

2. Если события A_1, A_2, A_3, \dots (конечное или счетное множество) принадлежат S , то S принадлежит сумма, произведение и дополнение (т. е. противоположное для A_i) этих событий.

3. *Вероятностью* называется функция $P(A)$, определенная на алгебре событий S , принимающая действительные значения и удовлетворяющая следующим аксиомам:

A1. Аксиома неотрицательности: вероятность любого события $A \in S$ неотрицательна, т. е.

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

A2. Аксиома нормированности: вероятность достоверного события равна единице, т. е.

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

A3. Аксиома аддитивности: вероятность суммы несовместных событий равна сумме вероятностей этих событий, т. е. если $A_i \cdot A_j = \emptyset$ ($i \neq j$), то

$$P\left(\sum_k A_k\right) = \sum_k P(A_k).$$

Совокупность объектов (Ω, S, P) , где Ω – пространство элементарных событий, S — алгебра событий, P – числовая функция, удовлетворяющая аксиомам *A1-A3*, называется *вероятностным пространством* случайного эксперимента.

Вероятностное пространство служит математической моделью любого случайного явления; заданием этого пространства завершается аксиоматика теории вероятностей.

Свойства вероятностей

Приведем ряд свойств вероятности, являющихся следствием аксиом Колмогорова.

C1. Вероятность невозможного события равна нулю, т. е.

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

Так как $A + \emptyset = A$ и $A \cdot \emptyset = \emptyset$, то согласно аксиоме *A3*, имеем $P(A) + P(\emptyset) = P(A)$, следовательно, $P(\emptyset) = 0$.

C2. Сумма вероятностей противоположных событий равна единице, т. е.

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

Поскольку $A + \bar{A} = \Omega$, то $P(A + \bar{A}) = P(\Omega)$ а так как $A \cdot \bar{A} = \emptyset$, то в силу аксиом *A2* и *A3* получаем $P(A) + P(\bar{A}) = 1$.

C3. Вероятность любого события не превосходит единицы, т. е.

$$P(A) \leq 1.$$

Из свойства C2 вытекает, что $P(A) = 1 - P(\bar{A})$. С учетом аксиомы A1 получаем $P(A) \leq 1$.

C4. Если $A \subseteq B$, т. е. событие A влечет за собой событие B , то $P(A) \leq P(B)$.

Так как $B = (B-A) + A$ при $A \subseteq B$ и $(B-A) \cdot A = \emptyset$, то согласно аксиоме A3 получаем $P(B) = P(B-A) + P(A)$. Но $P(B-A) \geq 0$ (аксиома A1), поэтому $P(B) \geq P(A)$.

C5. Если события A_1, A_2, \dots, A_n образуют полную группу несовместных событий, т. е. $\sum_{i=1}^n A_i = \Omega$ и $A_i \cdot A_j = \emptyset$, то $\sum_{i=1}^n P(A_i) = 1$.

Так как, $A_1 + A_2 + \dots + A_n = \Omega$ то, согласно аксиомам A2 и A3, имеем $P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = 1$.

Условные вероятности

Пусть A и B – два события, рассматриваемые в данном опыте. Наступление одного события (например, A) может влиять на возможность наступления другого (B). Для характеристики зависимостей одних событий от других вводится понятие условной вероятности.

Условной вероятностью события B при условии, что произошло событие A , называется отношение вероятности произведения этих событий к вероятности события A , причем $P(A) \neq 0$, обозначается символом $P(B|A)$.

Таким образом, по определению

$$P(B|A) = \frac{P(A \cdot B)}{P(A)}, \quad P(A) \neq 0 \quad (1.5)$$

Вероятность $P(B)$, в отличие от условной, называется *безусловной вероятностью*.

Аналогично определяется условная вероятность события A при условии B , т. е. $P(A|B)$:

$$P(A|B) = \frac{P(A \cdot B)}{P(B)}, \quad P(B) \neq 0 \quad (1.6)$$

Отметим, что условная вероятность, скажем $P(B|A)$, удовлетворяет аксиомам Колмогорова. Поэтому для условной вероятности справедливы все следствия (свойства) из аксиом. Формула (1.5) принимается по определению при аксиоматическом определении вероятности; в случае классического (геометрического, статистического) определения она может быть доказана.

Пример 1.14. В урне 2 белых и 7 черных шаров. Из нее последовательно вынимают два шара. Какова вероятность того, что 2-й шар окажется белым при условии, что 1-й шар был черным?

Решим задачу двумя способами:

1. Пусть A – 1-й шар черный, B – 2-й шар белый. Так как событие A произошло, то в урне осталось 8 шаров, из которых 2 белых.

$$P(B \setminus A) = \frac{2}{8} = \frac{1}{4}.$$

2. Найдем $P(B \setminus A)$ по формуле (1.5). Очевидно, что

$$P(A) = \frac{7}{9}.$$

Находим $P(AB)$: $n = 9 \cdot 8 = 72$ – общее число исходов (появление двух шаров).

Событию AB благоприятствуют $m = C_2^1 \cdot C_7^1 = 14$ исходов. Поэтому

$$P(AB) = \frac{14}{72} = \frac{7}{36}. \text{ Следовательно, } P(B \setminus A) = \frac{7}{36} : \frac{7}{9} = \frac{1}{4}.$$

Вероятность произведения событий. Независимость событий

Используя определение условной вероятности, получим

$$P(A \cdot B) = P(A) \cdot P(B \setminus A) = P(B) \cdot P(A \setminus B) \quad (1.7)$$

т. е. *вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло.*

Равенство (1.7) называют *правилом* или *теоремой* (для схемы случаев оно доказывается) *умножения вероятностей*. Это правило обобщается на случай n событий.

Так для 3-х событий A_1, A_2, A_3 получаем

$$P(A_1 \cdot A_2 \cdot A_3) = P((A_1 \cdot A_2) \cdot A_3) = P(A_1 \cdot A_2) \cdot P(A_3 \setminus A_1 \cdot A_2) = \\ P(A_1) \cdot P(A_2 \setminus A_1) \cdot P(A_3 \setminus A_1 \cdot A_2).$$

Пример 1.15. В коробке находится 4 белых, 3 синих и 2 черных шара. Наудачу последовательно вынимают 3 шара. Какова вероятность того, что 1-й шар будет белым, 2-й – синим, 3-й – черным?

Решение. Введем следующие обозначения. A_1 – первым вытащили белый шар, A_2 – вторым – синий, A_3 – третьим – черный. Тогда интересующее нас событие A представится в виде $A = A_1 \cdot A_2 \cdot A_3$. По правилу умножения вероятностей

$$P(A_1 \cdot A_2 \cdot A_3) = P(A_1) \cdot P(A_2 \setminus A_1) \cdot P(A_3 \setminus A_1 \cdot A_2).$$

$P(A_1) = \frac{4}{9}$; $P(A_2 \setminus A_1) = \frac{3}{8}$, так как шаров осталось 8, а число благоприятных случаев

для события A_2 равно 3; $P(A_3 \setminus A_1 \cdot A_2) = \frac{2}{7}$, так как уже два шара (белый и синий)

вытащены. Следовательно, $P(A) = \frac{4}{9} \cdot \frac{3}{8} \cdot \frac{2}{7} = \frac{1}{21} \approx 0.05$.

Правило умножения вероятностей имеет особо простой вид, если события,

образующие произведение, независимы.

Событие A называется *независимым от события B* , если его условная вероятность равна безусловной, т. е. если выполняется равенство

$$P(A|B) = P(A). \quad (1.8)$$

Если событие A не зависит от события B , то и событие B не зависит от события A (лемма о взаимной независимости событий).

Можно дать следующее (новое) определение независимости событий.

Два события называются *независимыми*, если появление одного из них не меняет вероятность появления другого.

Для независимых событий правило умножения вероятностей (1.7) принимает вид:

$$P(A \cdot B) = P(A) \cdot P(B) \quad (1.9)$$

т. е. вероятность произведения двух независимых событий равна произведению вероятностей этих событий.

Равенство (1.9) часто используют в качестве определения (еще одного!) независимости событий: события A и B называются *независимыми*, если

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

Можно показать, что если события A и B независимы, то независимы события \bar{A} и B , A и \bar{B} , \bar{A} и \bar{B} .

Понятие независимости может быть распространено на случай n событий.

Вероятность суммы событий

Известно, что вероятность суммы двух несовместных событий определяется аксиомой $A3$: $P(A+B) = P(A)+P(B)$. $A \cdot B = \emptyset$. Выведем формулу суммы вероятностей двух совместных событий:

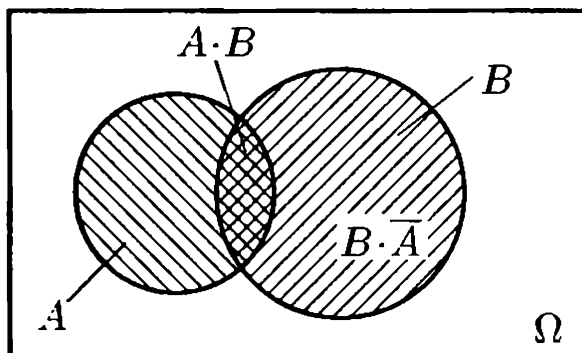
Теорема 1.1: Вероятность суммы двух совместных событий равна сумме их вероятностей без вероятности их произведения.

$$P(A+B) = P(A) + P(B) - P(A \cdot B). \quad (1.10)$$

Доказательство. Представим события $A + B$ и B в виде суммы двух *несовместных* событий:

$$A + B = A + B \cdot \bar{A}, \quad B = A \cdot B + B \cdot \bar{A}$$

В справедливости этих формул можно наглядно убедиться на рис. 10.



Тогда, согласно аксиоме АЗ, имеем

$$P(A + B) = P(A) + P(B \cdot \bar{A}) \text{ и } P(B) = P(A \cdot B) + P(B \cdot \bar{A}).$$

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

Формула (1.10) справедлива для любых событий А и В.

Можно получить формулу вероятности суммы трех и большего числа совместных событий; для трех событий она имеет вид:

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

Пример 1.16. Бросаются две игральные кости. Какова вероятность появления хотя бы одной шестерки?

Решение. Введем события А – появление шестерки на первой кости. В – на второй кости. Тогда А+В – появление хотя бы одной шестерки при бросании кости. Эти события совместные. По формуле (1.10) находим

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

Можно решить эту задачу другим способом (через обратное событие).

$$P(A + B) = 1 - P(\bar{A} \cdot \bar{B}) = 1 - \frac{5}{6} \cdot \frac{5}{6} = \frac{11}{36}.$$

Формула полной вероятности

Следствием обеих основных теорем – теоремы сложения вероятностей и теоремы умножения вероятностей – является так называемая формула полной вероятности. Вспомним, что события A_1, A_2, \dots, A_n образуют полную группу несовместных событий, если $A_i \cdot A_j = \emptyset, i \neq j$ и $\sum_{i=1}^n A_i = \Omega$

Теорема 1.2. Пусть события H_1, H_2, \dots, H_n образуют полную группу. Тогда для любого события А имеет место формула *полной вероятности*.

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A \setminus H_i) \quad (1.11)$$

Доказательство. Так как $\sum_{i=1}^n H_i = \Omega$, то в силу свойств операций над событиями $A = A \cdot \Omega = A \cdot (H_1 + H_2 + \dots + H_n) = A \cdot H_1 + A \cdot H_2 + \dots + A \cdot H_n$. Из того, что $H_i \cdot H_j = \emptyset$, следует, что и $(A \cdot H_i) \cdot (A \cdot H_j) = \emptyset, i \neq j$, т.е. события $A \cdot H_i$ и $A \cdot H_j$ также несовместны. Тогда по теореме сложения вероятностей $P(A) = \sum_{i=1}^n P(A \cdot H_i)$. По теореме умножения вероятностей

$P(A \cdot H_i) = P(H_i) \cdot P(A \setminus H_i)$, откуда и следует формула (1.11).

Отметим, что в формуле (1.11) события H_1, H_2, \dots, H_n обычно называют *гипотезами*; они исчерпывают все возможные предположения (гипотезы) относительно исходов как бы первого этапа опыта, событие A – один из возможных исходов второго этапа.

Пример 1.17. В сборочный цех завода поступает 40% деталей из I цеха и 60% – из II цеха. В I цехе производится 90% стандартных деталей, а во II – 95%. Найти вероятность того, что наудачу взятая сборщиком деталь окажется стандартной.

Взятие детали можно разбить на два этапа. Первый – это выбор цеха. Имеется две гипотезы: H_1 – деталь изготовлена I цехом, H_2 – II цехом. Вторым этапом – взятие детали. Событие A – взятая наудачу деталь стандартна. Очевидно, события H_1 и H_2 образуют полную группу, $P(H_1) = 0,4$; $P(H_2) = 0,6$. Числа 0,90 и 0,95 являются условными вероятностями события A при условии гипотез H_1 и H_2 соответственно. т. е. $P(A|H_1) = 0,90$ и $P(A|H_2) = 0,95$. По формуле (1.11) находим

$$P(A) = \sum_{i=1}^2 P(H_i) \cdot P(A|H_i) = 0,4 \cdot 0,90 + 0,6 \cdot 0,95 = 0,93.$$

Формула Байеса

При выводе формулы полной вероятности предполагалось, что вероятности гипотез известны до опыта. Формула Байеса позволяет производить переоценку первоначальных гипотез в свете новой информации, состоящей в том, что событие произошло. Поэтому формулу Байеса называют формулой уточнения гипотез. Она позволяет переоценить вероятности гипотез H_i , принятых до опыта и называемых *априорными* («a priori», доопытные, лат.) *по результатам уже проведенного опыта*, т. е. найти условные вероятности $P(H_i|A)$, которые называют *апостериорными* («a posteriori», послеопытные).

Теорема 1.3. Пусть события H_1, H_2, \dots, H_n образуют полную группу событий. Тогда условная вероятность события H_k ($k=1,2, \dots, n$) при условии, что событие A произошло, задается формулой

$$P(H_k | A) = \frac{P(H_k) \cdot P(A \setminus H_k)}{P(A)} \quad (1.12)$$

$P(A)$ – находится по формуле полной вероятности (1.11). Формула (1.12) называется *формулой Байеса*.

Доказательство. Применяя формулы условной вероятности и умножения вероятностей, имеем

$$P(H_k | A) = \frac{P(H_k \cdot A)}{P(A)} = \frac{P(H_k) \cdot P(A \setminus H_k)}{\sum_{k=1}^n P(H_k) \cdot P(A \setminus H_k)}$$

Формула Байеса позволяет пересчитывать вероятность гипотез в свете новой информации, которую дал результат A , в известном смысле она является обратной к формуле полной вероятности.

Пример 1.18. Предположим, произошла авиакатастрофа и эксперты заняты исследованием ее причин. Заранее известны 4 причины, по которым произошла катастрофа: либо причина H_1 , либо H_2 , либо H_3 , либо H_4 . По имеющейся статистике эти причины имеют следующие вероятности:

$$P(H_1)=0,2 \quad P(H_2)=0,4 \quad P(H_3)=0,3 \quad P(H_4)=0,1$$

При осмотре места катастрофы найдены следы воспламенения горючего, согласно статистике, вероятность этого события при тех или иных причинах такая:

$$P(A|H_1)=0,9 \quad P(A|H_2)=0 \quad P(A|H_3)=0,2 \quad P(A|H_4)=0,3$$

Вопрос: какая причина катастрофы наиболее вероятна?

Вычислим вероятности причин при условии наступления события A .

$$P(H_1|A)=0,2*0,9/(0,2*0,9+0,4*0+0,3*0,2+0,1*0,3)=2/3$$

$$P(H_2|A)=0$$

$$P(H_3|A)=2/9$$

$$P(H_4|A)=1/9$$

Отсюда видно, что наиболее вероятной является первая причина, так как ее вероятность максимальна.

Байесовский анализ



Точной даты рождения Томаса Байеса никто не знает. Предположительно, он родился в 1701 году. всю свою жизнь он занимался изучением теологии и математики. При его жизни работы священника не были широко известны. Лишь после его смерти в 1761 году его труд, посвященный теории вероятности, был рассмотрен английским Королевским обществом, которое объявило эту работу важной и значимой для развития науки.

Однако настоящая признательность пришла к работам Томаса Байеса только спустя 200 лет, когда были изобретены персональные компьютеры.

Теория байесовской вероятности была использована в процессе создания искусственного интеллекта. Если формулу вероятности немного упростить, то она станет доступной для многих любителей ставок на спорт.

Названий у байесовского анализа множество, однако сама суть этой теории выражается в одной формуле.

$$P(A|B) = P(A) \cdot P(B|A) / P(B)$$

Согласно формуле, вероятность того, что произойдет событие A в случае наступления события B , равна вероятности наступления события B , при условии, что гипотеза A истинна. Данное число нужно умножить на вероятность A и разделить на вероятность события B .

Байесовский анализ подойдет для любой сферы человеческой жизни, однако нагляднее всего представить действие этой формулы на примере прогноза погоды.

Предположим, что нам нужно оценить, насколько завтра вероятен дождь (событие A).

Допустим, мы берем за основу $P(A)=0.3$. Далее берем процент вероятности облачности (событие B). Предположим, $P(B) = 0.5$.

Вероятность что возможна облачность при условии, что пойдет дождь $P(B/A) = 1$. Однако анализ на этом не закончен. Ждем до следующего утра и посмотрим на небо.

Предположим, утром на небе мы заметили облака (произошло событие B).

Обновляем наш прогноз. Вероятность дождя нужно умножить на вероятность облачности в случае дождя и все поделить на вероятность облачности.

Итак, вероятность дождя будет составлять

$$P(A|B) = P(A) \cdot P(B|A) / P(B) = 0.3 \cdot 1 / 0.5 = 0.6$$

Независимые испытания. Схема Бернулли. Формула Бернулли

Последовательность n независимых испытаний, в каждом из которых может произойти некоторое событие A (успех) с вероятностью $P(A) = p$ или противоположное ему событие \bar{A} (неудача) с вероятностью $P(\bar{A}) = q = 1 - p$, называется *схемой Бернулли*.

Например, при стрельбе по мишени: событие A – попадание (успех), событие \bar{A} – промах (неудача).

Простейшая задача, относящаяся к схеме Бернулли, состоит в определении вероятности того, что в n независимых испытаниях событие A наступит m раз, где $0 \leq m \leq n$. Обозначается искомая вероятность так: $P_n(m)$ или $P_{n,m}$.

Например, при бросании игральной кости 3 раза $P_3(2)$ означает вероятность того, что в 3-х опытах событие A – выпадение, например, цифры 4 – произойдет 2 раза.

Теорема 1.4. Если производится n независимых испытаний, в каждом из которых вероятность появления события A равна p , а вероятность его не появления равна $q = 1 - p$, то вероятность того, что событие A произойдет m раз определяется *формулой Бернулли*

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}, \quad m = 0, 1, 2, \dots, n \quad (1.13)$$

Можно заметить, что вероятности $P_n(m)$, $m = 0, 1, \dots, n$ являются коэффициентами при x^m в разложении $(q+px)^n$ по формуле бинома Ньютона:

$$(q + px)^n = q^n + C_n^1 q^{n-1} px + C_n^2 q^{n-2} p^2 x^2 + \dots + C_n^m q^{n-m} p^m x^m + \dots + p^n x^n.$$

Поэтому, совокупность вероятностей $P_n(m)$ называют *биномиальным*

законом распределения вероятностей, а функцию $\varphi(x) = (q + px)^n$ – производящей функцией для последовательности независимых опытов.

Если в каждом из независимых испытаний вероятности наступления события A разные, то вероятность того, что событие A наступит m раз в n опытах, равна коэффициенту при m -ой степени многочлена

$$\varphi_n(z) = (q_1 + p_1 \cdot z)(q_2 + p_2 \cdot z) \cdot \dots \cdot (q_n + p_n \cdot z),$$

где φ_n – производящая функция.

Пример 1.19. Производится 3 независимых выстрела по цели. Вероятности попадания при разных выстрелах одинаковы и равны $p = 0,9$, Какова вероятность: а) промаха; б) одного попадания; в) двух попаданий; г) трех попаданий?

Решить задачу в случае, если вероятности попадания при разных выстрелах различны: $p_1=0,7$; $p_2=0,8$; $p_3=0,9$.

Решение.

а) $P_3(0) = C_3^0 \cdot 0,9^0 \cdot 0,1^3 = 0,001$ – вероятность трех промахов;

б) $P_3(1) = C_3^1 \cdot 0,9^1 \cdot 0,1^2 = 0,027$ – вероятность одного попадания;

в) $P_3(2) = C_3^2 \cdot 0,9^2 \cdot 0,1^1 = 0,243$ – вероятность двух попаданий;

г) $P_3(3) = C_3^3 \cdot 0,9^3 \cdot 0,1^0 = 0,729$ – вероятность трех попаданий;

Если вероятности при разных выстрелах различны, то производящая функция имеет вид:

$$\begin{aligned} \varphi_3(z) &= (0,3 + 0,7 \cdot z)(0,2 + 0,8 \cdot z)(0,1 + 0,9 \cdot z) = \\ &= 0,504z^3 + 0,398z^2 + 0,092z + 0,006. \end{aligned}$$

Откуда находим искомые вероятности:

$$P_3(3) = 0,504; P_3(2) = 0,398; P_3(1) = 0,092; P_3(0) = 0,006;$$

Предельные теоремы в схеме Бернулли

Использование формулы Бернулли (1.13) при больших значениях n и m вызывает большие трудности, так как это связано с громоздкими вычислениями. Рассмотрим три предельные теоремы, содержащие асимптотические формулы для вычисления биномиальной вероятности $P_n(m)$ при $n \rightarrow \infty$.

Теорема Пуассона. Если число испытаний неограниченно увеличивается ($n \rightarrow \infty$) и вероятность p наступления события A в каждом испытании неограниченно уменьшается ($p \rightarrow 0$), но так, что их произведение np является постоянной величиной ($np = a = \text{const}$), то вероятность $P_n(m)$ удовлетворяет предельному равенству

$$\lim_{n \rightarrow \infty} P_n(m) = \frac{a^m e^{-a}}{m!}. \quad (1.14)$$

Из предельного равенства (1.14) при больших n и малых p (редкие события) вытекает приближенная формула Пуассона

$$P_n(m) \approx \frac{a^m e^{-a}}{m!}, \quad a = np, \quad m = 0, 1, 2, \dots \quad (1.15)$$

Пример 1.20. Телефонная станция обслуживает 2000 абонентов. Вероятность позвонить любому абоненту в течение часа равна 0,003. Какова вероятность того, что в течение часа позвонят 5 абонентов?

Решение.

$$a = n \cdot p = 2000 \cdot 0.003 = 6 \qquad P_5 = \frac{6^5 \cdot e^{-6}}{5!} \approx 0,13$$

В тех случаях, когда число испытаний n велико, а вероятность p не близка к нулю ($p \neq 0, p \neq 1$), для вычисления биномиальных вероятностей используют теоремы Муавра-Лапласа.

Локальная теорема Муавра-Лапласа. Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, а число независимых испытаний достаточно велико, то вероятность $P_n(m)$ может быть вычислена по приближенной формуле

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \text{где } x = \frac{m - np}{\sqrt{npq}} \quad (1.16)$$

Пример 1.21. Вероятность попадания в мишень при одном выстреле для данного стрелка равна 0,7. Найти вероятность того, что при 200 выстрелах мишень будет поражена 160 раз.

Решение.

$$p=0.7; \quad q=1-p=0.3; \quad m=160.$$

Подставляя эти данные в формулу (1.16), получим $P_{200}(160) \approx 0,005$.

В тех случаях, когда требуется вычислить вероятность того, что в n независимых испытаниях событие A появится не менее k_1 раз, но не более k_2 раз, используют интегральную теорему Муавра-Лапласа (является частным случаем более общей теоремы – центральной предельной теоремы).

Интегральная теорема Муавра-Лапласа. Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k_1 \leq m \leq k_2)$ может быть найдена по приближенной формуле

$$P_n(k_1 \leq m \leq k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx, \quad \text{где } x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}} \quad (1.17)$$

Используя нормированную функцию Лапласа $\Phi_0(x)$ (будет введена позже при

рассмотрении нормального закона распределения)

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt,$$

равенство (1.17) примет вид

$$P_n(k_1 \leq m \leq k_2) \approx \Phi_0(x_2) - \Phi_0(x_1), \text{ где } x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}} \quad (1.18)$$

Эту формулу обычно используют на практике.

Имеются таблицы приближенных значений $\Phi_0(x)$ (интеграл не берется в элементарных функциях).

Пример 1.22. Проверкой установлено, что цех в среднем выпускает 96% продукции высшего качества. Приемщик проверяет 200 изделий этого цеха. Если среди них окажется более 10 изделий не высшего сорта, то вся партия изделий бракуется. Какова вероятность, что партия будет принята?

Решение.

Здесь $n = 200$, $p = 0.04$ (вероятность бракованного изделия), $q = 0.96$. Вероятность принятия всей партии можно найти по формуле (1.18).

$k_1 = 0$, $k_2 = 10$.

Находим $x_1 = \frac{0 - 200 \cdot 0.04}{\sqrt{200 \cdot 0.04 \cdot 0.96}} \approx -2.89$, $x_2 = \frac{10 - 200 \cdot 0.04}{\sqrt{200 \cdot 0.04 \cdot 0.96}} \approx 0.72$.

$$P_n(0 \leq m \leq 10) \approx \Phi_0(0.72) - \Phi_0(-2.89) = 0.26424 + 0.49807 = 0.7623$$

2. Случайные величины

Понятие случайной величины. Закон распределения случайной величины

Одним из важнейших понятий теории вероятностей (наряду со случайным событием и вероятностью) является понятие случайной величины.

Под *случайной величиной* понимают величину, которая в результате опыта принимает то или иное значение, причем неизвестно заранее, какое именно.

Случайные величины обозначаются прописными латинскими буквами X, Y, Z, \dots , а принимаемые ими значения соответственно малыми буквами $x_1, x_2, \dots, y_1, y_2, \dots$.

Примерами случайных величин могут служить:

- X – число очков, появляющихся при бросании игральной кости;
- Y – число выстрелов до первого попадания в цель;
- Z – время безотказной работы прибора и т. п. (рост человека, курс доллара, количество бракованных деталей в партии, температура воздуха, выигрыш игрока, прибыль фирмы, ...).

Случайная величина, принимающая конечное или счетное множество значений, называется *дискретной*.

Если же множество возможных значений случайной величины несчетно, то такая величина называется *непрерывной*.

То есть дискретная случайная величина принимает отдельные изолированные друг от друга значения, а непрерывная случайная величина может принимать любые значения из некоторого промежутка (например, значения на отрезке, на всей числовой прямой и т.д.). Отметим, что рассматриваются также случайные величины смешанного типа.

Дадим теперь строгое определение случайной величины, исходя из теоретико-множественной трактовки основных понятий теории вероятностей.

Случайной величиной X называется числовая функция, определенная на пространстве элементарных событий Ω , которая каждому элементарному событию w ставит в соответствие число $X(w)$, т. е. $X = f(w)$, $w \in \Omega$.

Для полного описания с. в. недостаточно лишь знания ее возможных значений; необходимо еще знать вероятности этих значений.

Любое правило (таблица, функция, график), позволяющее находить вероятности произвольных событий $A \subseteq S$ (S – σ -алгебра событий пространства Ω), в частности, указывающее вероятности отдельных значений случайной величины или множества этих значений, называется *законом распределения случайной величины* (или просто: *распределением*). Про случайную величину говорят, что «она подчиняется данному закону распределения».

Закон распределения дискретной случайной величины. Многоугольник распределения

Пусть X – дискретная случайная величина, которая принимает значения $x_1, x_2, x_3, \dots, x_n, \dots$ (множество этих значений конечно или счетно) с некоторой вероятностью p_i , где $i = 1, 2, 3, \dots, n$. Закон распределения дискретной случайной величины удобно задавать с помощью формулы $p_i = P\{X = x_i\}$, $i = 1, 2, 3, \dots, n, \dots$, определяющей вероятность того, что в результате опыта случайная величина X примет значение x_i . Для дискретной случайной величины X закон распределения может быть задан в виде *таблицы распределения*:

X	x_1	x_2	\dots		x_n	\dots
p	p_1	p_2	\dots		p_n	\dots

где первая строка содержит все возможные значения (обычно в порядке возрастания) случайной величины, а вторая – их вероятности. Такую таблицу называют *рядом распределения*.

Так как события $\{X = x_1\}, \{X = x_2\} \dots$ несовместны и образуют полную группу, то сумма их вероятностей равна единице, т.е. $\sum_i p_i = 1$.

Закон распределения дискретной случайной величины можно задать *графически*, если на оси абсцисс отложить возможные значения с. в., а на оси ординат – вероятности этих значений. Ломаную, соединяющую последовательно точки $(x_1, p_1), (x_2, p_2), \dots$ называют *многоугольником* (или *полигоном*) *распределения* (см. рис. 10).

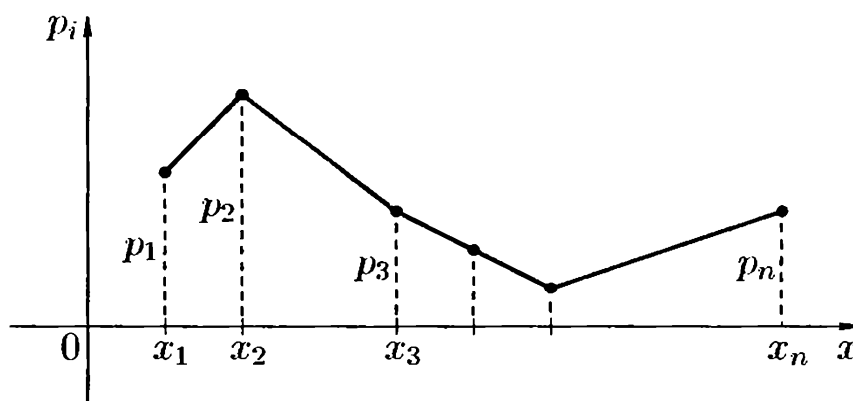


Рис. 10

Пример 2.1. В урне 8 шаров, из которых 5 белых, остальные – черные. Из нее вынимают наудачу 3 шара. Найти закон распределения числа белых шаров в выборке.

Решение. Возможные значения случайной величины X – числа белых шаров в выборке есть $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$. Вероятности их соответственно будут

$$p_1 = P\{X = 0\} = \frac{C_5^0 \cdot C_3^3}{C_8^3} = \frac{1}{56}, \quad p_2 = P\{X = 1\} = \frac{C_5^1 \cdot C_3^2}{C_8^3} = \frac{15}{56},$$

$$p_3 = P\{X = 2\} = \frac{C_5^2 \cdot C_3^1}{C_8^3} = \frac{30}{56}, \quad p_4 = P\{X = 3\} = \frac{C_5^3 \cdot C_3^0}{C_8^3} = \frac{10}{56}.$$

Закон распределения запишем в виде таблицы.

X	0	1	2	3
p	1/56	15/56	30/56	10/56

Функция распределения и ее свойства. Функция распределения дискретной случайной величины

Очевидно, ряд распределения случайной величины может быть построен только для дискретной случайной величины, т.к. для непрерывной случайной величины нельзя даже перечислить все ее возможные значения. Кроме того, можно показать, что вероятность каждого отдельно взятого значения непрерывной случайной величины равна нулю.

Для характеристики поведения непрерывной случайной величины целесообразно использовать вероятность события $\{X < x\}$ (а не $\{X = x\}$), где x – некоторое действительное число. С точки зрения практики нас мало интересует событие, состоящее, например, в том, что лампочка проработает ровно 900 часов, т. е. $X = 900$. Более важным является событие вида $\{X < 900\}$ (или $\{X > 900\}$). Такое событие имеет ненулевую вероятность.

Универсальным способом задания закона распределения вероятностей, пригодным как для дискретных, так и для непрерывных случайных величин, является ее функция распределения, обозначаемая $F(x)$.

Функцией распределения случайной величины X называется функция $F(x)$, которая для любого числа $x \in R$ равна вероятности события $\{X < x\}$.

$$F(x) = P\{X < x\} \quad \text{т.е.} \quad F(x) = P\{\omega: X(\omega) < x\}. \quad (2.1)$$

Функцию $F(x)$ называют также *интегральной функцией распределения*. $F(x)$ обладает следующими свойствами:

1. $F(x)$ ограничена, т. е.

$$0 \leq F(x) \leq 1.$$

2. $F(x)$ – неубывающая функция на R , т.е. если $x_2 > x_1$, то

$$F(x_2) \geq F(x_1).$$

3. $F(x)$ обращает в ноль на минус бесконечности и равна единице в плюс бесконечности, т. е.

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

4. Вероятность попадания случайной величины X в промежуток $[a, b)$ равна приращению ее функции распределения на этом промежутке, т. е.

$$P\{a \leq X < b\} = F(b) - F(a). \quad (2.2)$$

5. $F(x)$ непрерывна слева, т. е.

$$\lim_{x \rightarrow x_0 - 0} F(x) = F(x_0).$$

С помощью функции распределения можно вычислить вероятность события $\{X \geq x\}$:

$$P(X \geq x) = 1 - F(x). \quad (2.3)$$

Функция распределения дискретной случайной величины имеет вид

$$F(x) = \sum_{x_i < x} p_i. \quad (2.4)$$

Здесь суммирование ведется по всем i , для которых $x_i < x$. Равенство (2.4) непосредственно вытекает из определения (2.1).

Пример 2.2. По условию примера 2.1 найти функцию распределения $F(x)$ и построить ее график.

Решение. Будем задавать различные значения x и находить для них $F(x) = P\{X < x\}$:

1. Если $x \leq 0$, то, очевидно, $F(x) = P\{X < 0\} = 0$;

2. Если $0 < x \leq 1$, то $F(x) = P\{X < x\} = P\{X = 0\} = \frac{1}{56}$;

3. Если $1 < x \leq 2$, то $F(x) = P\{X = 0\} + P\{X = 1\} = \frac{1}{56} + \frac{15}{36} = \frac{16}{56}$;

4. Если $2 < x \leq 3$, то $F(x) = P\{X = 0\} + P\{X = 1\} + P\{X = 2\} = \frac{1}{56} + \frac{15}{36} + \frac{30}{56} = \frac{46}{56}$;

5. Если $3 < x$, то $F(x) = P\{X = 0\} + P\{X = 1\} + P\{X = 2\} + P\{X = 3\} +$
 $+\frac{1}{56} + \frac{15}{36} + \frac{30}{56} + \frac{10}{56} = 1.$

Функция распределения:

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ \frac{1}{56}, & \text{если } 0 < x \leq 1; \\ \frac{16}{56}, & \text{если } 1 < x \leq 2; \\ \frac{46}{56}, & \text{если } 2 < x \leq 3; \\ 1, & \text{если } 3 < x. \end{cases}$$

Как видим (рис. 11), функция распределения дискретной случайной величины X есть разрывная, со скачками p_i в точках x_i , функция, «непрерывная слева» (при подходе к точке разрыва слева функция $F(x)$ сохраняет значение). Ее график имеет ступенчатый вид.

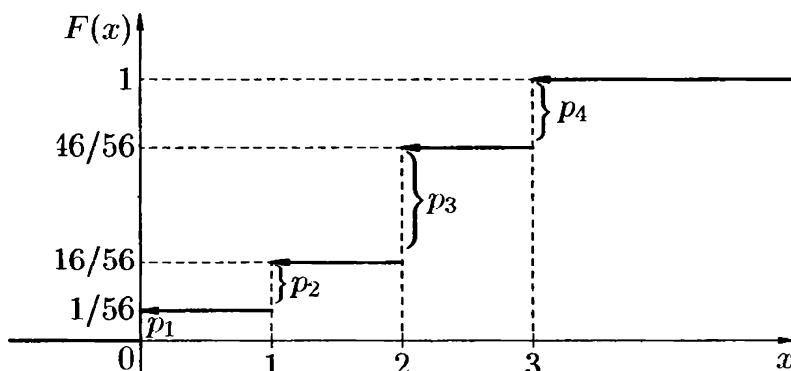


Рис.11

Отметим, что, пользуясь равенством (2.4), функцию распределения можно сразу записать в виде (2.5)

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{1}{56}, & 0 < x \leq 1; \\ \frac{1}{56} + \frac{15}{56}, & 1 < x \leq 2; \\ \frac{1}{56} + \frac{15}{56} + \frac{30}{56}, & 2 < x \leq 3; \\ \frac{1}{56} + \frac{15}{56} + \frac{30}{56} + \frac{10}{56}, & 3 < x. \end{cases} \quad (2.5)$$

Плотность распределения и ее свойства

Важнейшей характеристикой *непрерывной случайной величины* (помимо функции распределения) является плотность распределения вероятностей. Случайная величина X называется *непрерывной*, если ее функция распределения непрерывна и дифференцируема всюду, кроме, быть может, отдельных точек.

Плотностью распределения вероятностей $f(x)$ (плотностью распределения) непрерывной случайной величины X называется производная ее функции распределения.

$$f(x) = F'(x). \quad (2.6)$$

Функцию $f(x)$ называют также *дифференциальной функцией распределения*; она является одной из форм закона распределения случайной величины, существует только для непрерывных случайных величин.

Установим вероятностный смысл плотности распределения. Из определения производной следует

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta F(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Но согласно формуле (2.2) $F(x + \Delta x) - F(x) = P\{x \leq X < x + \Delta x\}$.

Отношение $\frac{P\{x \leq X < x + \Delta x\}}{\Delta x}$ представляет собой среднюю вероятность, которая приходится на единицу длины участка $[x, x + \Delta x]$, т.е. среднюю плотность распределения вероятности. Тогда

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P\{x \leq X < x + \Delta x\}}{\Delta x}, \quad (2.7)$$

т.е. плотность распределения есть предел отношения вероятности попадания случайной величины в промежуток $[x, x + \Delta x]$ к длине Δx этого промежутка, когда Δx стремится к нулю. Из равенства (2.7) следует, что

$$P\{x \leq X < x + \Delta x\} \approx f(x)\Delta x.$$

Плотность распределения $f(x)$ обладает следующими свойствами:

1. $f(x)$ неотрицательная, т. е.

$$f(x) \geq 0.$$

2. Вероятность попадания случайной величины в промежуток $[a, b]$ равна определенному интегралу от ее плотности в пределах от a до b , т. е.

$$P(a \leq X \leq b) = \int_a^b f(x) dx. \quad (2.8)$$

Геометрически эта вероятность равна площади S фигуры, ограниченной сверху кривой распределения $f(x)$ и опирающейся на отрезок $[a, b]$ (Рис. 12).

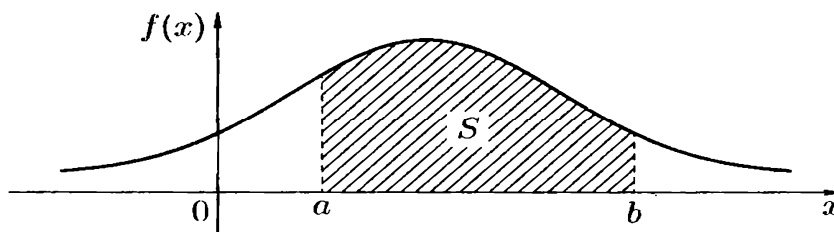


Рис.12

3. Функция распределения непрерывной случайной величины может быть выражена через ее плотность вероятности по формуле

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

4. Условие нормировки: несобственный интеграл от $f(x)$ в бесконечных пределах равен единице, т. е.

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

Пример 2.3. Плотность распределения случайной величины X задана функцией

$$f(x) = \frac{a}{1+a^2}.$$

Найти значение параметра a .

Решение. Согласно свойству 4 плотности распределения вероятности, имеем

$$\int_{-\infty}^{\infty} \frac{a}{1+x^2} dx = 1, \text{ т.е. } a \cdot \lim_{\substack{d \rightarrow +\infty, \\ c \rightarrow -\infty}} \int_c^d \frac{a}{1+x^2} dx = 1, \text{ т.е. } a \cdot \lim_{\substack{d \rightarrow +\infty, \\ c \rightarrow -\infty}} \operatorname{arctg} \Big|_c^d = 1$$

$$\text{или } a \cdot \left(\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right) = 1, \text{ отсюда } a = \frac{1}{\pi}.$$

Числовые характеристики случайных величин

При решении многих практических задач достаточно знать лишь некоторые *числовые параметры, характеризующие отдельные существенные свойства закона распределения случайной величины.* Такие числа принято называть *числовыми характеристиками случайной величины.*

Важнейшими среди них являются характеристики положения: математическое ожидание (центр распределения случайной величины), мода, медиана; характеристики рассеяния: дисперсия (отклонение значений случайной величины от ее центра), среднее квадратическое отклонение.

Математическое ожидание

Математическим ожиданием (или *средним значением*) дискретной случайной величины X , имеющей закон распределения $p_i = P\{X=x_i\}$, $i = 1, 2, 3, \dots, n$, называется число, равное сумме произведений всех ее значений на соответствующие им вероятности.

$$MX = M[X] = m_X = \sum_{i=1}^n x_i \cdot p_i \quad (2.9)$$

Вероятностный смысл математического ожидания состоит в том, что оно является средним значением случайной величины. Действительно, т. к. $\sum_{i=1}^n p_i = 1$

$$M[X] = \sum_{i=1}^n x_i \cdot p_i = \frac{\sum_{i=1}^n x_i \cdot p_i}{\sum_{i=1}^n p_i} = x_{\text{среднее}}$$

Математическим ожиданием непрерывной случайной величины X с плотностью вероятности $f(x)$, называется число

$$M[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx \quad (2.10)$$

Свойства математического ожидания

1. Математическое ожидание (м. о.) постоянной равно самой этой постоянной, т. е.

$$Mc = c.$$

2. Постоянный множитель выносится за знак м. о., т. е.

$$M(cX) = cMX.$$

3. М. о. суммы случайных величин равно сумме их м. о., т. е.

$$M(X + Y) = MX + MY.$$

4. М. о. отклонения случайной величины от ее м. о. равно нулю, т. е.

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

$X - MX$ называется центрированной случайной величиной и обозначается $\overset{0}{X}$.

5. М. о. произведения независимых случайных величин равно произведению их м. о., т. е. если X и Y независимы, то

$$M(X \cdot Y) = MX \cdot MY.$$

Свойства м. о. для дискретной случайной величины остаются справедливы и для непрерывных случайных величин.

Пример 2.4. В лотерее имеется 1000 билетов, из них выигрышных: 10 по 500 р., 50 по 50 р., 100 по 10 р., 150 по 1 р. Найти математическое ожидание выигрыша на один билет.

Решение. Ряд распределения случайной величины X – суммы выигрыша на один билет таков:

X	500	50	10	1	0
p	0.01	0.05	0.1	0.15	0.69

Находим $M[X] = 500 \cdot 0.01 + 50 \cdot 0.05 + 10 \cdot 0.1 + 1 \cdot 0.15 + 0 \cdot 0.69 = 8.65$ р.

Дисперсия

Дисперсией (рассеянием) случайной величины X называется математическое ожидание квадрата ее отклонения от своего математического ожидания.

$$DX = D[X] = D_X = M(X - MX)^2 \quad (2.11)$$

или $DX = M \overset{0}{X}^2$, или $DX = M(X - m_x)^2$.

Дисперсия характеризует разброс значений случайной величины X относительно ее математического ожидания. Из определения дисперсии следуют формулы для ее вычисления:

$$D[X] = \sum_{i=1}^n (x_i - m_x)^2 \cdot p_i \text{ – дискретная случайная величина.}$$

$$D[X] = \int_{-\infty}^{\infty} (x - m_x)^2 \cdot f(x) dx \text{ – непрерывная случайная величина.}$$

На практике дисперсию удобно находить по формуле

$$D[X] = MX^2 - (MX)^2.$$

Тогда формулы для вычисления дисперсии будут такие

$$D[X] = \sum_{i=1}^n x_i^2 \cdot p_i - m_x^2. \quad (2.12)$$

$$D[X] = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx - m_x^2. \quad (2.13)$$

Свойства дисперсии

1. Дисперсия постоянной равна нулю, т. е.

$$Dc = 0.$$

2. Постоянный множитель можно выносить за знак дисперсии, возведя его в квадрат, т. е.

$$DcX = c^2DX.$$

3. Дисперсия суммы независимых случайных величин равна сумме их дисперсий, т. е. если X и Y независимы, то

$$D(X + Y) = DX + DY.$$

4. Дисперсия случайной величины не изменится, если к этой случайной величине прибавить постоянную, т. е.

$$D(X + c) = DX.$$

5. Если случайные величины X и Y независимы, то

$$D(XY) = MX^2 \cdot MY^2 - (MX)^2 \cdot (MY)^2$$

Среднее квадратическое отклонение

Дисперсия DX имеет размерность квадрата случайной величины X , что в сравнительных целях неудобно. Когда желательно, чтобы оценка разброса (рассеяния) имела размерность случайной величины, используют еще одну числовую характеристику – среднее квадратическое отклонение.

Средним квадратическим отклонением или стандартным отклонением случайной величины X называется квадратный корень из ее дисперсии

$$\sigma_x = \sqrt{DX} \quad (2.14)$$

Для изучения свойств случайного явления, независящих от выбора масштаба измерения и положения центра группирования, исходную случайную величину X приводят к некоторому виду $Z = \frac{X - MX}{\sigma_x}$ – стандартная случайная величина. Ее математическое ожидание равно 0, а дисперсия равна 1.

Мода и медиана. Моменты случайных величин. Асимметрия и эксцесс. Квантили

Модой дискретной случайной величины X называется ее значение, принимаемое с наибольшей вероятностью по сравнению с двумя соседними значениями, обозначается через M_0X . Для непрерывной случайной величины M_0X – точка максимума (локального) плотности $f(x)$.

Если мода единственна, то распределение с. в. называется *унимодальным*, в противном случае — *полимодальным* (рис. 13).

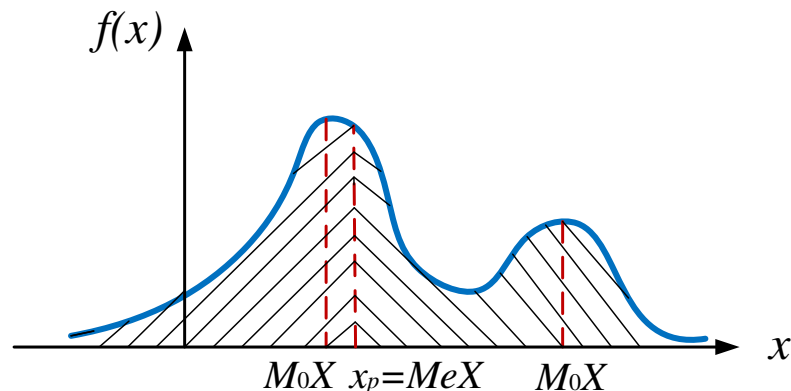


Рис. 13

Медианой M_eX непрерывной случайной величины X называется такое ее значение x_p , для которого

$$P\{X < x_p\} = P\{X > x_p\} = \frac{1}{2}$$

С помощью функции распределения $F(x)$ это равенств можно записать в виде $F(M_eX) = 1 - F(M_eX)$. Отсюда $F(M_eX) = 0,5$.

Для дискретной случайной величины медиана обычно не определяется.

Математическое ожидание и дисперсия являются частными случаями следующих более общих понятий – моментов случайных величин.

Начальным моментом порядка k случайной величины X называется математическое ожидание k -й степени этой величины, обозначается через α_k .

$$\alpha_k = M(X^k).$$

В частности, начальный момент 1-го порядка есть математическое ожидание.

Центральным моментом порядка k случайной величины X называется математическое ожидание величины $(X - MX)^k$, обозначается через μ_k .

$$\mu_k = M(X - MX)^k.$$

В частности, $\mu_2 = DX$, т. е. центральный момент 2-го порядка есть дисперсия.

Среди моментов высших порядков особое значение имеют центральные моменты 3-го и 4-го порядков, называемых соответственно коэффициентами асимметрии и эксцесса.

Коэффициентом асимметрии («скошенности») A_S случайной величины X называется величина

$$A_S = \frac{\mu_3}{\sigma_x^3} = \frac{M(X - MX)^3}{(DX)^{\frac{3}{2}}}.$$

Если $A_S > 0$, то кривая распределения более пологая справа от $M_Q X$, если $A_S < 0$, то кривая распределения более пологая слева от $M_Q X$ (рис. 14).

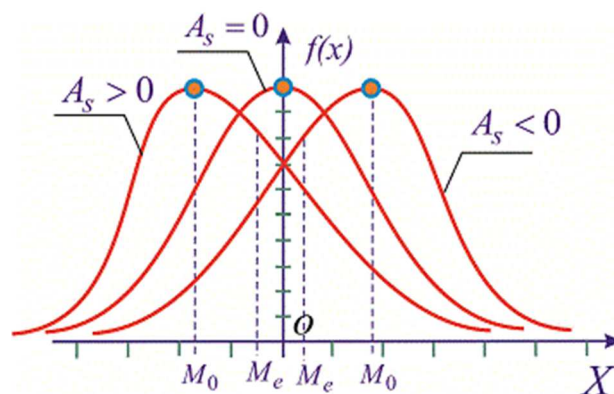


Рис. 14

Коэффициентом эксцесса («островершинности») E_S случайной величины X называется величина

$$E_s = \frac{\mu_4}{\sigma_x^4} - 3 = \frac{M(X - MX)^4}{(DX)^2} - 3.$$

Величина E_s характеризует островершинность или плосковершинность распределения (рис. 15).

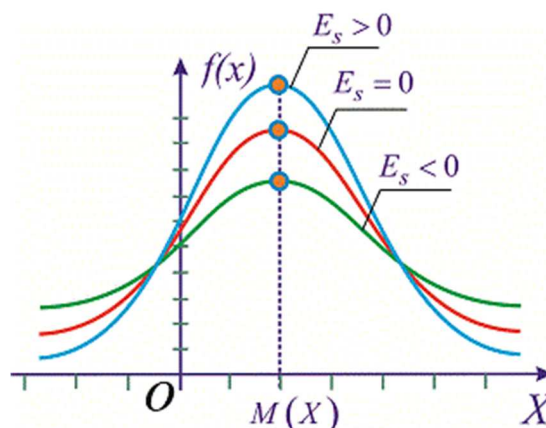


Рис.15

Кроме рассмотренных выше числовых характеристика случайных величин, в приложениях используются так называемые квантили.

Квантилем уровня p случайной величины X называется решение уравнения

$$F(x_p) = p,$$

где p – некоторое число, $0 < p < 1$).

Квантили $x_{0,25}$, $x_{0,5}$, $x_{0,75}$ имеют свои названия: *нижняя квартиль*, *медиана*, *верхняя квартиль* соответственно. Они делят числовую прямую на 4 части, вероятности попадания в которые равны 0,25 (рис. 16).

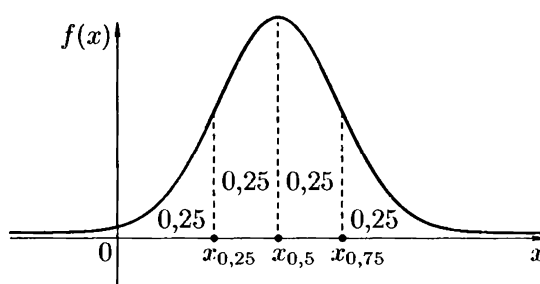


Рис. 16

Основные законы распределения случайных величин

Биномиальный закон распределения

Среди законов распределения дискретной случайной величины наиболее распространенным является биномиальное распределение, которое уже рассматривалось ранее (п. 1.16).

| Дискретная случайная величина X имеет *биномиальное распределение* (или

распределена по биномиальному закону), если она принимает значения $0, 1, 2, 3, \dots, n$ с вероятностями:

$$p_m = P\{X = m\} = C_n^m \cdot p^m \cdot q^{n-m}, \quad (2.15)$$

где $0 < p < 1, q = 1 - p, m = 0, 1, \dots, n$.

Случайная величина X , распределенная по биномиальному закону, является числом успехов с вероятностью p в схеме Бернулли проведения n независимых опытов.

Ряд распределения дискретной случайной величины X , имеющей биномиальное распределение:

$X=m$	0	1	2	...	m	...	n
$p_m=P\{X=m\}$	q^n	$C_n^1 p^1 q^{n-1}$	$C_n^2 p^2 q^{n-2}$...	$C_n^m p^m q^{n-m}$...	p^n

Функция распределения случайной величины X , распространенной по биномиальному закону имеет вид:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 0 \\ \sum_{m < x} C_n^m p^m q^{n-m}, & \text{при } 0 < x \leq n \\ 1, & \text{при } n < x. \end{cases}$$

Найдем числовые характеристики этого распределения, используя понятие производящей функции $\varphi(z)$ для дискретной случайной величины X , имеющей биномиальное распределение:

$$\begin{aligned} \varphi(z) &= \sum_{m=0}^n C_n^m p^m q^{n-m} z^m = \sum_{m=0}^n C_n^m (pz)^m q^{n-m} = (q + pz)^n, \\ \varphi(z) &= \sum_{k=0}^{\infty} p_k z^k = p_0 + p_1 z + p_2 z^2 + \dots, \end{aligned}$$

Можно получить:

$$MX = \alpha_1 = \varphi'(1), \quad DX = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2. \quad (2.16)$$

$$\varphi'(z) = n(q + pz)^{n-1} p, \quad \varphi''(z) = n(n-1)p^2(q + pz)^{n-2}.$$

Отсюда из (2.16)

$$MX = np \quad \quad \quad DX = npq$$

Распределение Пуассона

Дискретная случайная величина X имеет *распределение Пуассона*, если ее возможные значения: $0, 1, 2, \dots, m, \dots$ (счетное множество значений), а соответствующие вероятности выражаются формулой Пуассона

$$p_m = P\{X = m\} = \frac{a^m \cdot e^{-a}}{m!}, \quad (2.17)$$

где $m = 0, 1, 2, \dots$, a – параметр.

Распределение Пуассона является предельным для биномиального, когда $n \rightarrow \infty$ и $p \rightarrow 0$ так, что $np = a$ – постоянно. Примерами случайных величин, имеющих распределение Пуассона, являются: число вызовов на телефонной станции за время t ; число опечаток в большом тексте; число бракованных деталей в большой партии; число α -частиц, испускаемых радиоактивным источником, и т.д. При этом считается, что события появляются независимо друг от друга с постоянной *средней интенсивностью*, характеризующейся параметром $a = np$.

Случайная величина X , распределенная по закону Пуассона, имеет следующий ряд распределения

$X=m$	0	1	2	m	...
P_m	e^{-a}	$\frac{a \cdot e^{-a}}{1!}$	$\frac{a^2 \cdot e^{-a}}{2!}$...	$\frac{a^m \cdot e^{-a}}{m!}$...

Найдем математическое ожидание и дисперсию случайной величины X , распределенной по закону Пуассона. Производящей функцией распределения Пуассона будет

$$\varphi(z) = \sum_{m=0}^{\infty} \frac{a^m e^{-a}}{m!} z^m = e^{-a} \sum_{m=0}^{\infty} \frac{(az)^m}{m!} = e^{-a} \cdot e^{az} = e^{a(z-1)}.$$

Отсюда из (2.16)

$$MX = DX = a$$

Таким образом, если значения математического ожидания и дисперсии случайной величины близки между собой, есть основание считать, что случайная величина распределена по закону Пуассона.

Пример 2.5. Вероятность попадания в цель при одном выстреле равна 0,01. Какова вероятность того, что число попаданий при 200 выстрелах составит не менее 5 и не более 10?

Решение. Вероятность $p = 0.01$ очень мала, а число выстрелов достаточно велико. Поэтому искомую вероятность будем находить, используя формулу Пуассона. Случайная величина X – число попаданий. Требуется найти $P\{5 \leq X \leq 10\}$. По теореме сложения вероятностей

$$P\{5 \leq X \leq 10\} = P\{X = 5\} + P\{X = 6\} + \dots + P\{X = 10\}.$$

Имеем: $a = np = 200 \cdot 0.01 = 2$; $e^{-2} \approx 0.135$.

$$P\{5 \leq X \leq 10\} = 0.135 \left(\frac{2^5}{5!} + \frac{2^6}{6!} + \frac{2^7}{7!} + \frac{2^8}{8!} + \frac{2^9}{9!} + \frac{2^{10}}{10!} \right) \approx 0.053.$$

Равномерный закон распределения

Непрерывная случайная величина X имеет *равномерное распределение* на отрезке $[a, b]$, если ее плотность вероятности $f(x)$ постоянна на этом отрезке, а вне его равна нулю:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{при } x \in [a, b] \\ 0, & \text{при } x \notin [a, b] \end{cases} \quad (2.18)$$

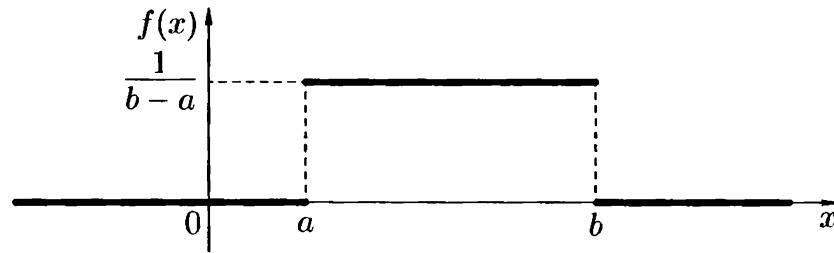


Рис. 17. График плотности вероятности равномерного распределения

Найдем функцию распределения $F(x)$. Согласно формуле $F(x) = \int_{-\infty}^x f(t) dt$ будем иметь:

$$F(x) = \begin{cases} 0, & \text{при } x \leq a, \\ \frac{x-a}{b-a}, & \text{при } a < x \leq b, \\ 1, & \text{при } b < x. \end{cases}$$

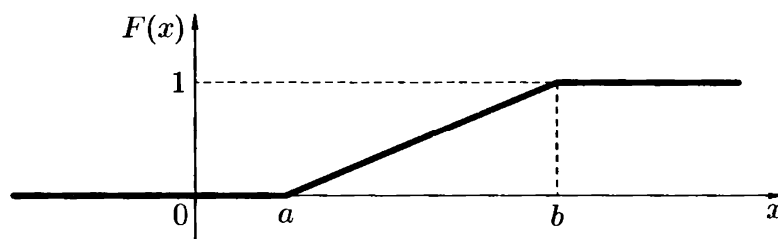


Рис. 18. График функции распределения равномерного распределения

Определим MX и DX

$$MX = \int_{-\infty}^a x \cdot 0 dx + \int_a^b \frac{x}{b-a} dx + \int_b^{+\infty} x \cdot 0 dx = \frac{a+b}{2}.$$

$$DX = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{dx}{b-a} = \frac{1}{b-a} \cdot \frac{1}{3} \left(x - \frac{a+b}{2}\right)^3 \Big|_a^b =$$

$$= \frac{1}{3(b-a)} \left(\frac{(b-a)^3}{8} - \frac{(a-b)^3}{8} \right) = \frac{(b-a)^2}{12}.$$

К случайным величинам, имеющим равномерное распределение, относятся: время ожидания пассажиром транспорта, курсирующего с определенным интервалом; ошибка округления числа до целого (она равномерно распределена на отрезке $[-0,5; 0,5]$). И вообще случайные величины, о которых известно, что все ее значения лежат внутри некоторого интервала и все они имеют одинаковую вероятность (плотность).

Показательный закон распределения

Непрерывная случайная величина X имеет *показательный* (или *экспоненциальный*) закон распределения, если ее плотность вероятности имеет вид

$$f(x) = \begin{cases} 0, & \text{при } x < 0, \\ \lambda e^{-\lambda x}, & \text{при } x \geq 0, \end{cases} \quad (2.19)$$

где $\lambda > 0$ – параметр распределения.

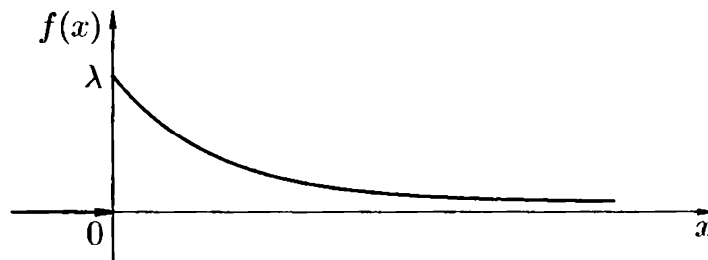


Рис. 19. График плотности вероятности показательного распределения

Функция распределения показательного распределения имеет вид

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

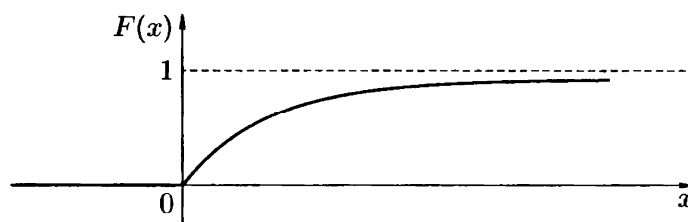


Рис.20. График функции распределения показательного распределения

Найдем математическое ожидание и дисперсию показательного распределения:

$$MX = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = \lim_{b \rightarrow \infty} \int_0^b x \cdot \lambda e^{-\lambda x} dx = [\text{интегрируем по частям}] =$$

$$= \lim_{b \rightarrow \infty} \left(-x \cdot e^{-\lambda x} \Big|_0^b - \frac{1}{\lambda} \cdot e^{-\lambda x} \Big|_0^b \right) = 0 - \frac{1}{\lambda} (0 - 1) = \frac{1}{\lambda}.$$

$$DX = \int_{-\infty}^{\infty} x^2 f(x) dx - (MX)^2 = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} =$$

$$= [\text{дважды интегрируем по частям}] =$$

$$= \lambda \left(\lim_{b \rightarrow \infty} \left(-\frac{x^2}{\lambda} e^{-\lambda x} + \frac{2}{\lambda} \left(-\frac{x}{\lambda} e^{-\lambda x} - \frac{1}{\lambda^2} e^{-\lambda x} \right) \right) \Big|_0^b \right) - \frac{1}{\lambda^2} =$$

$$= \lambda \left(0 + \frac{2}{\lambda} (0 + 0 - \frac{1}{\lambda^2} (0 - 1)) \right) - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Таким образом,

$$MX = \frac{1}{\lambda}, \quad DX = \frac{1}{\lambda^2}.$$

Пример 2.6. Случайная величина T – время работы прибора имеет показательное распределение. Найти вероятность того, что прибор проработает не менее 800 часов, если среднее время работы прибора 400 часов.

Решение. Имеем: $MT=400$, $\lambda=1/400$.

$$P\{T \geq 800\} = 1 - P\{T < 800\} = 1 - (1 - e^{-\frac{800}{400}}) = e^{-2} \approx 0.135.$$

Показательное распределение используется в приложениях теории вероятностей, особенно в теории массового обслуживания, в физике, в теории надежности. Оно используется для описания распределения случайной величины вида: длительность работы прибора до первого отказа, длительность времени обслуживания в системе массового обслуживания и т. д.

Нормальный закон распределения

Нормальный закон распределения (закон Гаусса) играет особо важную роль в теории вероятностей и чаще других применяется в решении практических задач. Его главная особенность в том, что он является предельным законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.

Непрерывная случайная величина X распределена по нормальному закону с параметрами a и $\sigma > 0$, если ее плотность распределения имеет вид:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in R. \quad (2.20)$$

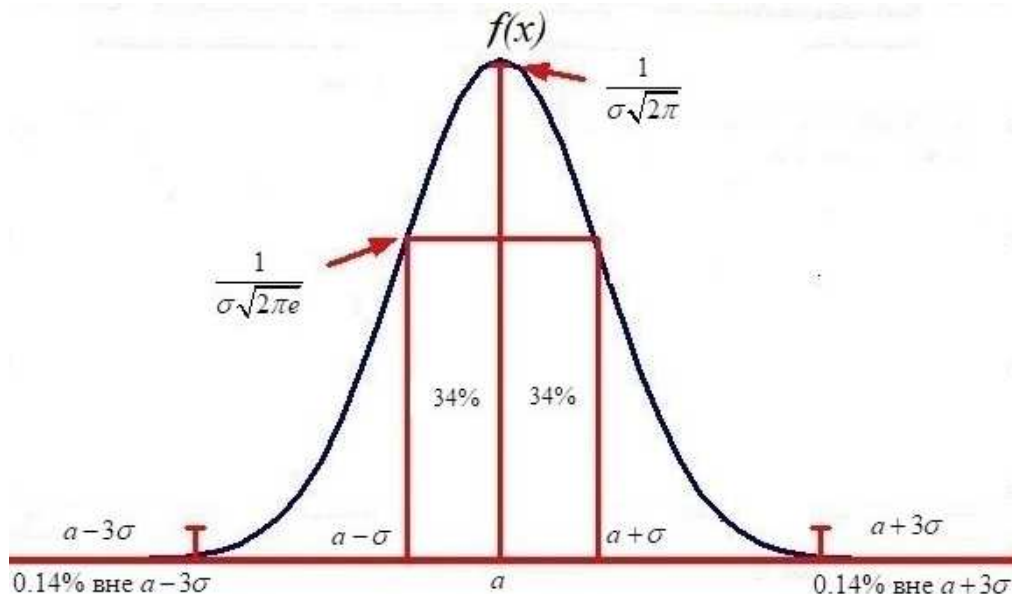


Рис.21. График плотности распределения вероятности нормального закона

Тот факт, что X имеет нормальное (или гауссовское) распределение с параметрами a и σ , сокращенно записывается так: $X \sim N(a, \sigma)$.

Функция распределения $F(x)$ имеет вид

$$F(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt.$$

Установим смысл параметров, a и σ нормального распределения. Для этого найдем математическое ожидание и дисперсию случайной величины $X \sim N(a, \sigma)$.

$$\begin{aligned} MX &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= [\text{подстановка } \frac{x-a}{\sigma\sqrt{2}} = t] = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t \sqrt{2} + a) e^{-t^2} \sigma \sqrt{2} dt = \\ &= \frac{\sigma \sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t e^{-t^2} dt + \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt = 0 + \frac{a}{\sqrt{\pi}} \cdot \sqrt{\pi} = a. \end{aligned}$$

$$\begin{aligned} DX &= \int_{-\infty}^{\infty} (x-a)^2 f(x) dx = \int_{-\infty}^{\infty} (x-a)^2 \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} 2\sigma^2 t^2 e^{-t^2} \sigma \sqrt{2} dt = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt = \end{aligned}$$

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

Таким образом, $MX = a$, $DX = \sigma^2$, σ – среднее квадратичное отклонение.

Если $a = 0$ и $\sigma = 1$, то нормальное распределение с такими параметрами называется *стандартным*. Плотность стандартной случайной величины имеет вид

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

Функция распределения $X \sim N\{0,1\}$ имеет вид

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (2.21)$$

и называется *функцией Лапласа*.

Она связана с нормированной функцией Лапласа

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \quad (2.22)$$

равенством

$$\Phi(x) = 0.5 + \Phi_0(x)$$

Можно показать, что для случайной величины $X \sim N(a, \sigma)$:

$$M_0 X = M_e = a, \quad A_s = \frac{\mu_3}{\sigma^3} = 0, \quad E_s = \frac{\mu_4}{\sigma^4} - 3 = 0$$

Как влияет изменение параметров a и σ на форму кривой Гаусса? Очевидно, что изменение a не изменяет форму нормальной кривой (графики функции $f(x)$ и $f(x-a)$ имеют одинаковую форму; график $f(x-a)$ получается из графика функции $f(x)$ путем сдвига последнего на a единиц вправо, если $a > 0$, и влево, если $a < 0$). С изменением σ максимальная ордината точки кривой изменяется. Так как площадь, ограниченная кривой распределения, равна единице при любом значении a , то с возрастанием σ кривая Гаусса становится более полой, растягивается вдоль оси Ox .

На рис. 22 изображены нормальные кривые при различных значениях σ ($\sigma_1 < \sigma < \sigma_2$) и некотором значении a (одинаковом для всех трех кривых).

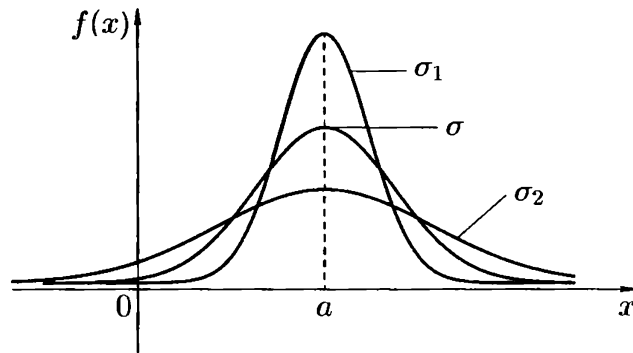


Рис. 22. Зависимость плотности распределения от σ

Нормальному закону подчиняются ошибки измерений, величины износа деталей в механизмах, рост человека, ошибки стрельбы, величина шума в радиоприемном устройстве, колебания курса акций и т. д.

Найдем вероятность попадания случайной величины $X \sim N(a, \sigma)$ на заданный участок (α, β) .

$$\begin{aligned}
 P\{\alpha < X < \beta\} &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = [\text{подстановка } \frac{x-a}{\sigma} = t] = \\
 &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} dt.
 \end{aligned}$$

Используя функцию Лапласа (2.22), получаем

$$P\{\alpha < X < \beta\} = \Phi_0\left(\frac{\beta-a}{\sigma}\right) - \Phi_0\left(\frac{\alpha-a}{\sigma}\right) \quad (2.23)$$

Равенство (2.23) можно переписать и так

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right) \quad (2.24)$$

На практике часто приходится вычислять вероятность попадания нормально распределенной случайной величины в интервал, симметричный относительно центра рассеяния a .

$$P\{|X - a| < l\} = \Phi_0\left(\frac{a+l-a}{\sigma}\right) - \Phi_0\left(\frac{a-l-a}{\sigma}\right) = 2\Phi_0\left(\frac{l}{\sigma}\right) \quad (2.25)$$

Полагая $l=3\sigma$, получим

$$P\{|X - a| < 3\sigma\} = 2\Phi_0(3)$$

По таблице значений для $\Phi_0(x)$ находим: $\Phi_0(3) = 0,49865$, т. е. отклонение случайной величины X от своего математического ожидания меньше, чем 3σ –

почти достоверное событие.

Вывод: практически достоверно, что случайная величина $X \sim N(a, \sigma)$ принимает свои значения в промежутке $(a-3\sigma, a+3\sigma)$. Это утверждение называется «правилом трех сигм».

Моменты нормального распределения

Выведем общие формулы для центральных моментов любого порядка. По определению:

$$\mu_s = \int_{-\infty}^{\infty} (x-m)^s f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-m)^s e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

Делая замену переменной

$$\frac{x-m}{\sigma\sqrt{2}} = t,$$

получим

$$\mu_s = \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^s e^{-t^2} dt. \quad (2.26)$$

Применим к выражению (2.26) формулу интегрирования по частям, заменив в нем t^s на $t^{s-1} \cdot t$. Будем иметь:

$$u = t^{s-1}; \quad dv = t \cdot e^{-t^2} dt.$$

$$du = (s-1) \cdot t^{s-2}; \quad v = \int t \cdot e^{-t^2} dt = \frac{1}{2} \int e^{-t^2} dt^2 = -\frac{1}{2} e^{-t^2}.$$

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

Имея в виду, что первый член внутри скобок равен нулю, получим:

$$\mu_s = \frac{(s-1)(\sigma\sqrt{2})^s}{2\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt \quad (2.27)$$

Из формулы (2.26) имеем следующее выражение для μ_{s-2} :

$$\mu_{s-2} = \frac{(\sigma\sqrt{2})^{s-2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt \quad (2.28)$$

Сравнивая правые части формул (2.27) и (2.28), видим, что они отличаются между собой только множителем $(s-1)\sigma^2$; следовательно,

$$\mu_s = (s-1)\sigma^2 \mu_{s-2} \quad (2.29)$$

Формула (2.29) представляет собой простое рекуррентное соотношение, позволяющее выразить моменты высших порядков через моменты низших

порядков. Пользуясь этой формулой и имея в виду, что $\mu_0 = 1$ и $\mu_1 = 0$, можно вычислить центральные моменты всех порядков. Так как $\mu_1 = 0$, то из формулы (2.29) следует, что все нечетные моменты нормального распределения равны нулю. Это, впрочем, непосредственно следует из симметричности нормального закона.

Для четных s из формулы (2.29) вытекают следующие выражения для последовательных моментов: $\mu_2 = \sigma^2$; $\mu_4 = 3\sigma^4$; $\mu_6 = 15\sigma^6$ и т.д.

Общая формула для момента s -го порядка при любом четном s имеет вид:

$$\mu_s = (s-1)!!\sigma^s,$$

где под символом $(s-1)!!$ понимается произведение всех нечетных чисел от 1 до $s-1$.

Поскольку все нечетные центральные моменты нормального распределения равны нулю, то асимметрия $A = \frac{\mu_3}{\sigma^3}$ нормального распределения будет равна нулю.

Это естественно, так как назначение асимметрии – характеризовать сравнительную скошенность данного закона по сравнению с нормальным.

Из выражения четвертого момента $\mu_4 = 3\sigma^4$ имеем:

$$E_s = \frac{\mu_4}{\sigma^4} - 3 = 0,$$

т.е. эксцесс нормального распределения равен нулю. Это также естественно, так как назначение эксцесса – характеризовать сравнительную крутость данного закона по сравнению с нормальным.

3. Системы случайных величин

Понятие о системе случайных величин и законе ее распределения

Совместное рассмотрение нескольких случайных величин приводит к системам случайных величин. Так, точка попадания снаряда характеризуется системой (X, Y) двух случайных величин: абсциссой X и ординатой Y ; успеваемость наудачу взятого абитуриента характеризуется системой n случайных величин (X_1, X_2, \dots, X_n) – оценками в его аттестате зрелости.

На многомерные случайные величины распространяются почти без изменений основные понятия и определения, относящиеся к одномерным случайным величинам. Ограничимся для простоты рассмотрением системы двух случайных величин; основные понятия обобщаются на случай большего числа компонент.

Упорядоченная пара (X, Y) двух случайных величин X и Y называется *двумерной случайной величиной* или *системой двух одномерных случайных величин* X и Y .

Полной характеристикой системы (X, Y) является ее *закон распределения вероятностей*, указывающий область возможных значений системы случайных величин и вероятности этих значений. Как и для отдельных случайных величин закон распределения системы может иметь разные формы (таблица, функция распределения, плотность, ...).

Зная закон распределения двумерной дискретной случайной величины, можно найти законы распределения каждой из компонент (обратное, вообще говоря, неверно).

Функция распределения двумерной случайной величины и ее свойства

Функцией распределения двумерной случайной величины (X, Y) называется функция $F(x, y)$, которая для любых действительных чисел x и y равна вероятности совместного выполнения двух событий $\{X < x\}$ и $\{Y < y\}$.

Таким образом, по определению

$$F(x, y) = P\{X < x, Y < y\}. \quad (3.1)$$

Геометрически функция $F(x, y)$ интерпретируется как вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной в точке (x, y) , лежащий левее и ниже ее (рис. 23).

Функция распределения двумерной дискретной случайной величины (X, Y) находится суммированием всех вероятностей p_{ij} , для которых $x_i < x, y_j < y$ т. е.

$$F(x, y) = \sum_{x_i < x} \sum_{y_j < y} p_{ij}.$$

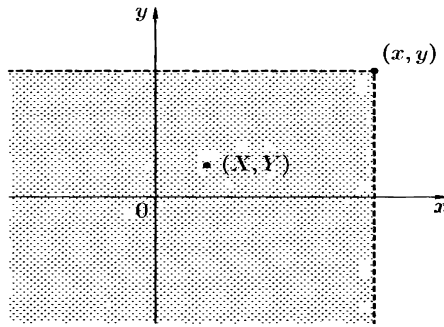


Рис. 23. Геометрическая интерпретация $F(x, y)$

Свойства функции распределения двумерной случайной величины:

1. Функция распределения $F(x, y)$ ограничена, т.е.

$$0 \leq F(x, y) \leq 1.$$

2. $F(x, y)$ не убывает по каждому из своих аргументов при фиксированном другом, т.е.

$$\begin{aligned} F(x_2, y) &\geq F(x_1, y) && \text{при } x_2 > x_1; \\ F(x, y_2) &\geq F(x, y_1) && \text{при } y_2 > y_1. \end{aligned}$$

3. Если хотя бы один из аргументов обращается в $-\infty$, то функция распределения $F(x, y)$ равна нулю, т.е.

$$F(x, -\infty) = F(-\infty, y) = (-\infty, -\infty) = 0.$$

4. Если оба аргумента обращаются в $+\infty$, то $F(x, y)$ равна 1, т.е.

$$F(+\infty, +\infty) = 1.$$

5. Если один из аргументов обращается в $+\infty$, функция распределения системы случайных величин становится функцией распределения случайной величины, соответствующей другому элементу, т.е.

$$F(x, +\infty) = F_1(x) = F_x(x), \quad F(+\infty, y) = F_2(y) = F_y(y).$$

6. $F(x, y)$ непрерывна слева по каждому из своих аргументов

$$\lim_{x \rightarrow x_0 - 0} F(x, y) = F(x_0, y), \quad \lim_{y \rightarrow y_0 - 0} F(x, y) = F(x, y_0).$$

Используя функцию распределения системы случайных величин X и Y , легко найти вероятность того, что в результате испытания случайная точка попадает в

полуполосу $x_1 \leq X < x_2$ и $Y < y$ (рис. 23, а) или в полуполосу $X < x$ и $y_1 \leq Y < y_2$ (рис.24, б).

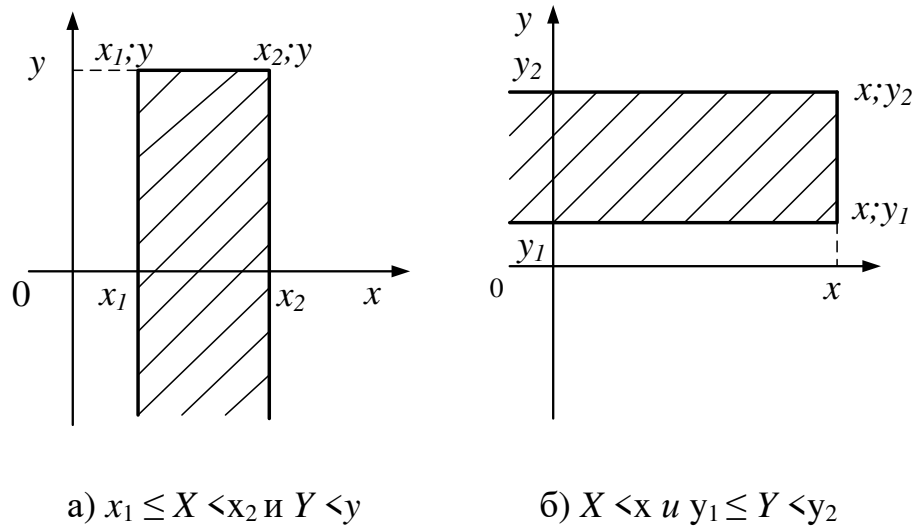


Рис. 24. Вероятность попадания в полуполосы

Вычитая из вероятности попадания случайной точки в квадрант с вершиной $(x_2; y)$ вероятность попадания точки в квадрант с вершиной $(x_1; y)$ (рис. 24, а), получим

$$P(x_1 \leq X < x_2, Y < y) = F(x_2, y) - F(x_1, y)$$

Аналогично имеем

$$P(X < x, y_1 \leq Y < y_2) = F(x, y_2) - F(x, y_1)$$

Таким образом, вероятность попадания случайной точки в полуполосу равна приращению функции распределения по одному из аргументов.

Рассмотрим прямоугольник $ABCD$ со сторонами, параллельными координатным осям (рис. 25). Пусть уравнения сторон таковы:

$$X=x_1, X=x_2, Y=y_1, Y=y_2$$

Найдем вероятность попадания случайной точки (X, Y) в этот прямоугольник.

Искомую вероятность можно найти так: из вероятности попадания случайной точки в полуполосу AB с вертикальной штриховкой (эта вероятность равна $F(x_2, y_2) - F(x_1, y_2)$) вычесть вероятность попадания точки в полуполосу CD с горизонтальной штриховкой (эта вероятность равна $F(x_2, y_1) - F(x_1, y_1)$)

Тогда вероятность попадания случайной точки (X, Y) в этот прямоугольник равна

$$P\{x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

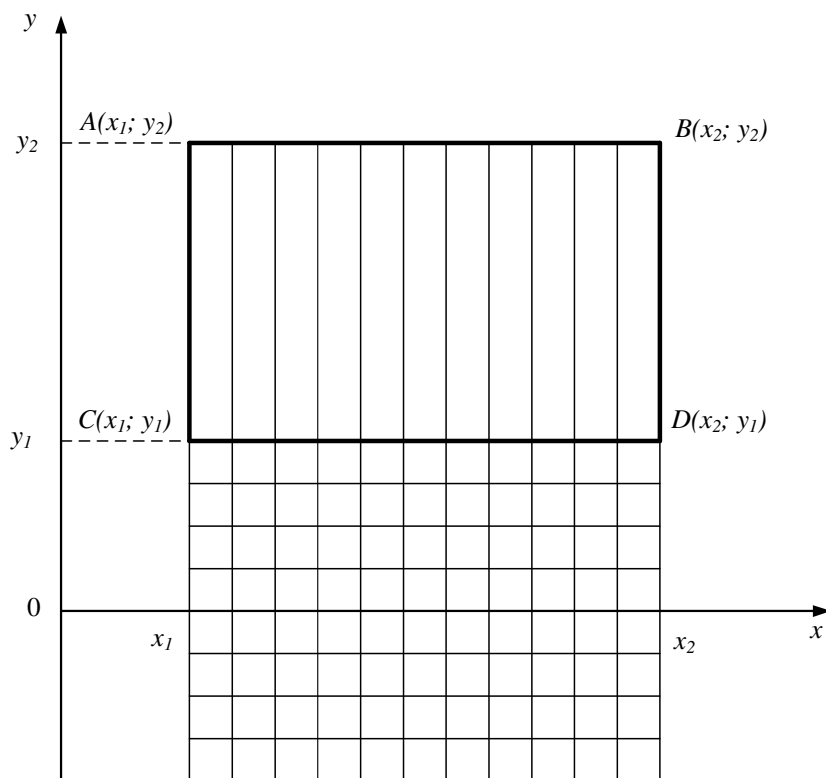


Рис. 25. Вероятность попадания случайной точки (X, Y) в прямоугольник

Плотность распределения вероятностей двумерной случайной величины и ее свойства

Плотностью распределения вероятностей (или совместной плотностью) непрерывной двумерной случайной величины (X, Y) называется вторая смешанная производная ее функции распределения

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = F''_{xy}(x, y), \quad (3.2)$$

Плотность распределения вероятностей непрерывной двумерной случайной величины (X, Y) есть предел отношения вероятности попадания случайной точки (X, Y) в элементарный прямоугольник со сторонами Δx и Δy , примыкающий к точке (x, y) , к площади этого прямоугольника, когда его размеры Δx и Δy стремятся к нулю (рис. 26).

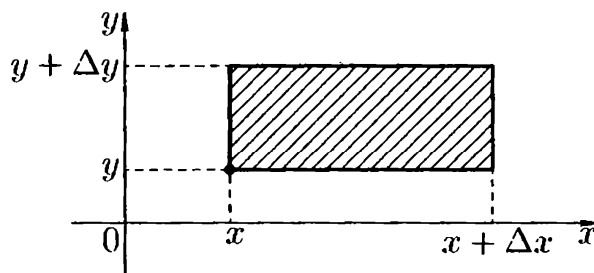


Рис. 26. Элементарный прямоугольник со сторонами Δx и Δy

Свойства плотности распределения двумерной случайной величины:

1. Плотность распределения двумерной случайной величины неотрицательна

$$f(x, y) \geq 0$$

2. Вероятность попадания случайной точки (X, Y) в область D равна двойному интегралу от плотности по области D

$$P\{(X, Y) \in D\} = \iint_D f(x, y) dx dy.$$

3. Функция распределения двумерной случайной величины может быть выражена через ее плотность распределения по формуле

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv.$$

4. Условие нормировки; двойной несобственный интеграл в бесконечных пределах от плотности вероятности двумерной с. в. равен единице

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$$

5. Плотности распределения одномерных составляющих X и Y могут быть найдены по формулам

$$\int_{-\infty}^{+\infty} f(x, y) dy = f_1(x) = f_X(x), \quad \int_{-\infty}^{+\infty} f(x, y) dx = f_2(y) = f_Y(y),$$

Условные законы распределения

Случайные величины X и Y называются *независимыми*, если независимыми являются события $\{X < x\}$ и $\{Y < y\}$ для любых действительных x и y . В противном случае случайные величины называются *зависимыми*.

Сформулируем условие независимости случайных величин.

Теорема 3.1. Для того, чтобы случайные величины X и Y были независимы, необходимо и достаточно, чтобы функция распределения системы (X, Y) была равна произведению функций распределения составляющих

$$F(x, y) = F_1(x) \cdot F_2(y).$$

Теорема 3.2. Необходимым и достаточным условием независимости двух дискретных случайных величин X и Y , образующих систему (X, Y) , является равенство

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\} \quad \text{для любых } i = 1, 2, \dots, n, j = 1, 2, \dots, m.$$

Теорема 3.3. Необходимым и достаточным условием независимости двух непрерывных случайных величин X и Y , образующих систему (X, Y) , является равенство

$$f(x, y) = f_1(x) \cdot f_2(y).$$

Если случайные величины X и Y , образующие систему (X, Y) , зависимы между собой, то для характеристики их зависимости вводится понятие условных законов распределения случайных величин.

Условным законом распределения одной из случайных величин, входящей в систему (X, Y) , называется ее закон распределения, найденный при условии, что другая случайная величина приняла определенное значение (или попала в какой-то интервал).

Пусть (X, Y) – непрерывная двумерная случайная величина с плотностью $f(x, y)$; $f_1(x)$ и $f_2(y)$ – плотности распределения соответственно случайной величины X и случайной величины Y .

Плотность вероятности условного распределения (условная плотность) случайной величины Y при условии $X = x$ определяется равенством

$$f(y/x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy}, \quad \text{где } f_1(x) \neq 0, \quad (3.3)$$

Аналогично определяется условная плотность распределения случайной величины X при условии $Y = y$

$$f(x/y) = \frac{f(x, y)}{f_2(y)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dx}, \quad \text{где } f_2(y) \neq 0, \quad (3.4)$$

Используя соотношения (3.3) и (3.4), можно записать

$$f(x, y) = f_1(x) \cdot f(y/x) = f_2(y) \cdot f(x/y), \quad (3.5)$$

Равенство (3.5) называют теоремой умножения плотностей распределений (она аналогична теореме умножения вероятностей для событий).

Пример. 3.1. Задана таблица распределения дискретной двумерной случайной величины (X, Y) :

$X \backslash Y$	$y_1=1$	$y_2=2$	$y_3=3$
$x_1=0.1$	$p_{11}=0.12$	$p_{12}=0.08$	$p_{13}=0.4$
$x_2=0.2$	$p_{21}=0.16$	$p_{22}=0.1$	$p_{23}=0.14$

Найти: а) безусловные законы распределения случайных величин X и Y ;

б) условный закон распределения случайной величины X при $Y = 2$.

Решение.

а) Так как $P_{x_i} = P\{X = x_i\} = \sum_{j=1}^m p_{ij}$, $P_{y_j} = P\{Y = y_j\} = \sum_{i=1}^n p_{ij}$, то

X	0.1	0.2
p	0.60	0.40

Y	1	2	3
p	0.28	0.18	0.54

б) Условный закон распределения случайной величины X при условии $Y = y_j$:

$$P\{X = x_i \setminus Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m$$

$$\text{или коротко: } p(x_i \setminus y_j) = \frac{p_{ij}}{p_{y_j}}, \quad \text{где } p_{y_j} = \sum_{i=1}^n p_{ij}$$

В нашем случае условный закон распределения случайной величины X при $Y = 2$ таков

$$P\{X = 0.1 \setminus Y = 2\} = \frac{0.08}{0.18} = \frac{4}{9} \approx 0.45, \quad P\{X = 0.2 \setminus Y = 2\} = \frac{0.1}{0.18} = \frac{5}{9} \approx 0.55.$$

Т.е.

X	0.1	0.2
$P_{Y=2}$	0.45	0.55

Очевидно несовпадение условного и безусловного законов распределения случайной величины X . Следовательно, случайные величины X и Y зависимы.

Числовые характеристики двумерной случайной величины

Для системы случайных величин также вводятся числовые характеристики, подобные тем, какие были для одной случайной величины. На практике чаще всего используются моменты I и II порядков: математическое ожидание, дисперсия и корреляционный момент. Математическое ожидание и дисперсия двумерной случайной величины служат соответственно средним значением и мерой рассеяния. Корреляционный момент выражает меру взаимного влияния случайных величин, входящих в систему (X, Y) .

Математическое ожидание

Для дискретных случайных величин

$$MX = m_x = \sum_{i=1}^n \sum_{j=1}^m x_i p_{ij}, \quad MY = m_y = \sum_{i=1}^n \sum_{j=1}^m y_j p_{ij}. \quad (3.6)$$

Для непрерывных случайных величин

$$MX = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dx dy, \quad MY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y) dx dy. \quad (3.7)$$

Дисперсия

Для дискретной двумерной случайной величины

$$DX = \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)^2 p_{ij}, \quad DY = \sum_{i=1}^n \sum_{j=1}^m (y_j - m_y)^2 p_{ij}. \quad (3.8)$$

Для непрерывной двумерной случайной величины

$$DX = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)^2 f(x, y) dx dy, \quad DY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - m_y)^2 f(x, y) dx dy. \quad (3.9)$$

Математические ожидания m_x и m_y являются частными случаями начального момента α_{ks} порядка $k+s$ системы (X, Y) , определяемого равенством

$$\alpha_{k,s} = M(X^k Y^s) \\ m_x = M(X^1 Y^0) = \alpha_{1,0}, \quad m_y = M(X^0 Y^1) = \alpha_{0,1}$$

Дисперсии DX и DY являются частными случаями центрального момента μ_{ks} порядка $k+s$ системы (X, Y) , определяемого равенством

$$\mu_{k,s} = M((X - m_x)^k (Y - m_y)^s) \\ DX = M(X - m_x)^2 = \mu_{2,0}, \quad DY = M(Y - m_y)^2 = \mu_{0,2}$$

Ковариация

Корреляционным моментом (или ковариацией) двух случайных величин X и Y называется математическое ожидание произведения отклонений этих случайных величин от их математического ожидания

$$K_{XY} = \text{cov}(X, Y) = M[(X - m_x)(Y - m_y)] = M[XY - X m_y - m_x Y + m_x m_y] = m_{xy} - m_x m_y$$

Для дискретной двумерной случайной величины

$$K_{XY} = \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)(y_j - m_y) p_{ij}. \quad (3.10)$$

Для непрерывной двумерной случайной величины

$$K_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y) dx dy. \quad (3.11)$$

Ковариацию часто удобно вычислять по формуле

$$K_{XY} = M(XY) - M(X) \cdot M(Y). \quad (3.12)$$

В качестве числовой характеристики зависимости случайных величин X и Y обычно используют безразмерную величину – коэффициент корреляции

$$r_{XY} = \frac{K_{XY}}{\sigma_x \sigma_y} = \frac{\text{cov}(X, Y)}{\sqrt{DX} \sqrt{DY}}. \quad (3.13)$$

Свойства коэффициента корреляции:

1. Коэффициент корреляции по абсолютной величине не превосходит 1, т.е.

$$|r_{XY}| \leq 1 \quad \text{или} \quad -1 \leq r_{XY} \leq 1.$$

2. Если X и Y независимы, то $r_{XY} = 0$. В этом случае X и Y называются некоррелированными.

3. Если случайные величины X и Y связаны линейной зависимостью, т. е.

$Y = aX + b$, $a \neq 0$, то $|r_{XY}| = 1$, причем $r_{XY} = 1$ при $a > 0$ (возрастанию X соответствует возрастание в среднем Y), $r_{XY} = -1$ при $a < 0$ (возрастанию X соответствует убывание в среднем Y).

4. Если $|r_{XY}| = 1$, то случайные величины X и Y связаны линейной функциональной зависимостью.

Пример 3.2. Закон распределения дискретной двумерной случайной величин задан таблицей:

XY	-1	0	1
0	0.15	0.40	0.05
1	0.2	0.1	0.1

Найти коэффициент корреляции r_{XY}

Решение

Находим законы распределения составляющих X и Y

X	0	1
p	0.6	0.4

Y	-1	0	1
p	0.35	0.5	0.15

Находим математическое ожидание составляющих

$$m_x = 0 \cdot 0.6 + 1 \cdot 0.4 = 0.4 \quad m_y = -1 \cdot 0.35 + 0 \cdot 0.5 + 1 \cdot 0.15 = -0.2$$

Находим дисперсии составляющих

$$DX = MX^2 - (MX)^2 = (0^2 \cdot 0.6 + 1^2 \cdot 0.4) - (0.4)^2 = 0.24$$

$$DY = MY^2 - (MY)^2 = ((-1)^2 \cdot 0.35 + 0^2 \cdot 0.5 + 1^2 \cdot 0.15) - (-0.2)^2 = 0.46$$

$$\text{Тогда } \sigma_x = \sqrt{0.24} \approx 0.49 \quad \sigma_y = \sqrt{0.46} \approx 0.68$$

Находим $M(XY) = \sum_{i=1}^2 \sum_{j=1}^3 x_i y_j p_{ij}$

$$M(XY) = 0 \cdot (-1) \cdot 0.15 + 0 \cdot 0 \cdot 0.40 + 0 \cdot 1 \cdot 0.05 + 1 \cdot (-1) \cdot 0.20 + 1 \cdot 0 \cdot 0.1 + 1 \cdot 1 \cdot 0.10 = -0.1$$

Находим корреляционный момент, используя формулу (3.12)

$$K_{XY} = -0.1 - 0.4 \cdot (-0.2) = -0.02$$

Тогда коэффициент корреляции равен $r_{XY} = \frac{K_{XY}}{\sigma_x \sigma_y} = \frac{-0.02}{0.49 \cdot 0.68} \approx -0.06$

Двумерное нормальное распределение

В теории вероятностей и её приложениях большую роль играет двумерное нормальное распределение. Плотность двумерной нормальной случайной величины (X, Y) имеет вид

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{XY}^2}} \exp \left\{ -\frac{1}{2(1-r_{XY}^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r_{XY}(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right] \right\} \quad (3.14)$$

Здесь m_x, m_y – математические ожидания величин X и Y ; σ_x, σ_y – средние квадратичные отклонения величин X и Y ; r_{XY} – коэффициент корреляции величин X и Y .

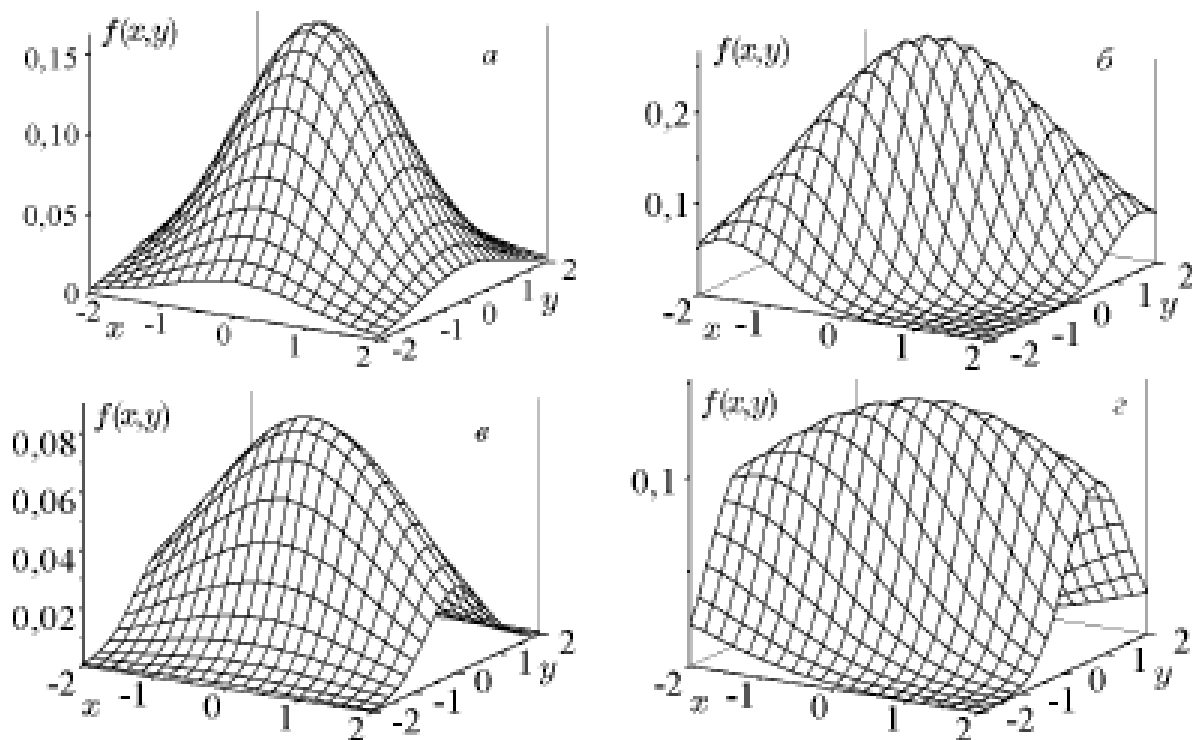


Рис. 27. Графики плотностей 2-мерного нормального распределения при $m_x = m_y = 0$

$\sigma_x = \sigma_y = 1, r_{XY} = 0$ (а); $\sigma_x = \sigma_y = 1, r_{XY} = 0.8$ (б);
 $\sigma_x = 2, \sigma_y = 1, r_{XY} = 0$ (в); $\sigma_x = 2, \sigma_y = 1, r_{XY} = 0.8$ (г)

Предположим, что случайные величины X и Y не коррелированы, то есть $r_{XY}=0$. Тогда имеем:

$$\begin{aligned}
 f(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{(x-m_x)^2}{\sigma_x^2} - \frac{(y-m_y)^2}{\sigma_y^2}\right\} = \\
 &= \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left\{-\frac{(x-m_x)^2}{\sigma_x^2}\right\} \cdot \frac{1}{\sigma_y\sqrt{2\pi}} \exp\left\{-\frac{(y-m_y)^2}{\sigma_y^2}\right\} = f_1(x) \cdot f_2(y)
 \end{aligned}
 \tag{3.15}$$

Получили, что плотность распределения системы двух случайных величин (X , Y) равна произведению плотностей распределения компонент X и Y , а это значит, что X и Y – независимые случайные величины.

Таким образом, доказана следующая теорема: из некоррелированности нормально распределенных случайных величин следует их независимость.

Поскольку из независимости любых случайных величин следует их некоррелированность, то можно сделать вывод, что термины «некоррелированные» и «независимые» величины для случая нормального распределения эквивалентны.

В общем случае, для любого закона, кроме нормального, из некоррелированности не следует независимость.

Приведём формулы для вероятности попадания нормально распределённой двумерной случайной величины в различные области на плоскости.

Пусть случайный вектор (X, Y) , компоненты которого независимы, распределён по нормальному закону (3.15). Тогда вероятность попадания случайной точки (X, Y) в прямоугольник R (рис. 28), стороны которого параллельны координатным осям, равна

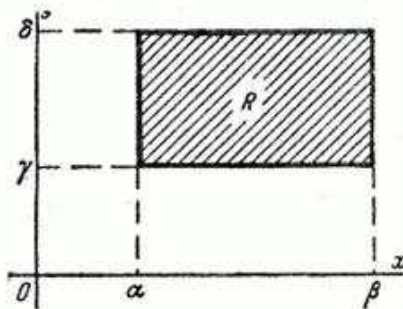


Рис. 28. Прямоугольник R , стороны которого параллельны координатным осям

$$P((X, Y) \in R) = \int_{\alpha}^{\beta} \int_{\gamma}^{\delta} f(x, y) dx dy = \int_{\alpha}^{\beta} \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left\{-\frac{(x-m_x)^2}{\sigma_x^2}\right\} dx \int_{\gamma}^{\delta} \frac{1}{\sigma_y\sqrt{2\pi}} \exp\left\{-\frac{(y-m_y)^2}{\sigma_y^2}\right\} dy
 \tag{3.16}$$

Откуда, применяя формулу для вероятности попадания на участок, находим:

$$P((X, Y) \in R) = \left[\Phi_0\left(\frac{\beta - m_x}{\sigma_x}\right) - \Phi_0\left(\frac{\alpha - m_x}{\sigma_x}\right) \right] \left[\Phi_0\left(\frac{\delta - m_y}{\sigma_y}\right) - \Phi_0\left(\frac{\gamma - m_y}{\sigma_y}\right) \right], \quad (3.17)$$

где $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{u^2}{2}\right) du$ – нормированная функция Лапласа.

Общий случай n -мерного нормального распределения

Плотность нормального распределения системы n случайных величин имеет вид:

$$f(x_1, x_2, \dots, x_n) = \frac{\sqrt{\det C}}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n C_{ij} (x_i - m_{X_i})(x_j - m_{X_j})\right\}, \quad (3.20)$$

где $\det C$ – определитель матрицы C , обратной к ковариационной матрице;

m_{X_i} – математическое ожидание случайной величины X_i – i -той компоненты n -мерного нормального случайного вектора;

m_{X_j} – математическое ожидание случайной величины X_j – j -той компоненты n -мерного нормального случайного вектора.

Из общего выражения вытекают все формы нормального закона для любого числа измерений и для любых видов зависимости между случайными величинами.

В частности, при $n = 2$ ковариационная матрица имеет вид:

$$k = \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y r_{XY} \\ \sigma_x \sigma_y r_{XY} & \sigma_y^2 \end{pmatrix}. \quad (3.21)$$

Матрица C , обратная к ковариационной матрице k , имеет вид:

$$C = \begin{pmatrix} \frac{1}{\sigma_x^2(1-r_{XY}^2)} & -\frac{r_{XY}}{\sigma_x \sigma_y(1-r_{XY}^2)} \\ -\frac{r_{XY}}{\sigma_x \sigma_y(1-r_{XY}^2)} & \frac{1}{\sigma_y^2(1-r_{XY}^2)} \end{pmatrix}. \quad (3.22)$$

Ее определитель $\det C = \frac{1}{\sigma_x^2 \sigma_y^2 (1-r_{XY}^2)}$

Подставляя $\det C$ и элементы матрицы C в общую формулу (3.20), получаем формулу для нормального распределения на плоскости (3.14).

Если случайные величины X_1, X_2, \dots, X_n независимы, то плотность распределения системы X_1, X_2, \dots, X_n равна

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_1 \sigma_2 \dots \sigma_n} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - m_{X_i})^2}{\sigma_i^2}\right\}. \quad (3.23)$$

При $n = 2$ эта формула принимает вид (3.15)

Распределение функций нормальных случайных величин

Часто возникают задачи, в которых по известному закону распределения (или числовым характеристикам) одной (или нескольких) случайной величины требуется определить распределение другой (или нескольких) случайной величины, функционально связанной с первой.

Рассмотрим распределение некоторых случайных величин, представляющих функции нормальных величин, используемые в математической статистике.

χ^2 -распределение (распределение Пирсона)

Пусть z_1, z_2, \dots, z_n – совместно независимые стандартные нормальные случайные величины, то есть: $z_i \sim N(0,1)$. Тогда случайная величина $X = \sum_{i=1}^n z_i^2 \sim \chi_n^2$ имеет распределение хи-квадрат с n степенями свободы. χ_n^2 – распределение впервые было получено в 1876 г. Гельмертом. Затем в 1900 г. снова было открыто и изучено Пирсоном.

Найдем функцию распределения случайной величины $X = \sum_{i=1}^n z_i^2 \sim \chi_n^2$

Пусть для $x > 0$ $F_n(x) = P(\chi_n^2 < x)$

Так как величины X_i независимы и каждая из них по условию имеет плотность $\frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}$, то совместная плотность величин X_1, X_2, \dots, X_n выразится произведением $(\frac{1}{\sqrt{2\pi}})^n e^{-\frac{\sum_{i=1}^n x_i^2}{2}}$. Поэтому для функции распределения можно записать:

$$F_n(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int \dots \int_{\sum_{i=1}^n x_i^2 \leq x} e^{-\frac{(x_1^2 + x_2^2 + \dots + x_n^2)}{2}} dx_1 \dots dx_n$$

Это n -мерный интеграл, распространённый на область $g_{n,x}$, определяемую неравенством $\sum_{i=1}^n x_i^2 \leq x$. Он выражает вероятность точки $Q(x_1, x_2, \dots, x_n)$ попасть в область $g_{n,x}$. Эта область – множество точек внутри и на поверхности сферы n -мерного пространства радиуса \sqrt{x} с центром в начале координат.

Найдем производную от функции $F_n(x)$.

$$F_n(x+h) - F_n(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int \dots \int_{x < \sum_{i=1}^n x_i^2 \leq x+h} e^{-\frac{(x_1^2 + x_2^2 + \dots + x_n^2)}{2}} dx_1 \dots dx_n, \quad (3.24)$$

где область интегрирования ограничена поверхностями двух концентрических сфер $g_{n,x}$ и $g_{n,x+h}$.

Применяя теорему (теорема Лагранжа) о среднем к (3.24), получим

$$F_n(x+h) - F_n(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{(x+\theta h)}{2}} \int \dots \int_{x < \sum_{i=1}^n X_i^2 \leq x+h} dx_1 \dots dx_n$$

Здесь $e^{-\frac{(x+\theta h)}{2}}$ – некоторое среднее значение подынтегральной функции в области интегрирования ($0 < \theta < 1$).

$$\text{Обозначим } S_n(x) = \int \dots \int_{\sum_{i=1}^n X_i^2 \leq x} dx_1 \dots dx_n$$

Тогда

$$F_n(x+h) - F_n(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{(x+\theta h)}{2}} [S_n(x+h) - S_n(x)] \quad (3.25)$$

$S_n(x)$ – объем сферы $g_{n,x}$ радиуса \sqrt{x} , он пропорционален n -ой степени радиуса.

Сделав замену переменных $X_i = Y_i \sqrt{x}$, $dX_i = \sqrt{x} dY_i$, получим:

$$S_n(x) = (\sqrt{x})^n \int \dots \int_{\sum_{i=1}^n Y_i^2 \leq 1} dy_1 \dots dy_n = C_{n,1} x^{\frac{n}{2}} \quad (3.26)$$

где $C_{n,1}$ – объем единичной сферы ($R=1$) n -мерного пространства.

Из (3.25) и (3.26) имеем:

$$\frac{F_n(x+h) - F_n(x)}{h} = (2\pi)^{-\frac{n}{2}} C_{n,1} e^{-\frac{(x+\theta h)}{2}} \frac{(x+h)^{\frac{n}{2}} - x^{\frac{n}{2}}}{h}$$

Переходя к пределу при $h \rightarrow 0$ и вычисляя его по правилу Лопиталья, получим:

$$f_n(x) = Ax^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \text{ где } A = (2\pi)^{-\frac{n}{2}} C_{n,1} \frac{n}{2}$$

Постоянную A находим из условия нормировки

$$\int_0^{\infty} f_n(x) dx = \int_0^{\infty} Ax^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = 1$$

Известно, что $\Gamma(p) = \int_0^{\infty} t^{p-1} e^{-t} dt$ – гамма-функция Эйлера ($\Gamma(p) = (p-1)!$ для

целых положительных p). Тогда с учетом этого будем иметь: $A = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})}$

Подставляя значение A в выражение для плотности $f_n(x)$ получим, что плотность распределения χ_n^2 имеет вид (рис. 28):

$$f_{\chi_n}(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & \text{при } x > 0 \\ 0, & \text{при } x \leq 0 \end{cases},$$

С возрастанием числа степеней свободы n распределение χ_n^2 приближается к нормальному закону распределения (при $n > 30$ распределение χ_n^2 практически не отличается от нормального).

Плотность вероятности случайной величины χ_n^2 зависит только от числа n – числа слагаемых (числа степеней свободы).

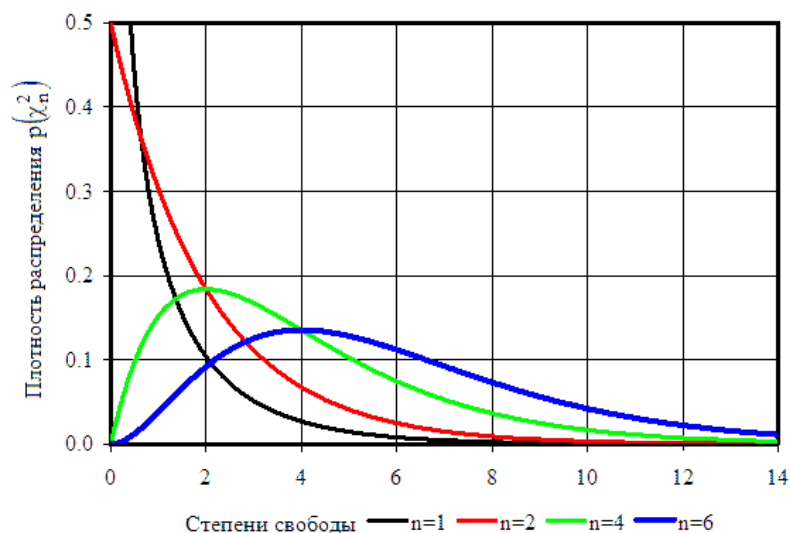


Рис. 29. Плотность распределения χ_n^2

$$M_{\chi_n^2} = n, \quad D_{\chi_n^2} = 2n.$$

На практике, как правило, используют не плотность вероятности распределения χ_n^2 , а его квантили.

Квантилью распределения χ_n^2 , отвечающей уровню значимости α , называется такое значение $\chi_n^2 = \chi_{\alpha,n}^2$ при котором

$$P\{\chi_n^2 > \chi_{\alpha,n}^2\} = \int_{\chi_{\alpha,n}^2}^{\infty} f_{\chi_n^2}(x) dx = \alpha.$$

С геометрической точки зрения нахождение квантили $\chi_{\alpha,n}^2$ заключается в выборе такого значения $\chi_n^2 = \chi_{\alpha,n}^2$, чтобы площадь заштрихованной на рис. 30 фигуры была равна α .

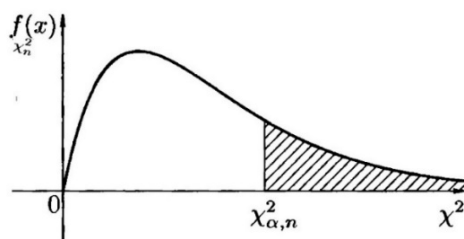


Рис. 30. Квантиль $\chi_{\alpha,n}^2$ распределения χ_n^2

Значения квантилей приводятся в соответствующих таблицах.

Распределение Стьюдента

Если ξ и η – независимые случайные величины, причем $\xi \sim N(0,1)$, а η имеет χ^2 - распределение соответственно с n степенями свободы, то случайная величина $T = \frac{\xi\sqrt{n}}{\sqrt{\eta}}$ подчиняется распределению Стьюдента (или t -распределению) с n степенями свободы.

Плотность вероятности распределения Стьюдента имеет вид (рис. 31):

$$f_{T_n}(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}; \quad t \in (-\infty; \infty).$$

График функции плотности распределения Стьюдента симметричен, а его форма напоминает форму колокола, как у стандартного нормального распределения, но он ниже и шире.

Следующие графики отражают плотность распределения Стьюдента при увеличении числа степеней свободы n . По мере возрастания n , кривая функции плотности все больше напоминает стандартное нормальное распределение и уже при $n > 30$ почти совпадает с нормальным распределением.

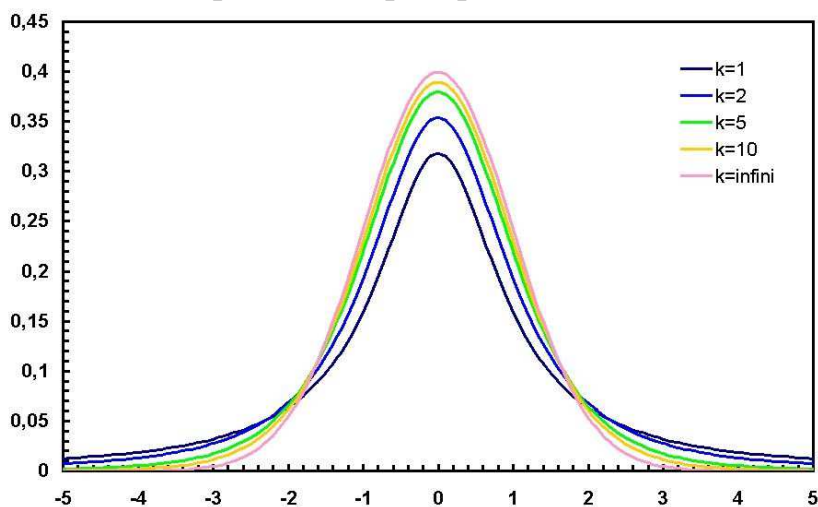


Рис. 31. Плотность t -распределения

$$MT_n = 0, \quad DT_n = \frac{n}{n-2}, \quad n > 2.$$

На практике используют квантили t -распределения: такое значение $t = t_{\frac{\alpha}{2}, n}$, что

$$P\{|t| > t_{\frac{\alpha}{2}, n}\} = 2 \int_{t_{\frac{\alpha}{2}, n}}^{\infty} f_{T_n}(t) dt = \alpha.$$

С геометрической точки зрения нахождение квантилей заключается в выборе

такого значения $t = t_{\frac{\alpha}{2}, n}$, чтобы площадь заштрихованной на рис. 32 фигуры была равна α .

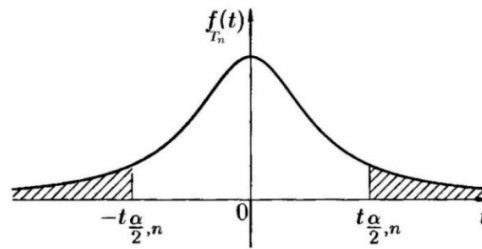


Рис. 32. Квантили $\pm t_{\frac{\alpha}{2}, n}$ t -распределения

Распределение Фишера-Снедекора

Пусть Y_1, Y_2 – две независимые случайные величины, имеющие распределение хи-квадрат: $Y_i \sim \chi^2(n_i)$, где $n_i \in N, i=1,2$. Тогда распределение

случайной величины $F = \frac{\frac{Y_1}{n_1}}{\frac{Y_2}{n_2}}$ называется распределением Фишера-Снедекора

(или F -распределением) со степенями свободы n_1 и n_2 (рис. 33).

При $n \rightarrow \infty$ F -распределение стремится к нормальному закону.

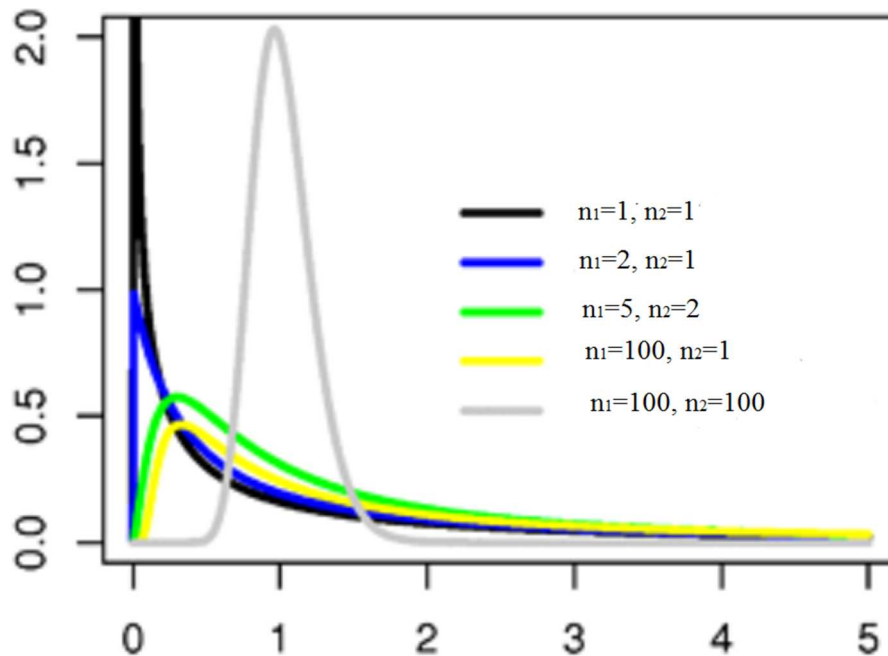


Рис. 33. Плотность F -распределения

$$MF = \frac{n}{n-2}, \quad n > 2 \quad DF = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \quad n > 4.$$

На практике обычно используют квантили F -распределения: такое значение $F = F_{\alpha, m, n}$, что

$$P\{F > F_{\alpha, m, n}\} = \int_{F_{\alpha, m, n}}^{\infty} f(F) dF = \alpha.$$

С геометрической точки зрения нахождение квантилей заключается в выборе такого значения $F = F_{\alpha, m, n}$, чтобы площадь заштрихованной на рис. 34 фигуры была равна α .

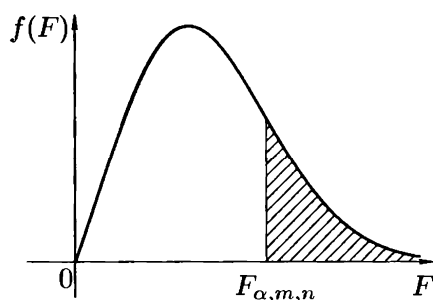


Рис. 34. Квантиль $F_{\alpha, m, n}$ F -распределения

4. Предельные теоремы теории вероятностей

Рассмотрим ряд утверждений и теорем из большой группы так называемых предельных теорем теории вероятностей, устанавливающих связь между теоретическими и экспериментальными характеристиками случайных величин при большом числе испытаний над ними. Они составляют основу математической статистики. Предельные теоремы условно делят на две группы. Первая группа теорем, называемая *законом больших чисел*, устанавливает устойчивость средних значений: при большом числе испытаний их средний результат перестает быть случайным и может быть предсказан с достаточной точностью. Вторая группа теорем, называемая *центральной предельной теоремой*, устанавливает условия, при которых закон распределения суммы большого числа случайных величин неограниченно приближается к нормальному

Неравенство Чебышева

Теорема 4.1. Если случайная величина X имеет математическое $MX=a$ и дисперсию DX , то для любого $\varepsilon > 0$ справедливо неравенство Чебышева

$$P\{|X - MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2} \quad (4.1)$$

Доказательство. Докажем неравенство (4.1) для непрерывной случайной величины X с плотностью $f(x)$.

$$\begin{aligned} P\{|X - a| \geq \varepsilon\} &= \int_{-\infty}^{a-\varepsilon} f(x)dx + \int_{a+\varepsilon}^{+\infty} f(x)dx = \int_{|x-a| \geq \varepsilon} f(x)dx = \\ &= \int_{|x-a| \geq \varepsilon} 1 \cdot f(x)dx \leq \int_{|x-a| \geq \varepsilon} \frac{(x-a)^2}{\varepsilon^2} f(x)dx, \end{aligned}$$

Т.к. область интегрирования $|x-a| \geq \varepsilon$ можно записать в виде $(x-a)^2 \geq \varepsilon^2$, откуда следует $1 \leq \frac{(x-a)^2}{\varepsilon^2}$. Имеем

$$P\{|X - a| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \int_{|x-a| \geq \varepsilon} (x-a)^2 f(x)dx \leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} (x-a)^2 f(x)dx,$$

т.к. интеграл от неотрицательной функции при расширении области интегрирования может только возрасти. Таким образом,

$$P\{|X - a| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} ((x-a)^2 f(x)dx = \frac{1}{\varepsilon^2} DX$$

Отметим, что неравенство Чебышева можно записать в другой форме.

$$P\{|X - MX| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2} \quad (4.2)$$

В форме (4.2) оно устанавливает нижнюю границу вероятности события, а в форме (4.1) – верхнюю.

Для случайной величины $X = m$, имеющей *биномиальное распределение* с математическим ожиданием $MX = a = np$ и дисперсией $DX = npq$, неравенство Чебышева принимает вид

$$P\{|m - np| < \varepsilon\} \geq 1 - \frac{npq}{\varepsilon^2} \quad (4.3)$$

для частоты $\frac{m}{n}$ события в n независимых испытаниях, в которых оно может произойти с вероятностью $p = M\left(\frac{m}{n}\right) = a$, дисперсия которых $D\left(\frac{m}{n}\right) = \frac{pq}{n}$, неравенство Чебышева имеет вид

$$P\left\{\left|\frac{m}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{pq}{n\varepsilon^2} \quad (4.4)$$

Пример 4.1. Оценить с помощью неравенства Чебышева вероятность того, что отклонение случайной величины X от своего математического ожидания будет меньше $3\sigma_x$.

Решение. Полагая в формуле (4.2) $\varepsilon = 3\sigma_x$, получаем

$$P\{|X - MX| < 3\sigma_x\} \geq 1 - \frac{\sigma_x^2}{(3\sigma_x)^2} = 1 - \frac{1}{9} \approx 0.8889.$$

Эта оценка, как известно, называется *правилом трех сигм*. Для случайной величины $X \approx N(\alpha, \sigma)$ эта вероятность равна 0.9973.

Закон больших чисел в форме Чебышева

Основное утверждение закона больших чисел содержится в теореме Чебышева (1886 г.). В ней и других теоремах закона больших чисел используется понятие «сходимости случайных величин по вероятности».

Случайные величины $X_1, X_2, \dots, X_n, \dots$ *сходятся по вероятности* к величине A (случайной или неслучайной), если для любого $\varepsilon > 0$ вероятность события $\{|X_n - A| < \varepsilon\}$ при $n \rightarrow \infty$ стремится к единице, т. е.

$$\lim_{n \rightarrow \infty} P\{|X_n - A| < \varepsilon\} = 1 \quad \text{или} \quad P\{|X_n - A| < \varepsilon\} \rightarrow 1$$

Сходимость по вероятности символически записывают так: $X_n \xrightarrow[n \rightarrow \infty]{p} A$.

Теорема 4.2.

Если случайные величины $X_1, X_2, \dots, X_n, \dots$ независимы и существует такое число $C > 0$, что $DX_i \leq C, i = 1, 2, \dots$, то для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n MX_i\right| < \varepsilon\right\} = 1,$$

т. е. среднее арифметическое этих случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий.

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p} \frac{1}{n} \sum_{i=1}^n MX_i$$

Доказательство.

Так как $DX_i \leq C, i = 1, 2, \dots$, то

$$\begin{aligned} D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) &= \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n DX_i = \frac{1}{n^2} (DX_1 + DX_2 + \dots + DX_n) \leq \\ &\leq \frac{1}{n^2} (C + C + \dots + C) = \frac{1}{n^2} n \cdot C = \frac{C}{n}. \end{aligned}$$

Тогда, применяя к случайной величине $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ неравенство Чебышева (4.2), имеем

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n MX_i\right| < \varepsilon\right\} \geq 1 - \frac{D\left(\frac{1}{n} \sum_{i=1}^n X_i\right)}{\varepsilon^2} \geq 1 - \frac{C}{n\varepsilon^2}$$

Переходя к пределу при $n \rightarrow \infty$ и учитывая, что вероятность любого события не превышает 1, получаем

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n MX_i\right| < \varepsilon\right\} = 1.$$

Следствие. Если случайные величины $X_1, X_2, \dots, X_n, \dots$ независимы и одинаково распределены, $MX_i = a, DX_i = \sigma^2$, то для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - a\right| < \varepsilon\right\} = 1,$$

т. е. среднее арифметическое случайной величины сходится по вероятности к математическому ожиданию a :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p} a.$$

На теореме Чебышева основан широко применяемый в статистике *выборочный метод*, суть которого в том, что о качестве большого количества однородного материала можно судить по небольшой его пробе.

Теорема Чебышева подтверждает связь между случайностью и необходимостью: среднее значение *случайной величины* практически не отличается от *неслучайной величины* (среднее значение математического ожидания).

Пример 4.2. Глубина моря измеряется прибором, не имеющим систематической ошибки. Среднее квадратическое отклонение измерений не

превосходит 15 м. Сколько нужно сделать независимых измерений, чтобы с вероятностью, не меньшей 0,9, можно было утверждать, что среднее арифметическое этих измерений отличается от a (глубины моря) по модулю меньше, чем на 5 м?

Решение. Обозначим через X_i результаты n независимых измерений глубины моря. Нужно найти число n , которое удовлетворяет неравенству

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n MX_i\right| < \varepsilon\right\} = 1 - \frac{C}{n\varepsilon^2},$$

где $MX_i = a$, что означает отсутствие при измерениях систематической ошибки (т.е. измерения производятся с одинаковой точностью).

По условию $\varepsilon = 5$, $C = 225$ (дисперсия), $\sigma = \sqrt{C} = \sqrt{225} = 15$ м. Тогда

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i - a\right| < 5\right\} = 1 - \frac{225}{25n} \geq 0.9,$$

т.е. $0.1 \geq \frac{9}{n}$, $n \geq 90$. Измерения нужно производить не менее 90 раз.

Теорема Бернулли

Теорема Бернулли исторически является первой и наиболее простой формой закона больших чисел. Она теоретически обосновывает свойство устойчивости относительной частоты

Теорема 4.3. (Закон больших чисел в форме Я. Бернулли, 1713 г.). Если вероятность наступления события A в одном испытании равна p , число наступления этого события при n независимых испытаниях равно n_A , то для любого числа $\varepsilon > 0$ имеет место равенство

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1, \quad (4.5)$$

т.е. относительная частота $P^*(A)$ события A сходится по вероятности к вероятности p события A :

$$P^*(A) \xrightarrow[n \rightarrow \infty]{p} P(A).$$

Теорема Бернулли теоретически обосновывает возможность приближенного вычисления вероятности события с помощью его относительной частоты.

Обобщением теоремы Бернулли на случай, когда вероятности p_i появления события A в каждом из n испытаний различны, является *теорема Пуассона*:

$$P^*(A) \xrightarrow[n \rightarrow \infty]{p} \frac{1}{n} \sum_{i=1}^n p_i.$$

где p_i – вероятность события A в i -м испытании.

Пример 4.3. Вероятность наличия опечатки на одной странице рукописи

равна 0,2. Оценить вероятность того, что в рукописи, содержащей 400 страниц, частота появления опечатки отличается от соответствующей вероятности по модулю меньше, чем 0,05.

Воспользуемся формулой (4.4). В данном случае, $p=0,2$; $q=0,8$; $n=400$; $\varepsilon=0,05$.

$$P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{pq}{n\varepsilon^2}; \quad 1 - \frac{0,2 \cdot 0,8}{400 \cdot 0,05^2} = 0,84$$

Центральная предельная теорема

Центральная предельная теорема (ЦПТ) представляет собой вторую группу предельных теорем, которые устанавливают связь между законом распределения суммы случайных величин и его предельной формой – нормальным законом распределения.

Сформулируем ЦПТ для случая, когда члены суммы имеют одинаковое распределение (именно эта теорема чаще других используется на практике, так в математической статистике выборочные случайные величины имеют одинаковые распределения, так как получены из одной и той же генеральной совокупности).

Теорема 4.4. Пусть $X_1, X_2, \dots, X_n, \dots$ – независимые одинаково распределенные случайные величины с математическими ожиданиями $M(X_i) = m$ и дисперсиями $D(X_i) = \sigma^2, i = 1, 2, \dots, n, \dots$. Тогда для любого действительного числа x существует предел

$$\lim_{n \rightarrow \infty} P\left(\frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}} < x\right) = \Phi(x),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ – функция стандартного нормального распределения (функция Лапласа).

Эту теорему иногда называют теоремой Линдеберга-Леви.

В ряде прикладных задач не выполнено условие одинаковости распределенности. В таких случаях центральная предельная теорема обычно остается справедливой, однако на последовательность случайных величин приходится накладывать те или иные условия. Суть этих условий состоит в том, что ни одно слагаемое не должно быть доминирующим, вклад каждого слагаемого в среднее арифметическое должен быть пренебрежимо мал по сравнению с итоговой суммой.

История получения центральных предельных теорем для случайных величин растянулась на два века – от первых работ Муавра в 30-х годах 18-го века до необходимых и достаточных условий, полученных Линдебергом и Феллером в 30-х годах 20-го века.

Теорема 4.5. Теорема Линдберга-Феллера, устанавливающая условия асимптотичности нормальности функции распределения суммы независимых случайных величин, обладающих конечными дисперсиями.

Пусть $X_1, X_2, \dots, X_n, \dots$ – независимые случайные величины с математическими ожиданиями $M(X_i) = m_i$ и дисперсиями $D(X_i) = \sigma_i^2 \neq 0, i = 1, 2, \dots, n, \dots$. Предельное соотношение, т.е. центральная предельная теорема, выполнено тогда и только тогда, когда при любом $\tau > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-m_k| > \tau B_n} (x-m_k)^2 dF_k(x) = 0,$$

где $F_k(x)$ обозначает функцию распределения случайной величины X_k .

Достаточность была доказана Дж. Линдбергом, необходимость – В. Феллером.

ЦПТ позволяет при больших n (обычно ей пользуются если $n > 10$) вычислять вероятности различных событий, связанных с суммами случайных величин. Так, перейдя от случайной величины $S_n = \sum_{i=1}^n X_i$ к стандартной случайной величине, получим:

$$\begin{aligned} P\{\alpha \leq \sum_{i=1}^n X_i \leq \beta\} &= P\left\{\frac{\alpha - na}{\sigma\sqrt{n}} \leq \frac{\sum_{i=1}^n X_i - na}{\sigma\sqrt{n}} \leq \frac{\beta - na}{\sigma\sqrt{n}}\right\} \approx \\ &\approx \Phi\left(\frac{\beta - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{\alpha - na}{\sigma\sqrt{n}}\right). \end{aligned}$$

Или

$$P\{\alpha \leq S_n \leq \beta\} = \Phi\left(\frac{\beta - MS_n}{\sqrt{DS_n}}\right) - \Phi\left(\frac{\alpha - MS_n}{\sqrt{DS_n}}\right) \quad (4.6)$$

Пример 4.4. Независимые случайные величины X_i распределены равномерно на отрезке $[0,1]$. Найти закон распределения случайной величины $Y = \sum_{i=1}^{100} X_i$, а также вероятность того, что $55 < Y < 70$.

Решение. Условия ЦПТ соблюдаются, поэтому случайная величина Y имеет приближенно плотность распределения

$$f_Y(y) \approx \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}.$$

По известным формулам для математического ожидания и дисперсии в

случае равномерного распределения находим:

$$MX_i = \frac{0+1}{2} = \frac{1}{2}, \quad DX_i = \frac{(1-0)^2}{12} = \frac{1}{12}, \quad \sigma X_i = \frac{1}{\sqrt{12}} = \frac{1}{2\sqrt{3}}.$$

Тогда

$$m_y = M\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} MX_i = 100 \cdot \frac{1}{2} = 50.$$

$$\sigma_y^2 = D\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} DX_i = 100 \cdot \frac{1}{12} = \frac{25}{3}, \quad \sigma_y = \frac{5\sqrt{3}}{3}.$$

Таким образом, $f_Y(y) \approx \frac{3}{5\sqrt{6\pi}} e^{-\frac{3(y-50)^2}{50}}$.

Используя формулу (4.6), находим

$$P\{55 < Y < 70\} = \Phi\left(\frac{70-50}{\frac{5\sqrt{3}}{3}}\right) - \Phi\left(\frac{55-50}{\frac{5\sqrt{3}}{3}}\right) = \Phi(6.9282) - \Phi(1.73) \approx 0.04.$$

Доска Гальтона

Доска Гальтона (рис. 35) – устройство, изобретённое английским учёным Фрэнсисом Гальтоном (первый экземпляр изготовлен в 1873 г.) и предназначенное для демонстрации центральной предельной теоремы.

Доска Гальтона представляет собой ящик с прозрачной передней стенкой. В заднюю стенку в шахматном порядке вбиты штырьки, образующие треугольник. Сверху в ящик через воронку (выход из которой расположен ровно посередине между левой и правой стенками) кидаются шарики. В идеальном случае сталкиваясь со штырьком, шарик каждый раз с одинаковой вероятностью может повернуть либо направо, либо налево. Нижняя часть ящика разделена перегородками (число которых равно числу штырьков в нижнем ряду), в результате чего шарики, скатываясь на дно ящика, образуют столбики, которые тем выше, чем ближе к середине доски (при достаточно большом числе шариков внешний вид столбиков приближается к кривой нормального распределения).

Обозначим как n общее число столкновений шарика со штырьками; как k число раз, когда шарик поворачивает направо (таким образом, он оказывается в k -м по порядку столбике). Тогда число способов, которыми он может добраться до k -го столбика, определяется биномиальным коэффициентом C_n^k . Отсюда следует, что вероятность оказаться в k -м столбике равна $C_n^k p^k (1-p)^{n-k}$, где p – вероятность поворота направо (обычно можно считать, что $p=0.5$). Это функция вероятности биномиального распределения, которое в

соответствии с центральной предельной теоремой при достаточно большом n аппроксимирует нормальное распределение.

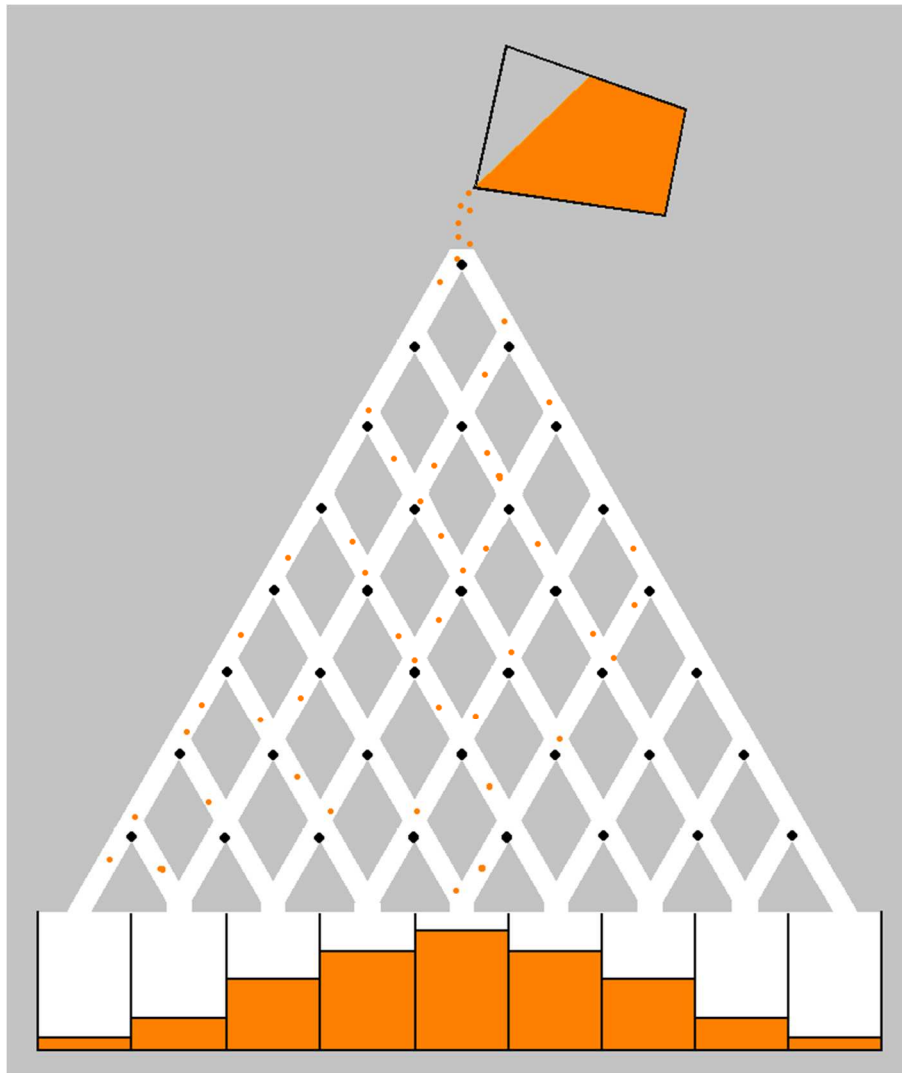


Рис. 35. Доска Гальтона

II. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Введение



Математическая статистика как наука начинается с работ знаменитого немецкого математика Карла Фридриха Гаусса (1777-1855), который на основе теории вероятностей исследовал и обосновал метод наименьших квадратов, созданный им в 1795 г. и примененный для обработки астрономических данных (с целью уточнения орбиты малой планеты Церера).

Серьезные работы по теории ошибок измерений конца XVIII в., по обработке результатов биологических исследований в XIX в. позволили выделить математическую статистику в особую науку. В конце XIX в. – начале XX в. крупный вклад в математическую статистику внесли английские исследователи, прежде всего К. Пирсон (1857-1936) и Р.А. Фишер (1890-1962). В частности, Пирсон разработал критерий «хи-квадрат» проверки статистических гипотез, а Фишер – дисперсионный анализ, теорию планирования эксперимента, метод максимального правдоподобия оценки параметров.

Большой вклад в математическую статистику внесли ученые русской математической школы П. Л. Чебышев (1821-1894), А.А. Марков (1856-1922), А. М. Ляпунов (1857-1918). Их трудами она была поставлена на прочную математическую основу и сделана эффективным средством познания.

В 30-е годы XX в. поляк Ежи Нейман (1894-1977) и англичанин Э. Пирсон развили общую теорию проверки статистических гипотез, а советские математики академик А.Н. Колмогоров (1903-1987) и член-корреспондент АН СССР Н.В. Смирнов (1900-1966) заложили основы непараметрической статистики. В сороковые годы XX в. румын А. Вальд (1902-1950) построил теорию последовательного статистического анализа.

Математическая статистика бурно развивается и в настоящее время. Так, за последние 50 лет можно выделить четыре принципиально новых направления исследований:

- Разработка и внедрение математических методов планирования экспериментов;
- Развитие статистики объектов нечисловой природы как самостоятельного направления в прикладной математической статистике;
- Развитие статистических методов, устойчивых по отношению к малым отклонениям от используемой вероятностной модели;
- Широкое развертывание работ по созданию компьютерных пакетов программ, предназначенных для проведения статистического анализа данных.

5. Выборки и их характеристики

Предмет математической статистики

Математическая статистика – это математическая наука, посвященная разработке методов описания и анализа статистических экспериментальных данных, полученных в результате наблюдений массовых случайных явлений.

Выводы о закономерностях, которым подчиняются явления, изучаемые методами математической статистики, всегда основываются на ограниченном, выборочном числе наблюдений. Поэтому естественно предположить, что эти выводы при большем числе наблюдений могут оказаться иными. Чтобы быть в состоянии высказать более определенное суждение об изучаемом явлении, математическая статистика опирается на теорию вероятностей.

Обе эти математические дисциплины изучают массовые случайные явления. Связующим звеном между ними являются предельные теоремы теории вероятностей. При этом теория вероятностей выводит из математической модели свойства реального процесса, а математическая статистика устанавливает свойства математической модели, исходя из данных наблюдений (говорят «из статистических данных»).

Предметом математической статистики является изучение случайных величин (или случайных событий, процессов) по результатам наблюдений.

Задачи математической статистики:

1. Указать способы получения, группировки и обработки статистических данных, собранных в результате наблюдений, специально поставленных опытов или произведённых измерений.

2. Разработка методов анализа статистических сведений в зависимости от целей исследования. Например, целью исследования может быть:

- оценка неизвестной вероятности события;
- оценка параметров распределения случайной величины;
- оценка неизвестной функции распределения случайной величины;
- проверка гипотез о параметрах распределения или о виде неизвестного распределения;
- оценка зависимости случайной величины от одной или нескольких случайных величин и т.д.

Для обработки статистических данных созданы специальные программные пакеты (*STADIA*, СтатЭксперт, Эвриста, *SYSTAT*, *STAT- GRAPHICS* и др.), которые выполняют трудоемкую работу по расчету различных статистик, построению таблиц и графиков. Простейшие статистические функции имеются в программируемых калькуляторах и популярных офисных программах (*EXCEL*).

Результаты исследования статистических данных методами математической статистики *используются для принятия решения* (в задачах планирования,

управления, прогнозирования и организации производства, при контроле качества продукции, при выборе оптимального времени настройки или замены действующей аппаратуры и т.д.), т.е. для научных и практических выводов.

Генеральная и выборочная совокупности

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить стандартность детали, а количественным – контролируемый размер детали. Каждый такой признак (и их комбинации) образует случайную величину, наблюдения над которой мы и производим.

Совокупность всех подлежащих изучению объектов или возможных результатов всех мыслимых наблюдений, производимых в неизменных условиях над одним объектом, называется *генеральной совокупностью*.

Более строго: генеральная совокупность – это случайная величина $X(w)$, заданная на пространстве элементарных событий Ω с выделенным в нем классом S подмножеств событий, для которых указаны их вероятности.

Зачастую проводить *сплошное обследование*, когда изучаются все объекты (например – перепись населения), трудно или дорого, экономически нецелесообразно (например – не вскрывать же каждую упаковку для проверки качества продукции), а иногда невозможно. В этих случаях наилучшим способом обследования является *выборочное наблюдение*: выбирают из генеральной совокупности часть ее объектов («выборку») и подвергают их изучению.

Выборочной совокупностью (выборкой) называется совокупность объектов, отобранных случайным образом из генеральной совокупности.

Более строго: *выборка* – это последовательность X_1, X_2, \dots, X_n независимых одинаково распределенных случайных величин., распределение каждой из которых совпадает с распределением генеральной случайной величины.

Число объектов (наблюдений) в совокупности называется ее *объемом*.

Например, если из 1000 деталей отобрано для обследования 100 деталей, то объем генеральной совокупности $N = 1000$, а объем выборки $n=100$. Число объектов генеральной совокупности N значительно превосходит объем выборки n .

Конкретные значения выборки, полученные в результате наблюдений (испытаний), называют *реализацией выборки* и обозначают строчными буквами x_1, x_2, \dots, x_n . Метод статистического исследования, состоящий в том, что на основе изучения выборочной совокупности делается заключение о всей генеральной совокупности, называется *выборочным*.

Для получения хороших оценок характеристик генеральной совокупности необходимо, чтобы выборка была *репрезентативной* (или *представительной*), т.е. достаточно полно представлять изучаемые признаки генеральной совокупности. Условием обеспечения репрезентативности выборки является, согласно закону больших чисел, соблюдение случайности отбора, т. е. все объекты генеральной совокупности должны иметь равные вероятности попасть в выборку.

Есть три разновидности лжи: ложь, гнусная ложь и статистика

Приписывается английскому писателю и государственному деятелю, премьер-министру Великобритании от партии консерваторов (1874—1880) Бенджамину Дизраэли, лорду Биконсфильду (1804—1881). Но в его работах и высказываниях такая фраза не обнаружена.

В литературе впервые встречается у американского писателя Марка Твена (1835—1910) в его «Автобиографии» (опубл. 1924), где эта фраза приводится со ссылкой на Бенджамина Дизраэли.

Имеется ввиду, что мы получим, если будем использовать нерепрезентативную выборку.

Различают выборки *с возвращением (повторные)* и *без возвращения (бесповторные)*. В первом случае отобранный объект возвращается в генеральную совокупность перед извлечением следующего; во втором – не возвращается. На практике чаще используется бесповторная выборка.

Заметим, если объем выборки значительно меньше объема генеральной совокупности, различие между повторной и бесповторной выборками очень мало, его можно не учитывать.

В зависимости от конкретных условий для обеспечения репрезентативности применяют различные способы отбора: *простой*, при котором из генеральной совокупности извлекают по одному объекту; *типический*, при котором генеральную совокупность делят на «типические» части и отбор осуществляется из каждой части (например, мнение о референдуме спросить у случайно отобранных людей, разделенных по признаку пола, возраста,...); *механический*, при котором отбор производится через определенный интервал (например, мнение спросить у каждого шестидесятого...); *серийный*, при котором объекты из генеральной совокупности отбираются «сериями», которые должны исследоваться при помощи сплошного обследования.

На практике пользуются сочетанием вышеупомянутых способов отбора.

Пример 5.1. Десять абитуриентов проходят тестирование по математике. Каждый из них может набрать от 0 до 5 баллов включительно. Пусть X_k – количество баллов, набранных каждым ($k = 1, 2, \dots, 10$) абитуриентом.

Тогда значения 0, 1, 2, 3, 4, 5 – все возможные количества баллов, набранных одним абитуриентом, – образуют генеральную совокупность.

Выборка X_1, X_2, \dots, X_{10} – результат тестирования 10 абитуриентов.

Реализациями выборки могут быть следующие наборы чисел: {5, 3, 0, 1, 4, 2, 5, 4, 1, 5} или {4, 4, 5, 3, 3, 1, 5, 5, 2, 5} или {3, 4, 5, 0, 1, 2, 3, 4, 5, 4} и т.д.

Статистическое распределение выборки. Эмпирическая функция распределения

Пусть изучается некоторая случайная величина X . С этой целью над ней производится ряд независимых опытов (наблюдений). В каждом из этих опытов величина X принимает то или иное значение.

Пусть она приняла n_1 раз значение x_1 , n_2 раз – значение x_2 , ..., n_k раз – значение x_k . При этом $n_1 + n_2 + \dots + n_k = n$ – *объем выборки*. Значения x_1, x_2, \dots, x_k называются *вариантами* случайной величины X .

Вся совокупность значений случайной величины X представляет собой первичный статистический материал, который подлежит дальнейшей обработке, прежде всего – упорядочению.

Операция расположения значений случайной величины (признака) по неубыванию называется *ранжированием* статистических данных.

Полученная таким образом последовательность $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ (где $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ и $x_{(1)} = \min_{1 \leq i \leq n} X_i, \dots, x_{(n)} = \max_{1 \leq i \leq n} X_i$) называется *вариационным рядом*.

Числа n_i , показывающие, сколько раз встречаются варианты x_i в ряде наблюдений, называются *частотами*, а отношение их к объему выборки – *частостями* или *относительными частотами* (p_i^*), т.е.

$$p_i^* = \frac{n_i}{n}, \text{ где } n = \sum_{i=1}^k n_i \quad (5.1)$$

Перечень вариантов и соответствующих им частот или частостей называется *статистическим распределением выборки* или *статистическим рядом*.

Записывается статистическое распределение в виде таблицы. Первая строка содержит варианты, а вторая – их частоты n (или частости p_i^*).

Пример 5.2. В результате тестирования группа абитуриентов набрала баллы: 5, 3, 0, 1, 4, 2, 5, 4, 1, 5. Записать полученную выборку в виде: а) вариационного ряда; б) статистического ряда.

Решение.

а) Проранжировав статистические данные, получим вариационный ряд:
(0, 1, 1, 2, 3, 4, 4, 5, 5, 5).

б) Подсчитав частоту и частость вариантов $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3, x_5 = 4, x_6 = 5$, получим статистическое распределение выборки (дискретный статистический ряд).

x_i	0	1	2	3	4	5	$\sum_{i=1}^6 n_i = 10$
n_i	1	2	1	1	2	3	

или

x_i	0	1	2	3	4	5	$\sum_{i=1}^6 p_i^* = 1$
p_i^*	1/10	2/10	1/10	1/10	2/10	3/10	

Статистическое распределение выборки является оценкой неизвестного распределения. В соответствии с теоремой Бернулли относительные частоты p_i^* сходятся при $n \rightarrow \infty$ к соответствующим вероятностям p_i . Поэтому при больших значениях n статистическое распределение мало отличается от истинного распределения.

В случае, когда число значений признака велико или признак является непрерывным (т. е. когда X может принять любое значение в некотором интервале), составляют *интервальный статистический ряд*. В первую строку таблицы статистического распределения вписывают частичные промежутки $[x_0, x_1], [x_1, x_2], \dots, [x_{k-1}, x_k]$ которые берут обычно одинаковыми по длине. Для определения величины интервала (h) можно использовать формулу Стерджеса:

$$h = \frac{x_{\max} - x_{\min}}{1 + \log_2 n} = \frac{x_{\max} - x_{\min}}{1 + 3.322 \lg n}$$

За начало первого интервала рекомендуется брать величину $x_{\text{нач}} = x_{\min} - \frac{h}{2}$

Пример 5.3. Измерили рост 30 наудачу отобранных студентов. Результаты измерений таковы:

178, 160, 154, 183, 155, 153, 167, 186, 163, 155,
157, 175, 170, 166, 159, 173, 182, 167, 171, 169,
179, 165, 156, 179, 158, 171, 175, 173, 164, 172

Построить интервальный статистический ряд.

Решение. Проранжируем полученные данные.

153, 154, 155, 155, 156, 157, 158, 159, 160, 163, 164, 165, 166, 167, 167, 169, 170, 171, 171, 172, 173, 173, 175, 175, 178, 179, 179, 182, 183, 186.

X – рост студента является непрерывной случайной величиной. При более точном измерении рост значения случайной величины X обычно не повторяется (вероятность наличия на Земле двух человек, рост которых, скажем равен $\sqrt{3} \approx 1.732050808... \text{ м}$, равна нулю!).

По формуле Стерджеса находим длину частичного интервала h :

$$h = \frac{186 - 153}{1 + \log_2 30} \approx \frac{33}{5.907} \approx 5.59.$$

Примем $h = 6$. Тогда $x_{\text{нач}} = 153 - 6/2 = 150$. Исходные данные разбиваем на 6 интервалов. Подсчитав число студентов n_i , попавших в каждый из полученных промежутков, получим интервальный статистический ряд.

Рост	[150 – 156]	[156 – 162]	[162 – 168]	[168 – 174]	[174 – 180]	[180 – 186]
Частота	4	5	6	7	5	3
Частость	0.13	0.17	0.20	0.23	0.17	0.10

Одним из способов обработки вариационного ряда является построение эмпирической функции распределения, определяющей для каждого значения x частоту события $\{X < x\}$:

$$F_n^*(x) = p^* \{X < x\} = \frac{n_x}{n},$$

где n – объем выборки; n_x – число наблюдений, меньших x ; ($x \in R$).

Эмпирическая функция распределения $F_n^*(x)$ является оценкой теоретической функции распределения $F(x)$ случайной величины X . Имеет место

$$\lim_{n \rightarrow \infty} P\{|F_n^*(x) - F(x)| < \varepsilon\} = 1.$$

Построим эмпирическую функцию $F_n^*(x)$ для условий и результатов примера

5.2. Здесь $n=10$. Имеем $F_{10}^*(x) = \frac{0}{10} = 0$ при $x \leq 0$ (наблюдений меньших 0 нет);

$F_{10}^*(x) = \frac{1}{10} = 0$ при $0 < x \leq 1$ (здесь $n_x=1$) и т.д. Окончательно получаем

$$F_{10}^*(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ 0.1, & \text{если } 0 < x \leq 1; \\ 0.3, & \text{если } 1 < x \leq 2; \\ 0.4, & \text{если } 2 < x \leq 3; \\ 0.5, & \text{если } 3 < x \leq 4; \\ 0.7, & \text{если } 4 < x \leq 5; \\ 1, & \text{если } 5 < x. \end{cases}$$

График эмпирической функции $F_{10}^*(x)$ приведен на рис. 1.

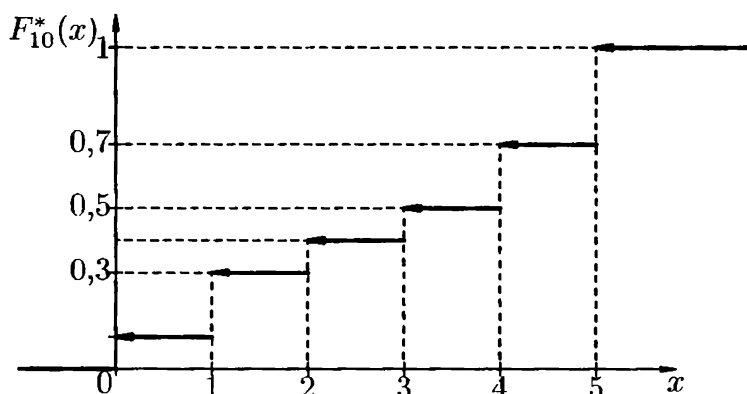


Рис. 1. Эмпирическая функция распределения

Статистическое распределение изображается графически в виде так называемых *полигона частот (частостей)* и *гистограммы*.

Полигон, как правило, служит для изображения дискретного (т. е. варианты отличаются на постоянную величину) статистического ряда. *Полигоном частот* называют ломаную, отрезки которой соединяют точки с координатами (x_1, n_1) , (x_2, n_2) , ..., (x_k, n_k) ; *полигоном частностей* – с координатами (x_1, p_1^*) , (x_2, p_2^*) , ..., (x_k, p_k^*) (рис. 2.)

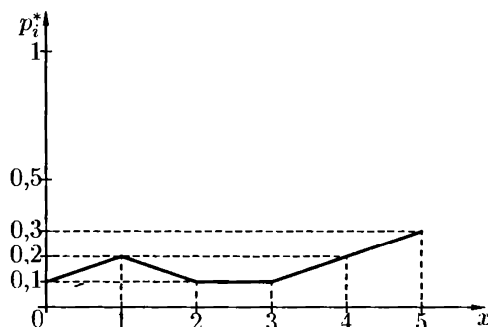


Рис. 2. Полигон частостей

Более употребительной является гистограмма. *Гистограммой частот (частостей)* называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины h , а высоты равны отношению $\frac{n_i}{h}$ (гистограмма частот) или $\frac{n_i}{n \cdot h} = \frac{p_i^*}{h}$ (гистограмма частостей).

Очевидно, площадь гистограммы частот равна объему выборки, а площадь гистограммы частостей равна единице.

Построим гистограмму частостей для условий и результатов примера 5.3. Здесь длина интервала равна $h=6$, а высоты h_i прямоугольников:

$$h_1 = \frac{0.13}{6} \approx 0.022, h_2 = \frac{0.17}{6} \approx 0.028, h_3 = \frac{0.20}{6} \approx 0.033,$$

$$h_4 = \frac{0.23}{6} \approx 0.038, h_5 = \frac{0.17}{6} \approx 0.028, h_6 = \frac{0.1}{6} = 0.017.$$

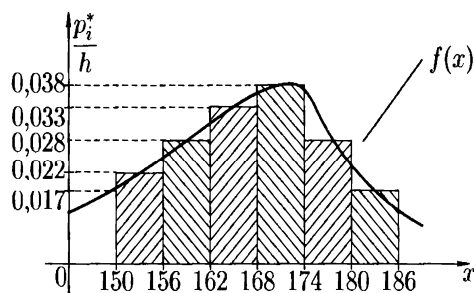


Рис. 3. Гистограмма частостей

Гистограмма частот является статистическим аналогом функции плотности $f(x)$ случайной величины X .

Числовые характеристики статистического распределения

Можно определить для выборки ряд числовых характеристик, подобных тем, что в теории вероятностей определялись для случайных величин.

Пусть статистическое распределение выборки объема n имеет вид:

x_i	x_1	x_2	x_3	...	x_k
n_i	n_1	n_2	n_3	...	n_k

Выборочным средним \bar{x} называется среднее арифметическое всех значений выборки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i = \sum_{i=1}^k x_i \cdot p_i^* \quad (5.2),$$

где $p_i^* = \frac{n_i}{n}$ – частость.

В случае интервального статистического ряда в равенстве (5.2) в качестве x_i берут середины его интервалов, а n_i – соответствующие им частоты.

Выборочной дисперсией D_B называется среднее арифметическое квадратов отклонений значений выборки от выборочной средней \bar{x} , т. е.

$$D_B = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x})^2 \cdot n_i = \sum_{i=1}^k (x_i - \bar{x})^2 \cdot p_i^* \quad (5.3)$$

Можно показать, что D_B может быть подсчитана также по формуле:

$$D_B = \frac{1}{n} \sum_{i=1}^k x_i^2 \cdot n_i - (\bar{x})^2 \quad (5.4)$$

Выборочное среднее квадратическое отклонение выборки определяется формулой

$$\sigma_B = \sqrt{D_B} \quad (5.5)$$

Особенность выборочного среднего квадратического отклонения состоит в том, что оно измеряется в тех же единицах, что и изучаемый признак.

В практических задачах часто используется *исправленная выборочная дисперсия*

$$S^2 = \frac{n}{n-1} D_B \quad (5.6)$$

Размахом вариации называется число $R = x_{max} - x_{min}$

Модой M_o^* вариационного ряда называется вариант, имеющий наибольшую частоту.

Медианой M_e^* вариационного ряда называется значение признака (случайная величина X), приходящееся на середину ряда.

6. Элементы теории оценок

Оценка неизвестных параметров

Пусть закон распределения изучаемой случайной величины X зависит от одного или нескольких параметров. Требуется по выборке x_1, x_2, \dots, x_n , полученной в результате n наблюдений, оценить неизвестный параметр θ .

Статистической оценкой $\tilde{\theta}$ параметра θ теоретического распределения называют его приближенное значение, зависящее от данных выбора.

$$\tilde{\theta} = \tilde{\theta}(X_1, X_2, \dots, X_n)$$

Функцию результатов наблюдений (т. е. функцию выборки) называют *статистикой*.

Так, $F^*(x)$ есть оценка $F(x)$, гистограмма – плотности $f(x)$.

Оценка $\tilde{\theta}$ является случайной величиной, так как является функцией независимых случайных величин X_1, X_2, \dots, X_n . Если произвести другую выборку, то функция примет, вообще говоря, другое значение.

Если число опытов невелико, то замена неизвестного параметра θ его оценкой $\tilde{\theta}$, например, математического ожидания средним арифметическим, приводит к ошибке. Это ошибка в среднем тем больше, чем меньше число опытов.

К оценке любого параметра предъявляется ряд требований, которым она должна удовлетворять, чтобы быть «близкой» к истинному значению параметра, т. е. быть в каком-то смысле «доброкачественной» оценкой.

Свойства статистических оценок

Качество оценки определяют, проверяя, обладает ли она свойствами несмещенности, состоятельности и эффективности.

1. Оценка $\tilde{\theta}$ параметра θ называется *несмещенной*, если $M\tilde{\theta} = \theta$.

Если $M\tilde{\theta} \neq \theta$, то оценка $\tilde{\theta}$ называется *смещенной*.

Чтобы оценка $\tilde{\theta}$ не давала систематической ошибки (ошибки одного знака) в сторону завышения ($M\tilde{\theta} > \theta$) или занижения ($M\tilde{\theta} < \theta$), надо потребовать, чтобы математическое ожидание оценки было равно оцениваемому параметру.

Если $M\tilde{\theta}_n \rightarrow \theta$, то оценка $\tilde{\theta}_n$ называется *асимптотически несмещенной*. Требование несмещенности особенно важно при малом числе опытов.

2. Оценка $\tilde{\theta}_n$ параметра θ называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру: $\tilde{\theta}_n \xrightarrow[n \rightarrow \infty]{p} \theta$, т.е. для любого $\varepsilon > 0$ выполнено

$$\lim_{n \rightarrow \infty} P\{|\tilde{\theta}_n - \theta| < \varepsilon\} = 1.$$

Это означает, что при увеличении объема выборки, мы все ближе приближаемся к истинному значению параметра θ . Свойство состоятельности обязательно для любого правила оценивания (несостоятельные оценки не используются).

3. Несмещенная оценка $\tilde{\theta}$ параметра θ называется *эффективной*, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра θ .

На практике не всегда удается удовлетворить всем перечисленным выше требованиям (несмещенность, состоятельность, эффективность), и поэтому приходится довольствоваться оценками, не обладающими сразу всеми тремя свойствами. Однако выбору оценки всегда должно предшествовать ее критическое рассмотрение со всех точек зрения.

Точечные оценки математического ожидания и дисперсии

Пусть изучается случайная величина X с математическим ожиданием $a = MX$ и дисперсией DX , причем оба параметра неизвестны.

Статистика, используемая в качестве приближенного значения неизвестного параметра генеральной совокупности, называется ее *точечной оценкой*. То есть точечная оценка характеристики генеральной совокупности – это число, определяемое по выборке.

Под X_i будем понимать значение случайной величины X в i -м опыте. Случайные величины X_1, X_2, \dots, X_n можно рассматривать как n независимых «экземпляров» величины X . Поэтому $MX_1 = MX_2 = \dots = MX_n = MX = a, DX_1 = DX_2 = \dots = DX_n = DX$.

Найдем математическое ожидание оценки выборочного среднего \bar{X}_B

$$M\bar{X}_B = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} M\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n MX_i = \frac{1}{n} \cdot n \cdot a = a.$$

Отсюда по определению получаем, что \bar{X}_B – несмещенная оценка MX .

Состоятельность оценки $\tilde{\theta}_n$ может быть установлена с помощью следующей теоремы.

Теорема 6.1. Если оценка $\tilde{\theta}_n$ параметра θ является несмещенной и $D\tilde{\theta}_n \rightarrow 0$ при $n \rightarrow \infty$, то $\tilde{\theta}_n$ – состоятельная оценка.

Доказательство. Запишем неравенство Чебышева для случайной величины $\tilde{\theta}_n$ для любого $\varepsilon > 0$

$$P(|\tilde{\theta}_n - \theta| < \varepsilon) \geq 1 - \frac{D\tilde{\theta}_n}{\varepsilon^2}.$$

Так как по условию $\lim_{n \rightarrow \infty} D\tilde{\theta}_n = 0$, то $\lim_{n \rightarrow \infty} P(|\tilde{\theta}_n - \theta| < \varepsilon) \geq 1$. Но вероятность любого события не превышает 1 и, следовательно,

$$\lim_{n \rightarrow \infty} (|\tilde{\theta}_n - \theta| < \varepsilon) = 1,$$

т.е. $\tilde{\theta}_n$ – состоятельная оценка параметра θ .

При нормальном распределении случайной величины X эта оценка будет и эффективной. На практике во всех случаях в качестве оценки математического ожидания используется среднее арифметическое, т.е. \bar{X}_B .

Можно доказать, что $MD_B = \frac{n}{n-1} DX$.

Отсюда следует, что $MD_B \neq DX$, т.е. выборочная дисперсия является смещенной оценкой дисперсии DX . Поэтому выборочную дисперсию исправляют, умножив ее на $\frac{n}{n-1}$, получая формулу:

$$S^2 = \frac{n}{n-1} D_B - \text{исправленная дисперсия.}$$

Дробь $\frac{n}{n-1}$ называют поправкой Бесселя. При малых n поправка Бесселя довольно значительно отличается от единицы, с увеличением же n она быстро стремится к единице. При $n > 50$ практически нет разницы между D_B и S^2 .

Таким образом, S^2 – несмещенная оценка DX . Можно доказать, что она также является и состоятельной оценкой DX .

Методы нахождения точечных оценок

Рассмотрим наиболее распространенные методы получения точечных оценок параметров распределения: метод моментов, метод максимального правдоподобия и метод наименьших квадратов.

Метод моментов

Метод моментов для нахождения точечных оценок неизвестных параметров заданного распределения состоит в приравнивании теоретических моментов распределения соответствующим эмпирическим моментам, найденных по выборке.

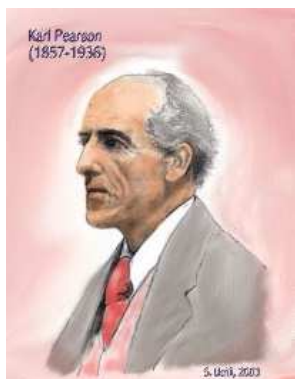
Так, если распределение зависит от одного параметра θ (например, задан вид плотности распределения $f(x, \theta)$), то для нахождения его оценки надо решить относительно θ одно уравнение:

$$MX = \bar{X}_B$$

Если распределение зависит от двух параметров (например, вид плотности распределения $f(x, \theta_1, \theta_2)$) – надо решить относительно θ_1 и θ_2 систему уравнений:

$$\begin{cases} MX = \bar{x}_B, \\ DX = D_B. \end{cases}$$

Метод моментов был предложен в 1894 г. К. Пирсоном. Оценки метода моментов обычно состоятельны, однако их эффективность часто значительно меньше единицы.



Английский математик, статистик, биолог и философ; основатель математической статистики, один из основоположников биометрики. Автор свыше 650 опубликованных научных работ.

Пример 6.1. Найти оценки параметров нормального распределения случайной величины X методом моментов.

Решение. Требуется по выборке x_1, x_2, \dots, x_n найти точечные оценки неизвестных параметров $a = MX = \theta_1$ и $\sigma = \sqrt{DX} = \theta_2$.

По методу моментов приравниваем их, соответственно, к выборочному среднему и выборочной дисперсии ($\alpha_1 = MX$ – начальный момент 1-го порядка, $\mu_2 = DX$ – центральный момент 2-го порядка). Получаем

$$\begin{cases} MX = \bar{x}_B, \\ DX = D_B. \end{cases}$$

Т.е.

$$\begin{cases} a = \bar{x}_B, \\ \sigma^2 = D_B. \end{cases}$$

Таким образом, искомые оценки параметров нормального распределения:

$$\tilde{\theta}_1 = \bar{x}_B \text{ и } \tilde{\theta}_2 = \sqrt{D_B}.$$

Метод максимального правдоподобия

Пусть x_1, x_2, \dots, x_n – выборка, полученная в результате проведения n независимых наблюдений за случайной величиной X . И пусть вид закона распределения величины X , например, вид плотности $f(x, \theta)$, известен, но неизвестен параметр θ , которым определяется этот закон. Требуется по выборке оценить параметр θ .

В основе метода максимального правдоподобия, предложенного Р. Фишером, лежит понятие функции правдоподобия.

Функцией правдоподобия, построенной по выборке x_1, x_2, \dots, x_n , называется функция аргумента θ вида

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta)$$

$$L(x, \theta) = \prod_{i=1}^n f(x_i, \theta),$$

где $f(x, \theta)$ – плотность распределения случайной величины X в случае, если X – непрерывная.

Если X – дискретная случайная величина, то функция правдоподобия имеет вид

$$L(x_1, x_2, \dots, x_n; \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \cdot \dots \cdot p(x_n, \theta) = \prod_{i=1}^n p(x_i, \theta),$$

где $p(x_i, \theta) = p\{X = x_i, \theta\}$.

Из определения следует, что чем больше значение функции $L(x, \theta)$, тем более вероятно (правдоподобнее) появление (при фиксированном θ) в результате наблюдений чисел x_1, x_2, \dots, x_n .

За точечную оценку параметра θ , согласно методу максимального правдоподобия, берут такое его значение $\tilde{\theta}$, при котором функция правдоподобия достигает максимума.

Эта оценка, называемая *оценкой максимального правдоподобия*, является решением уравнения

$$\left. \frac{dL(x, \theta)}{d\theta} \right|_{\theta=\tilde{\theta}} = 0.$$

Так как функции $L(x, \theta)$ и $\ln L(x, \theta)$ достигают максимума при одном и том же значении θ , то вместо отыскания максимума функции $L(x, \theta)$ ищут (что проще) максимум функции $\ln L(x, \theta)$.

Пример 6.2. Найти оценку параметра, a распределения Пуассона методом максимального правдоподобия.

Решение. В данном случае $P\{X = m\} = \frac{a^m \cdot e^{-a}}{m!}$. Поэтому

$$p(x_i, \theta) = P\{X = x_i, \theta\} = \frac{\theta^{x_i} \cdot e^{-\theta}}{x_i!}$$

при $x_i \in \mathbb{N}$. Составляем функцию правдоподобия для дискретной случайной величины X .

$$L(x, \theta) = \frac{\theta^{x_1} \cdot e^{-\theta}}{x_1!} \cdot \frac{\theta^{x_2} \cdot e^{-\theta}}{x_2!} \cdot \dots \cdot \frac{\theta^{x_n} \cdot e^{-\theta}}{x_n!} = e^{-n\theta} \cdot \theta^{\sum_{i=1}^n x_i} \cdot \frac{1}{x_1! \cdot \dots \cdot x_n!}.$$

Тогда

$$\ln L(x, \theta) = -n \cdot \theta + \sum_{i=1}^n x_i \cdot \ln \theta - \ln(x_1! \cdot x_2! \cdot \dots \cdot x_n!)$$

и

$$\frac{d \ln L(x, \theta)}{d \theta} = -n + \frac{1}{\theta} \cdot \sum_{i=1}^n x_i.$$

$$\left(-n + \frac{1}{\theta} \cdot \sum_{i=1}^n x_i\right) \Big|_{\theta=\tilde{\theta}} = 0.$$

Отсюда находим

$$\tilde{\theta} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_B.$$

А так как

$$\frac{d^2 \ln L(x, \theta)}{d \theta^2} \Big|_{\theta=\tilde{\theta}} = -\frac{1}{\tilde{\theta}^2} \sum_{i=1}^n x_i < 0,$$

то оценка $\tilde{\theta} = \bar{x}_B$ является оценкой максимального правдоподобия.

Итак $\tilde{\theta} = \tilde{a} = \bar{x}_B$.

Метод наименьших квадратов

Метод нахождения оценки $\tilde{\theta}$ неизвестного параметра θ , основанный на минимизации суммы квадратов отклонений выборочных данных от определяемой оценки θ , называется методом наименьших квадратов (МНК).

Другими словами, в МНК требуется найти такое значение $\tilde{\theta}$, которое бы минимизировало сумму

$$F(\theta) = \sum_{i=1}^n (X_i - \theta)^2 \rightarrow \min.$$

Пример 6.3. Найти оценку параметра a распределения Пуассона методом наименьших квадратов.

Решение. Найдем точку минимума функции $F(\theta) = \sum_{i=1}^n (X_i - \theta)^2$;

$$F'(\theta) = -2 \sum_{i=1}^n (X_i - \theta);$$

Из уравнения $F'(\theta) = 0$ находим критическую точку: $-2 \sum_{i=1}^n (X_i - \theta) = 0$.

Отсюда $\theta_{кр} = \frac{1}{n} \sum_{i=1}^n X_i$. А так как $F''(\theta_{кр}) = (-2 \sum_{i=1}^n (X_i - \theta_{кр}))' = 2n > 0$ при любом значении θ , то $\theta_{кр} = \frac{1}{n} \sum_{i=1}^n X_i$ — точка минимума функции $F(\theta)$.

Таким образом оценкой параметра a в распределении Пуассона согласно МНК является $\tilde{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$.

Можно доказать, что $M(\tilde{\theta}) = \theta = a$, $D(\tilde{\theta}) = \frac{\theta}{n}$.

Интервальное оценивание параметров

Точечные оценки неизвестного параметра θ хороши в качестве первоначальных результатов обработки наблюдений. Их недостаток в том, что неизвестно, с какой точностью они дают оцениваемый параметр.

Для выборок небольшого объема вопрос о точности оценок очень существенен, так как между параметром θ и его оценкой $\tilde{\theta}$ может быть большое расхождение в этом случае. Кроме того, при решении практических задач часто требуется определить и надежность этих оценок. Тогда и возникает задача о приближении параметра $\tilde{\theta}$ не одним числом, а целым интервалом $(\tilde{\theta}_1, \tilde{\theta}_2)$.

Оценка неизвестного параметра называется *интервальной*, если она определяется двумя числами – концами интервала.

Задачу интервального оценивания можно сформулировать так: по данным выборки построить числовой интервал $(\tilde{\theta}_1, \tilde{\theta}_2)$, относительно которого с заранее выбранной вероятностью γ можно сказать, что внутри этого интервала находится точное значение оцениваемого параметра.

Интервал $(\tilde{\theta}_1, \tilde{\theta}_2)$ накрывающий с вероятностью γ истинное значение параметра θ , называется *доверительным интервалом*, а вероятность γ – *надежностью оценки* или *доверительной вероятностью*.

Величина γ выбирается заранее, ее выбор зависит от конкретно решаемой задачи. Надежность γ принято выбирать равной 0,9; 0,95; 0,99 или 0,999. Тогда практически достоверно нахождение параметра θ в доверительном интервале $(\tilde{\theta}_1 - \varepsilon, \tilde{\theta}_2 + \varepsilon)$.

Доверительные интервалы для параметров нормального распределения

Построим доверительные интервалы для параметров нормального распределения, т. е. когда выборка производится из генеральной совокупности, имеющей нормальное распределение с параметрами a и σ .

Доверительный интервал для математического ожидания при известной дисперсии

Пусть случайная величина $X \sim N(a, \sigma)$; σ – известна, доверительная вероятность (надежность) γ – задана.

Случайные величины X_1, X_2, \dots, X_n – независимы (под X_i будем понимать значение X в i -м опыте), закон распределения любой из них совпадает с законом распределения X . А это значит, что $MX_1 = MX_2 = \dots = MX_n = MX = a$, $DX_1 = DX_2$

$$= \dots = DX_n = DX.$$

Выборочное среднее

$$\bar{X}_B = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

также будет распределено по нормальному закону. Параметры распределения \bar{X} таковы: $M(\bar{X}) = a$, $D(\bar{X}) = \frac{\sigma^2}{n}$. Действительно,

$$M(\bar{X}) = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n M X_i = \frac{1}{n} \sum_{i=1}^n M X = \frac{1}{n} \cdot n \cdot M X = a,$$

$$D(\bar{X}) = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D X_i = \frac{1}{n^2} \sum_{i=1}^n D X = \frac{1}{n^2} \cdot n \cdot D X = \frac{\sigma^2}{n}.$$

Таким образом, $\bar{X} \sim N\left(a, \frac{\sigma}{\sqrt{n}}\right)$.

Пользуясь формулой (2.25)

$$P\{|X - a| < l\} = \Phi_0\left(\frac{a+l-a}{\sigma}\right) - \Phi_0\left(\frac{a-l-a}{\sigma}\right) = 2\Phi_0\left(\frac{l}{\sigma}\right),$$

можно записать $\gamma = P\{|\bar{X} - a| < \varepsilon\} = 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 2\Phi_0(t)$, где $t = \frac{\varepsilon\sqrt{n}}{\sigma}$.

Из последнего равенства находим $\varepsilon = \frac{t \cdot \sigma}{\sqrt{n}}$. поэтому

$$\gamma = P\left\{|\bar{X} - a| < \frac{t \cdot \sigma}{\sqrt{n}}\right\} = 2\Phi_0(t) \text{ или}$$

$$P\left\{\bar{X} - t \cdot \frac{\sigma}{\sqrt{n}} < a < \bar{X} + t \cdot \frac{\sigma}{\sqrt{n}}\right\} = 2\Phi_0(t) = \gamma.$$

В соответствии с определением доверительного интервала получаем, что доверительный интервал для $\alpha = MX$ есть

$$\left(\bar{X} - t \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + t \cdot \frac{\sigma}{\sqrt{n}}\right),$$

где t определяется из уравнения

$$\Phi_0(t) = \frac{\gamma}{2} \text{ или } \Phi(t) = \frac{1+\gamma}{2}.$$

При заданном γ по таблице нормированной функции Лапласа $\Phi_0(t)$ или функции Лапласа $\Phi(t)$ находим аргумент t .

Видно, что с возрастанием объема выборки n число ε убывает и, значит, точность оценки увеличивается; увеличение надежности γ влечет уменьшение точности оценки.

Пример 6.4. Произведено 5 независимых наблюдений над случайной

величиной $X \sim N(a, 20)$. Результаты наблюдений: $x_1 = -25, x_2 = 34, x_3 = -20, x_4 = 10, x_5 = 21$. Найти оценку для $a = MX$ и построить для него 95% доверительный интервал.

Решение.

$$\bar{x} = \frac{1}{5}(-25 + 34 - 20 + 10 + 21) = 4$$

$$\Phi_0(t) = \frac{0.95}{2} = 0.475$$

Из таблицы для нормированной функции Лапласа находим $t=t_\gamma=1.96$.

Тогда $\varepsilon = \frac{1.96 \cdot 20}{\sqrt{5}} \approx 17.5$ и доверительный интервал для α таков

$$(4-17.5; 4+17.5), \text{ т.е. } (-13.5; 21.5)$$

Доверительный интервал для математического ожидания при неизвестной дисперсии

Пусть случайная величина $X \sim N(a, \sigma)$; σ – неизвестна, γ – задана.

Найдем такое число ε , чтобы выполнялось соотношение

$$P\{|\bar{X} - a| < \varepsilon\} = \gamma \quad (6.1)$$

Введем случайную величину

$$T = \frac{\bar{X} - a}{\frac{S}{\sqrt{n}}}$$

где S – исправленное среднее квадратическое отклонение случайной величины X , вычисленное по выборке:

$$S = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \bar{X})^2}$$

Доказывается, что случайная величина T имеет распределение Стьюдента с $n-1$ степенью свободы.

Перейдем в левой части равенства (6.1) от случайной величины \bar{X} к случайной величине T

$$P\{|T| < \frac{\varepsilon \sqrt{n}}{S}\} = \gamma \text{ или } P\{|T| < t_\gamma\} = \gamma,$$

где

$$t_\gamma = \frac{\varepsilon \cdot \sqrt{n}}{S}$$

Тогда искомый доверительный интервал будет иметь вид

$$\left(\bar{X} - t_\gamma \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_\gamma \cdot \frac{S}{\sqrt{n}}\right)$$

Величина t_γ (квантиль уровня $1 - \gamma$) находится из равенства

$$2 \int_0^{t_\gamma} f_T(t, n-1) dt = \gamma.$$

где $f_T(t, n-1)$ – функция плотности t -распределения Стьюдента.

Величина t_γ находится из таблиц квантилей распределения Стьюдента, по доверительной вероятности γ и числу степеней свободы $n-1$.

Историческая справка



t -распределение – это распределение вероятностей, связанное с нормальным распределением. Возникает оно, когда требуется оценить среднее статистической выборки, когда размер выборки, используемой для оценки, мал и дисперсии неизвестны. Определяется это распределение случайной величины следующим образом:

$$t = \frac{Z}{\sqrt{\chi^2/n}},$$

где $Z \rightsquigarrow N(0, 1)$, то есть нормальное распределение с математическим ожиданием 0 и стандартным отклонением 1, χ^2 – распределение хи-квадрат (еще одно распределение вероятностей, связанное с нормальным), n – число степеней свободы распределения хи-квадрат. История развития этого распределения вероятностей, как мы уже говорили в начале, весьма любопытна. Уильям Сили Госсет – английский математик и химик, который после окончания университета начал работать на заводе Guinness (знаменитого пива “Гиннесс”), занимаясь контролем качества в процессе создания пива. Малый размер выборки, с которой обычно приходилось иметь дело, был виновником его занятий, которые в конечном счете привели к разработке t -распределения. В 1908 году, когда ему было 32 года, Госсет опубликовал статью The probable error of a mean (“Вероятная ошибка среднего”) в журнале Biometrika, но не под своим именем, а под псевдонимом Стьюдент.

Почему был избран псевдоним? Как это часто бывает в таких случаях, существует несколько теорий, пытающихся объяснить это. Первая и, видимо, наиболее распространенная, говорит, что основная причина в том, что Guinness ранее нес ущерб от утечки информации из-за публикаций сотрудников, что компания запретила своим сотрудникам публиковать статьи, независимо от их темы. Эта история имеет не одно продолжение, которое отличается в различных источниках. Некоторые говорят, что публикация Госсета под псевдонимом Стьюдент позволила скрыть от компании Guinness, что ее сотрудник опубликовал статью. Другие говорят, что Госсет договорился с пивоварней о ее публикации (однако содержание статьи не было бы полезно для конкуренции), но компания попросила использовать псевдоним, чтобы другие сотрудники не были осведомлены об этой публикации.

Пример 6.5. По условию примера 6.4, считая, что случайная величина $X \sim N(a, \sigma)$, построить для неизвестного $MX = a$ доверительный интервал. Считать $\gamma = 0,95$.

Решение. Оценку \bar{x} для MX уже знаем: $\bar{x} = 4$. Находим значение S .

$$S = \frac{1}{2} \sqrt{(-25 - 4)^2 + (34 - 4)^2 + (-20 - 4)^2 + (10 - 4)^2 + (21 - 4)^2} \approx 25.7$$

По таблице квантилей t -распределения Стьюдента (см. **Приложения**) для $\alpha = 1 - \gamma = 1 - 0.95 = 0.05$ и $k = n - 1 = 5 - 1 = 4$ находим $t_\gamma = 2.78$. Следовательно,

$$\varepsilon = 2.78 \cdot \frac{25.7}{2.24} \approx 31.9. \text{ Искомый доверительный интервал таков:}$$

$$4 - 31.9; 4 + 31.9 = (-27.9; 35.9).$$

Доверительный интервал для среднего квадратического отклонения нормального распределения

Пусть случайная величина $X \sim N(a, \sigma)$; σ – неизвестна, γ – задана.

Можно показать, что если $MX = a$ неизвестно, то доверительный интервал для неизвестного σ имеет вид:

$$\left(\frac{\sqrt{n-1} \cdot S}{\chi_2}, \frac{\sqrt{n-1} \cdot S}{\chi_1} \right),$$

где n – объем выборки, $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$ – исправленное среднее квадратическое отклонение, квантили $\chi_1^2 = \chi_{\frac{1+\gamma}{2}; n-1}^2$, $\chi_2^2 = \chi_{\frac{1-\gamma}{2}; n-1}^2$ определяются по таблице квантилей χ^2 -распределения (см. **Приложения**) при $k = n-1$ и $\alpha_1 = \frac{1+\gamma}{2}$ и $\alpha_2 = \frac{1-\gamma}{2}$ соответственно.

Пример 6.6. Для оценки параметра нормально распределенной случайной величины была сделана выборка объема в 30 единиц и вычислено $S = 1,5$. Найти доверительный интервал, покрывающий σ с вероятностью $\gamma = 0,90$.

Решение. Имеем $n = 30$, $\gamma = 0.9$. По таблице квантилей χ^2 -распределения находим:

$$\chi_1^2 = \chi_{\frac{1+0.9}{2}; 30-1}^2 = \chi_{0.95; 29}^2 = 17.7; \chi_2^2 = \chi_{\frac{1-0.9}{2}; 30-1}^2 = \chi_{0.05; 29}^2 = 42.6$$

Доверительный интервал имеет вид:

$$\left(\frac{\sqrt{29} \cdot 1.5}{\sqrt{42.6}}; \frac{\sqrt{29} \cdot 1.5}{\sqrt{17.7}} \right) \text{ или } 1.238 < \sigma < 1.920.$$

7. Проверка статистических гипотез

Статистическая гипотеза. Статистический критерий

Процедура сопоставления высказанного предположения (гипотезы) с выборочными данными называется *проверкой гипотез*. Статистическими методами гипотезу можно только *опровергнуть* или *не опровергнуть*, но не доказать.

Статистические гипотезы делятся на гипотезы о параметрах распределения известного вида (это так называемые *параметрические* гипотезы) и гипотезы о виде неизвестного распределения (*непараметрические* гипотезы).

Нулевой (основной) называют выдвинутую гипотезу H_0 .

Конкурирующей (альтернативной) называют гипотезу H_1 , которая противоречит нулевой.

Простой называют гипотезу, содержащую только одно предположение, *сложной* – гипотезу, состоящую из конечного или бесконечного числа простых гипотез.

Основной прием проверки статистических гипотез заключается в том, что по имеющейся выборке вычисляется значение некоторой случайной величины, имеющей известный закон распределения.

Статистическим критерием называется случайная величина K с известным законом распределения, служащая для проверки нулевой гипотезы.

Критической областью называют область значений критерия, при которых нулевую гипотезу отвергают, *областью принятия гипотезы* – область значений критерия, при которых гипотезу принимают. Для ее построения задают α – *уровень значимости критерия* (рис. 4).

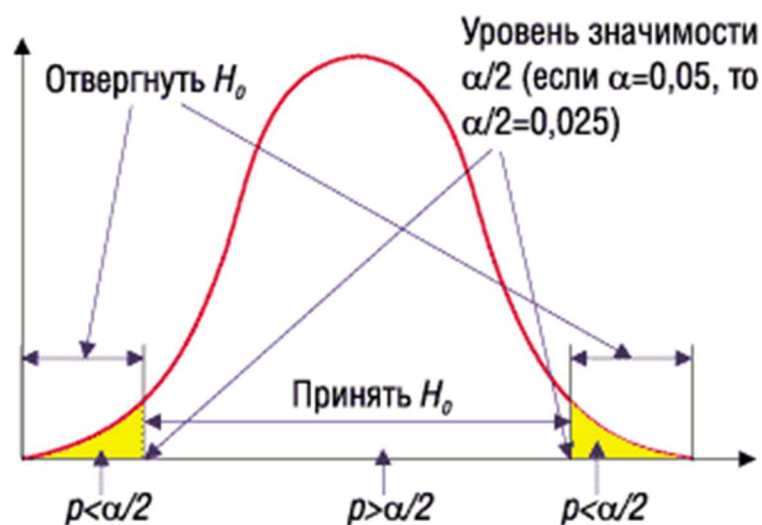


Рис. 4. Проверка статистической гипотезы

Типы критических областей (рис. 5-7):

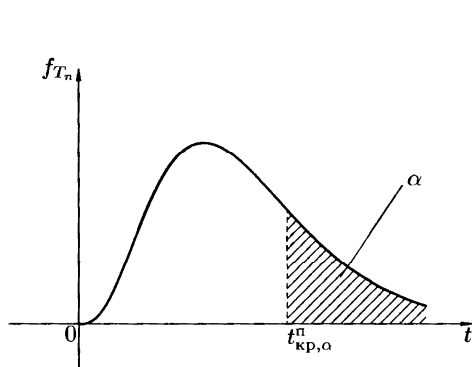


Рис 5. Правосторонняя критическая область S

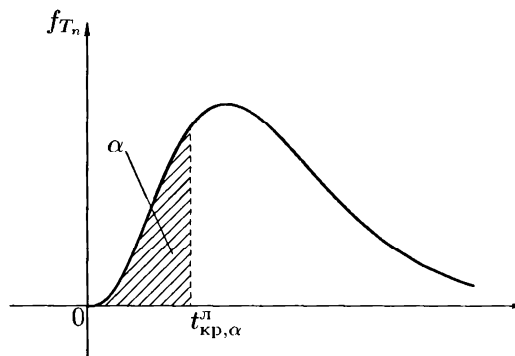


Рис 6. Левосторонняя критическая область S

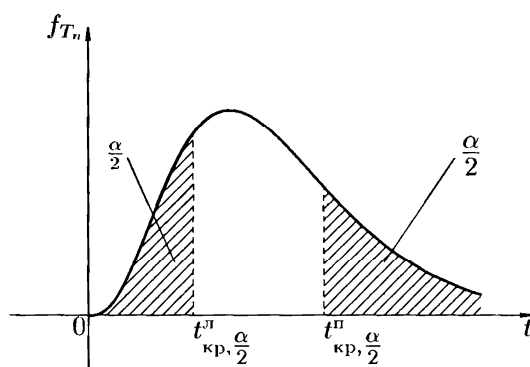


Рис 7. Двухсторонняя критическая область S

При проверке гипотезы может быть принято неправильное решение, т. е. могут быть допущены ошибки двух родов:

Ошибка первого рода состоит в том, что отвергается нулевая гипотеза H_0 , когда на самом деле она верна.

Ошибка второго рода состоит в том, что отвергается альтернативная гипотеза H_1 , когда она на самом деле верна.

Статистическое решение	Фактическая ситуация	
	Гипотеза H_0 верна	Гипотеза H_0 неверна
Гипотеза H_0 не отклоняется	Правильное решение (<i>true-negative</i>) Доверительная вероятность равна $1-\alpha$	Ошибка 2-рода (<i>false-negative</i>) Вероятность ошибки равна β
Гипотеза H_0 отклоняется	Ошибка 1-рода (<i>false-positive</i>) Вероятность ошибки равна α	Правильное решение (<i>true-positive</i>) Мощность критерия равна $1-\beta$

Вероятность ошибки 1-го рода (обозначается через α) называется *уровнем значимости критерия*.

Очевидно, $\alpha = P(H_1|H_0)$. Чем меньше, α , тем меньше вероятность отклонить верную гипотезу. Допустимую ошибку 1-го рода обычно задают заранее.

Обычно, используются стандартные значения: $\alpha = 0,05$; $\alpha = 0,01$; $0,005$; $0,001$.

Вероятность ошибки 2-го рода обозначается через β , т.е. $\beta = P(H_0|H_1)$.

Последствия ошибок 1-го, 2-го рода могут быть совершенно различными: в одних случаях надо минимизировать α , в другом – β .

Какая из ошибок является на практике более опасной, зависит от конкретной задачи. Например, если проверяется правильность выбора метода лечения больного, то ошибка первого рода означает отказ от правильной методики, что может замедлить лечение, а ошибка второго рода (применение неправильной методики) чревата ухудшением состояния больного и является более опасной.

Величину $1 - \beta$, т. е. вероятность недопущения ошибки 2-го рода (отвергнуть неверную гипотезу H_0 принять верную H_1), называется *мощностью критерия* (рис. 8).

$$1 - \beta = P(H_1|H_1) = P((x_1, x_2, \dots, x_n) \in S|H_1)$$

Чем больше мощность критерия, тем вероятность ошибки 2-го рода меньше, что, конечно, желательно (как и уменьшение α). Поэтому после выбора уровня значимости следует строить критическую область так, чтобы мощность критерия была максимальной.

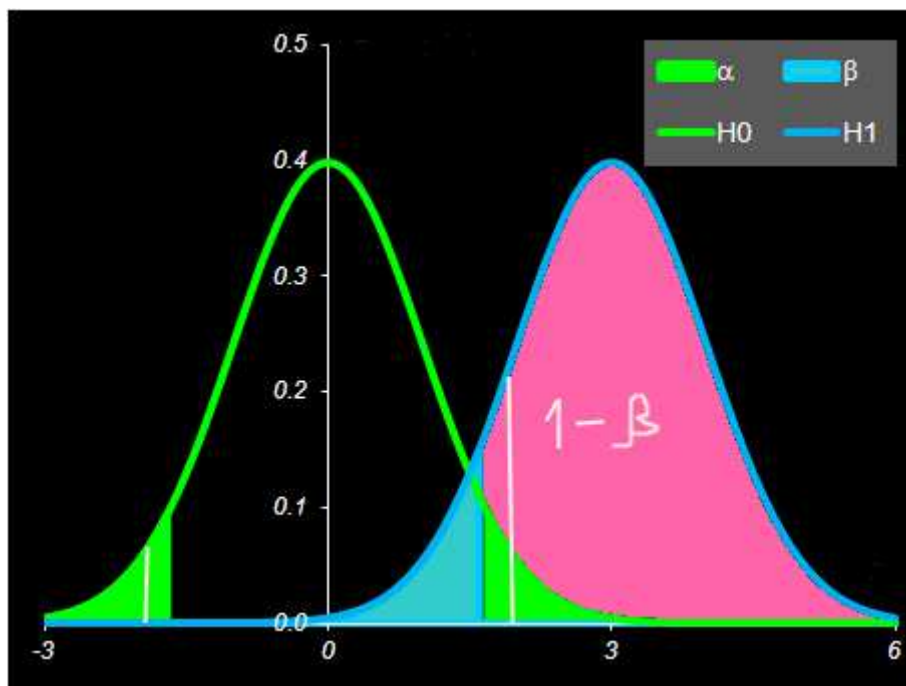
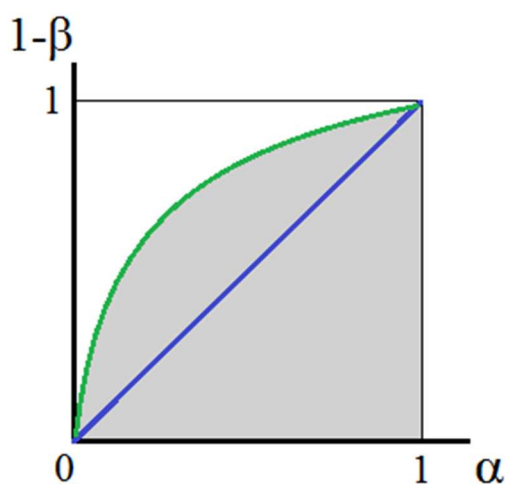


Рис. 8. Пример проверки гипотезы (двухсторонняя критическая область)

Отметим, что *одновременное уменьшение ошибок 1-го и 2-го рода возможно лишь при увеличении объема выборок*. Поэтому обычно при заданном уровне значимости α отыскивается критерий с наибольшей мощностью.

Сейчас большинство статистических тестов, реализованных в виде компьютерных программ, выдают в качестве результата не только рассчитанное по выборке значение статистики, но также предельный (наименьший) уровень значимости α , при котором нулевая гипотеза отвергается в пользу альтернативной. Эту величину называют “достигаемым уровнем значимости” (*p-value*, *p-значение*). *ROC-кривая* (*кривая ошибок, receiver operating characteristic*) показывает зависимость мощности ($1-\beta$) статистического критерия от уровня значимости α при изменении *p-значения*, т.е. она позволяет оценить качество теста (рис. 9). Начало координат соответствует нулевому значению *p-значения*; при его увеличении мы двигаемся по ROC-кривой вправо и вверх. Итак, *p-значение* – это наименьший уровень значимости, при котором нулевая гипотеза отвергается для данного значения статистики критерия.



$1-\beta$ – мощность критерия (вероятность правильного отбрасывания)
 α – ошибка 1-го рода (вероятность ложного отбрасывания)

Диагональ этого квадрата, выходящая из начала координат, соответствует абсолютно бесполезному тесту, который эквивалентен подбрасыванию монетки: при любом выборе *p-значения* мощность теста оказывается равной вероятности совершить ошибку первого рода.

Рис.9. Оперативная характеристика критерия (*ROC-кривая*)

ROC-кривые полезных тестов лежат левее и выше этой диагонали. Правее и ниже лежат *ROC-кривые* тестов, которые хуже монетки (они станут лучше, если принимать решение “наоборот”).

Из рис. 9 следует:

Если $\alpha=0$, то $1-\beta=0$ и $\beta=1$ и будет всегда совершаться ошибка 2-рода.

Если $\alpha=1$, то $1-\beta=1$ и $\beta=0$ и будет всегда совершаться ошибка 1-рода.

Видно, что одновременно уменьшить α и β при фиксированном объеме выборки невозможно.

Итак, процесс проверки гипотезы состоит из следующих этапов:

1. Выбирается статистический критерий *K*.

2. Вычисляется его наблюдаемое значение $K_{\text{набл}}$ по имеющейся выборке.
3. Поскольку закон распределения K известен, определяется (по известному уровню значимости α) критическое значение $k_{\text{кр}}$, разделяющее критическую область и область принятия гипотезы.
4. Если вычисленное значение $K_{\text{набл}}$ попадает в область принятия гипотезы, то нулевая гипотеза принимается, если в критическую область – нулевая гипотеза отвергается.

Проверка гипотезы о равенстве центров распределения нормальных генеральных совокупностей при неизвестном σ

Рассмотрим две случайные величины $X \sim N(m_X, \sigma_X)$ и $Y \sim N(m_Y, \sigma_Y)$. Будем считать, что $\sigma_X = \sigma_Y = \sigma$.

Пусть имеются две независимые выборки объемом n_1 и n_2 соответственно из генеральных совокупностей X и Y .

Необходимо проверить нулевую гипотезу $H_0: m_X = m_Y$ относительно альтернативной гипотезы $H_1: |m_X - m_Y| > 0$

Для оценки m_X и m_Y используем их наилучшие оценки по выборке \bar{X} и \bar{Y} , а для оценки σ^2

$$S_X^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{X})^2 \qquad S_Y^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{Y})^2$$

Доказано, что лучшей оценкой для σ^2 в данном случае является

$$S^2 = \frac{S_X^2 \cdot (n_1 - 1) + S_Y^2 \cdot (n_2 - 1)}{n_1 + n_2 - 2}$$

Если случайная величина $(\bar{X} - \bar{Y})$ подчиняется нормальному закону и гипотеза H_0 справедлива, то статистика t имеет t -распределение Стьюдента и ее можно записать в виде

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2}}}$$

Выбрав уровень значимости α критерия по таблице t -распределения Стьюдента можно определить критическое значение $t_{n_1+n_2-2, \alpha}$, для которого

$$P(|t| > t_{n_1+n_2-2, \alpha}) = \alpha$$

Если вычисленное значение $|t| > t_{n_1+n_2-2, \alpha}$, то гипотеза H_0 отвергается, т.е. можно считать с надежностью $p=1-\alpha$ можно считать расхождение средних значимым (неслучайным).

Пример 7.1. В результате двух серий измерений с количеством измерений $n_1=25$ и $n_2=50$ получены следующие средние значения исследуемой величины: $\bar{X} = 9.79$ и $\bar{Y} = 9.60$. Можно ли с надежностью $p=0.99$ объяснить это расхождение случайными причинами, если средние квадратические отклонения в обоих случаях измерений неизвестны, но равны. По выборкам найдены их оценки $S_x^2 = 0.0784$, $S_y^2 = 0.1089$.

Решение. Находим наблюдаемое значение критерия t

$$t = \frac{9.79 - 9.60}{\sqrt{\left(\frac{1}{25} + \frac{1}{50}\right) \cdot \frac{(25-1) \cdot 0.0784 + (50-1) \cdot 0.1089}{25 + 50 - 2}}} = 2.47$$

Уровню значимости $\alpha=1-p=1-0.99=0.01$ и числу степеней свободы $k=25+50-2=73$ в таблице t -распределения соответствует $t_{73,0.01} = 2.65$.

Т.к. $2.47 < 2.65$, то с надежностью 0.99 нельзя считать расхождение средних значимым.

Проверка гипотез о законе распределения

Во многих случаях закон распределения изучаемой случайно величины неизвестен, но есть основания предположить, что он имеет вполне определенный вид: нормальный, биномиальный или какой-либо другой.

Пусть необходимо проверить гипотезу H_0 том, что случайная величина X подчиняется определенному закону распределения, заданному функцией распределения $F_0(x)$, т. е. $H_0: F_X(x) = F_0(x)$. Под альтернативной гипотезой H_1 будем понимать в данном случае то, что просто не выполнена основная ($H_0: F_X(x) \neq F_0(x)$).

Для проверки гипотезы о распределении случайной величины X проведем выборку, которую оформим в виде статистического ряда:

x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

(7.1)

Требуется сделать заключение: согласуются ли результаты наблюдений с высказанным предположением. Для этого используем специально подобранную величину – критерий согласия.

Критерием согласия называют статистический критерий проверки гипотезы о предполагаемом законе неизвестного распределения. Он используется для проверки согласия предполагаемого вида распределения с опытными данными на основании выборки.

Существуют различные критерии согласия: Пирсона, Колмогорова, Фишера, Смирнова и др.

Критерий согласия Пирсона – наиболее часто употребляемый критерий для

проверки простой гипотезы о законе распределения.

Критерий χ^2 Пирсона

Для проверки гипотезы H_0 поступают следующим образом.

Разбивают всю область значений случайной величины X на m интервалов $\Delta_1, \Delta_2, \dots, \Delta_m$ и подсчитывают вероятности p_i ($i = 1, 2, \dots, m$) попадания случайной величины X в интервал Δ_i , используя формулу $P\{\alpha \leq X \leq \beta\} = F_0(\beta) - F_0(\alpha)$. Тогда теоретическое число значений случайной величины X , попавших в интервал Δ_i можно рассчитать по формуле $n \cdot p_i$. Таким образом, имеем статистический ряд распределения случайной величины X (7.1) и теоретический ряд распределения:

Δ_1	Δ_2	...	Δ_m
$n'_1 = np_1$	$n'_2 = np_2$...	$n'_m = np_m$

Если эмпирические частоты n_i сильно отличаются от теоретических $np_i = n'_i$, то проверяемую гипотезу H_0 следует отвергнуть; в противном случае – принять.

В качестве меры расхождения К. Пирсон предложил величину (критерий Пирсона)

$$\chi^2 = \sum_{i=1}^n \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^n \frac{n_i^2}{np_i} - n. \quad (7.2)$$

Согласно теореме Пирсона, при $n \rightarrow \infty$ статистика (7.2) имеет χ^2 -распределение с $k = m - r - 1$ степенями свободы, где m – число групп (интервалов) выборки, r – число параметров предполагаемого распределения. В частности, если предполагаемое распределение нормально, то оценивают два параметра (α и σ), поэтому число степеней свободы $k = m - 3$.

Правило применения критерия χ^2 сводится к следующему:

1. По формуле (7.2) вычисляют $\chi^2_{\text{набл}}$ – выборочное значение статистики критерия.

2. Выбрав уровень значимости α критерия, по таблице χ^2 -распределения находим критическую точку (квантиль) $\chi^2_{\alpha, k}$.

3. Если $\chi^2_{\text{набл}} \leq \chi^2_{\alpha, k}$, то гипотеза H_0 не противоречит опытным данным, в противном случае гипотеза H_0 отвергается.

Необходимым условием применения критерия Пирсона является наличие в каждом из интервалов не менее 5 наблюдений. Если в отдельных интервалах их меньше, то число интервалов надо уменьшить путем объединения (укрупнения) соседних интервалов.

Пример 7.2. Измерены 100 обработанных деталей. Отклонения от заданного

размера приведены в таблице:

$[x_i, x_{i+1})$	$[-3, -2)$	$[-2, -1)$	$[-1, -0)$	$[0, 1)$	$[1, 2)$	$[2, 3)$	$[3, 4)$	$[4, 5)$
n_i	3	10	15	24	25	13	7	3

Проверить при уровне значимости $\alpha=0.01$ гипотезу H_0 о том, что отклонения от проектного размера подчиняются нормальному закону.

Решение.

Число наблюдений в крайних интервалах меньше 5, поэтому объединим их с соседними.

$[x_i, x_{i+1})$	$[-3, -1)$	$[-1, -0)$	$[0, 1)$	$[1, 2)$	$[2, 3)$	$[3, 5)$
n_i	13	15	24	25	13	10

Случайную величину – отклонение – обозначим через X . Для вычисления вероятностей p_i необходимо вычислить параметры, определяющие нормальный закон распределения (a и σ). Их оценки вычислим по выборке.

$$\bar{x} = \frac{1}{100}(-2 \cdot 13 + (-0.5) \cdot 15 + \dots + 4 \cdot 10) = 0.885 \approx 0.9$$

$$D_B = \frac{1}{100}(4 \cdot 13 + 0.25 \cdot 15 + \dots + 16 \cdot 10) - (0.885)^2 \approx 2.89, \quad \sigma \approx 1.676 \approx 1.7.$$

Находим p_i . ($i = \overline{1, 6}$). Так случайная величина $X \sim N(a, \sigma)$ определена на $(-\infty, \infty)$, то крайние интервалы в ряде распределения заменяем на $(-\infty, -1)$ и $(3, +\infty)$

$$\text{Тогда } p_1 = P\{-\infty < X < -1\} = \Phi_0\left(\frac{-1-0.9}{1.7}\right) - \Phi_0(-\infty) = \frac{1}{2} - \Phi_0(1.12) = 0.1314.$$

Аналогично получаем: $p_2 = 0.1667, p_3 = 0.2258, p_4 = 0.2183, p_5 = 0.1503.$

$$p_6 = P\{3 < X < \infty\} = \Phi_0(\infty) - \Phi_0\left(\frac{3-0.9}{1.7}\right) = \frac{1}{2} - \Phi_0(1.24) = 0.1075.$$

Полученные результаты приведем в следующей таблице:

$[x_i, x_{i+1})$	$[-\infty, -1)$	$[-1, -0)$	$[0, 1)$	$[1, 2)$	$[2, 3)$	$[3, \infty)$
n_i	13	15	24	25	13	10
$n'_i - np_i$	13.14	16.67	22.58	21.83	15.03	10.75

Вычисляем $\chi^2_{\text{набл}}$

$$\chi^2_{\text{набл}} = \sum_{i=1}^6 \frac{n_i^2}{np_i} - n = \left(\frac{13^2}{13.14} + \frac{15^2}{16.67} + \dots + \frac{10^2}{10.75}\right) - 100 = 101.45 - 100 \approx 1.045.$$

Находим число степеней свободы. По выборке рассчитаны два параметра, т.е. $r = 2$. Количество интервалов 6, т.е. $m = 6$. Следовательно, $k = 6 - 2 - 1 = 3$. По таблице χ^2 -распределения для $\alpha=0.01$ находим $\chi^2_{\alpha, k} = 11.3$. Итак, $\chi^2_{\text{набл}} < \chi^2_{\alpha, k}$,

следовательно нет оснований отвергнуть проверяемую гипотезу.

Критерий Колмогорова

Критерий Колмогорова для простой гипотезы является наиболее простым критерием проверки гипотезы о виде закона распределения. Он связывает эмпирическую функцию распределения $F_n^*(x)$ с функцией распределения $F(x)$ непрерывной случайной величины X .

Пусть x_1, x_2, \dots, x_n – конкретная выборка из распределения с неизвестной функцией распределения $F(x)$ и $F_n^*(x)$ – эмпирическая функция распределения. Выдвигается простая гипотеза $H_0: F(x) = F_n^*(x)$.

Сущность критерия Колмогорова состоит в том, что вводят в рассмотрение функцию

$$D_n = \max_{-\infty < x < \infty} |F_n^*(x) - F_0(x)|, \quad (7.3)$$

называемой статистикой Колмогорова, представляющей собой максимальное отклонение эмпирической функции распределения $F_n^*(x)$ от гипотетической (т. е. соответствующей теоретической) функции распределения $F_0(x)$. Колмогоров доказал, что при $n \rightarrow \infty$ закон распределения случайной величины $\sqrt{n} \cdot D_n$ независимо от вида распределения случайная величина X стремится к закону распределения Колмогорова:

$$P\{\sqrt{n} \cdot D_n < x\} \rightarrow K(x),$$

где $K(x)$ – функция распределения Колмогорова, для которой составлена таблица, ее можно использовать для расчетов уже при $n \geq 20$:

α	0.1	0.05	0.02	0.01	0.001
x_0	1.224	1.358	1.520	1.627	1.950

Найдем D_0 такое, что $P(D_n > D_0) = \alpha$.

Рассмотрим уравнение $K(x) = 1 - \alpha$. С помощью функции Колмогорова найдем корень x_0 этого уравнения. Тогда по теореме Колмогорова

$$P\{\sqrt{n} \cdot D_n < x_0\} = 1 - \alpha, \quad P\{\sqrt{n} \cdot D_n > x_0\} = \alpha, \quad \text{откуда } D_0 = \frac{x_0}{\sqrt{n}}.$$

Если $D_n < D_0$, то гипотезу H_0 нет оснований отвергать; в противном случае – ее отвергают.

Пример 7.3. Монету бросали 4040 раз (Бюффон). Получили $n_1 = 2048$ выпадений герба и $n_2 = 1992$ выпадений решки. Проверить, используя, а) критерий Колмогорова и б) критерий Пирсона, согласуются ли эти данные с гипотезой H_0 о симметричности монеты ($\alpha = 0.05$).

Решение. Случайная величина X принимает два значения: $x_1 = -1$ (решка) и

$x_2 = 1$ (герб). Гипотеза $H_0: P\{X = -1\} = P\{X = 1\} = \frac{1}{2}$.

а) По таблице распределения Колмогорова находим корень уравнения $K(x) = 1 - \alpha$ при $\alpha = 0.05$. Следует $x_0 = 1.358$. Тогда $D_0 = \frac{x_0}{\sqrt{n}} = \frac{1.358}{\sqrt{4040}} \approx 0.021$.

Для нахождения по выборке D_n строим функции $F_0(x)$ и $F_n^*(x)$ и вычисляем величину $D_n = \max_{-\infty < x < \infty} |F_n^*(x) - F_0(x)|$.

x_i	решка $x_1 = -1$	герб $x_2 = 1$
p_i	0.5	0.5

x_i	решка	герб
n_i	1992	2048
p_i^*	0.493	0.507

$$F_0(x) = \begin{cases} 0, & \text{при } x \leq -1, \\ 0.5, & \text{при } -1 < x \leq 1, \\ 1, & \text{при } 1 < x. \end{cases} \quad F_n^*(x) = \begin{cases} 0, & \text{при } x \leq -1, \\ 0.493, & \text{при } -1 < x \leq 1, \\ 1, & \text{при } 1 < x. \end{cases}$$

Максимальное отклонение $F_0(x)$ от $F_n^*(x)$ равно 0.007, т.е. $D_n = 0.007$. Поскольку $D_n < D_0$, то нет оснований отвергать гипотезу H_0 . Опытные данные согласуются с гипотезой H_0 о симметричности монеты.

б) Вычисляем статистику χ^2

$$\chi_{\text{набл}}^2 = \sum_{i=1}^2 \frac{n_i^2}{np_i} - n = \frac{1992^2}{\frac{1}{2} \cdot 4040} + \frac{2048^2}{\frac{1}{2} \cdot 4040} - 4040 = 0.776.$$

По таблице χ^2 -распределения находим $\chi_{0.05,1}^2 = 3.8$.

$\chi_{\text{набл}}^2 < \chi_{\alpha,k}^2$, следовательно нет оснований отвергнуть проверяемую гипотезу о симметричности монеты.

Приложения

Значение функции $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

Сотые доли x										
x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0112	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3079	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3553	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1,6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4700	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4762	4767
Десятые доли x										
x	0	1	2	3	4	5	6	7	8	9
2,	4773	4821	4861	4893	4918	4938	4953	4965	4974	4981
3,	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000 ¹

Квантили t-распределения Стьюдента
(k – число степеней свободы)

k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	3,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,05	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Квантили $\chi_{\alpha,k}^2$ распределения χ_k^2
 (k – число степеней свободы)

k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,26
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,3	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Учебное издание

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Составитель
ШЕВЕЛЁВ Геннадий Ефимович

Научный редактор
*доктор технических наук,
профессор О.Г. Берестнева*


В авторской редакции

Компьютерная верстка *Г.Е. Шевелев*



Национальный исследовательский
Томский политехнический университет
Система менеджмента качества
Издательства Томского политехнического университета
сертифицирована в соответствии с требованиями ISO 9001:2008



ИЗДАТЕЛЬСТВО  ТПУ. 634050, г. Томск, пр. Ленина, 30.
Тел./факс: 8(3822)56-35-35, www.tpu.ru