

1. Введение¹

1.1 Предмет математической статистики

Математическая статистика — это раздел математики, посвященный методам сбора, анализа и обработки статистических данных для научных и практических целей.

Статистические данные представляют собой данные, полученные в результате обследования большого числа объектов или явлений (то есть, математическая статистика имеет дело с массовыми явлениями).

Методы анализа массовых явлений — предмет многих научных дисциплин; но только в том случае, когда для анализа привлекаются формальные (абстрактные) математические модели, эти методы становятся статистическими.

Математическая статистика подразделяется на две обширные области:

описательная статистика	аналитическая статистика (теория статистических выводов)
методы описания статистических данных, представления их в форме таблиц, распределений и пр.	обработка данных, полученных в ходе эксперимента, и формулировка выводов, имеющих прикладное значение для конкретной области человеческой деятельности. Теория статистических выводов тесно связана с другой математической наукой — теорией вероятностей и базируется на ее математическом аппарате

Трудно найти современную область научных исследований, где бы ни использовались методы математической статистики. В последнее время они нашли широкое применение в медицине, биологии, социологии, т. е. в областях, сравнительно недавно считавшихся далекими от математики.

1.2. Понятие о выборочном методе

Сплошное наблюдение, т.е. изучение всех членов совокупности, сначала кажется единственно возможным способом получения о ней достаточно точной информации. На самом деле это не всегда так. Рассмотрим некоторые примеры.

Пусть на заводе за день изготавливается большая партия лампочек. Не контролировать срок их службы, конечно, нельзя. Однако если это будет сделано в стенах завода в отношении каждой лампочки, то, получив полную картину о долговечности лампочек, мы их все выведем из строя и ни одна не дойдет до потребителя. С такой «проверкой», безусловно, нельзя согласиться. В аналогичных условиях находятся предприятия, производящие консервы, ткани, искусственные волокна, строительные материалы и т. д.

Сплошное наблюдение нецелесообразно не только в случаях, когда оно приводит к уничтожению всех подлежащих рассмотрению объектов. Например, при составлении баланса денежных доходов и расходов населения нашей страны, при планировании денежного обращения, розничного товарооборота, транспортных тарифов, при проведении мероприятий по повышению материального и культурного уровня жизни народа необходимы данные о бюджетах семей трудящихся. Сбор этих данных осуществляется статистическими органами. Один работник-статистик в состоянии вести ежедневные записи доходов, расходов, потребления и т. д. не более чем в 20—25 семьях одновременно. Для обследования только бюджетов трудящихся нашей страны понадобилось бы несколько миллионов работников. Кроме того, для обработки собранных данных необходимо большое число специалистов. Выделить такое количество рабочей силы практически невозможно и нецелесообразно.

Средством для получения необходимой информации в подобных случаях остается **несплошное** наблюдение. Широкое применение находит *выборочный метод исследования*.

Суть этого метода: если по результатам изучения сравнительно небольшой ее части можно получить с достаточной для практики достоверностью необходимую информацию о всей совокупности, то нет необходимости в сплошном наблюдении.

Часть объектов исследования, определенным образом избранная из более обширной совокупности, называется выборкой, а исходная совокупность, из которой взята выборка, — генеральной (основной) совокупностью.

Вся подлежащая изучению совокупность объектов называется *генеральной совокупностью*.

Та часть объектов, которая попала на исследование, называется *выборочной совокупностью* (или просто *выборкой*).

¹ Основой для данного текста является учебник “Мат. статистика для институтов физкультуры”

Важнейшая характеристика выборки — объем выборки, т. е. число элементов в ней.

Число элементов в генеральной совокупности называется *объемом генеральной совокупности* (обозначается N). Относительно N , как правило, делается предположение, что он бесконечно велик, т. е. выборка получается из бесконечной генеральной совокупности.

Число элементов в выборке называется *объемом выборки* (обозначается n).

1.3. Способы образования выборочной совокупности

Чтобы иметь право судить о генеральной совокупности по выборке, последняя должна быть образована случайно. Этого можно достичь различными способами.

Существуют различные виды выборок:

1. собственно-случайная;
2. механическая;
3. типическая;
4. серийная...

1. Члены генеральной совокупности можно предварительно занумеровать, а каждый номер записать на отдельной карточке. Отбирая наудачу после тщательного перемешивания из пачки таких карточек по одной карточке, получим выборочную совокупность любого нужного объема, которая называется *собственно-случайной*.

Номера на отобранных карточках укажут, какие члены генеральной совокупности попали в выборку. При этом возможны два принципиально различных способа отбора карточек в зависимости от того, возвращается или не возвращается обратно вынутая карточка после записи ее номера. Выборочная совокупность, образованная по первой схеме называется *собственно-случайной*, а *повторным отбором членов*, по второй — *собственно-случайной с бесповторным отбором членов*. Для краткости далее их будем часто называть соответственно *повторной* и *бесповторной* выборками.

Собственно-случайная бесповторная выборка образуется и в том случае, когда из тщательно перемешанной пачки сразу взято нужное число карточек.

Собственно-случайную выборку заданного объема n можно образовать и с помощью так называемых таблиц случайных чисел или генератора случайных чисел на компьютере.

При образовании собственно-случайной выборки каждый член генеральной совокупности с одинаковой вероятностью может попасть в выборку.

2. Выборка, в которую члены из генеральной совокупности отбираются через определенный интервал, называется *механической*.

Например, если объем выборки должен составлять 5% объема генеральной совокупности (5%-ная выборка), то отбирается ее каждый 20-й член, при 10%-ной выборке — каждый 10-й член генеральной совокупности и т.д. Механическую выборку можно образовать, если имеется определенный порядок следования членов генеральной совокупности, например, если они следуют друг за другом в определенной последовательности во времени. Именно так появляются готовые детали со станка, приборы с конвейера и т. п. При этом необходимо убедиться, что в следующих один за другим членах генеральной совокупности значения признака не изменяются с той же (или кратной ей) периодичностью, что и периодичность отбора элементов в выборку.

Пусть из продукции станка в выборку попадает каждая пятая деталь, а после каждой десятой детали рабочий производит смену (или заточку) режущего инструмента и подналадку станка. Эти операции рабочего направлены на улучшение качества деталей, износ режущего инструмента происходит более или менее равномерно. Следовательно, в выборочную совокупность попадут детали, на качество которых работа станка влияет в одну и ту же сторону, а показатели выборочной совокупности могут неправильно отразить соответствующие показатели генеральной совокупности.

3. Если из предварительно разбитой на непересекающиеся группы генеральной совокупности образовать собственно-случайные выборки из каждой группы (с повторным или бесповторным отбором членов), то отобранные элементы составят выборочную совокупность, которая называется *типической*.

Оказывается, что выборочная совокупность с большей достоверностью воспроизводит однородную генеральную совокупность. Качество изделий различных цехов, участков, станков и смен может оказаться существенно различным. Поэтому при изучении качества изделия, выпускаемых предприятием, целесообразно образовывать выборку не из общей массы изготавливаемой предприятием продукции, а из продукции отдельно каждого цеха, смены (ночной, дневной) и даже участка, станка, т. е. образовать типическую выборку.

4. Если генеральную совокупность предварительно разбить на непересекающиеся серии (группы), а

затем, рассматривая серии как элементы, образовать собственно-случайную выборку (с повторным или бесповторным отбором серий), то все члены отобранных серий составят выборочную совокупность, которая называется *серийной*.

Пример.

Предположим, что на заводе 150 станков (10 цехов по 15 станков) производят одинаковые изделия. Если в выборку отбирать изделия из тщательно перемешанной продукции всех 150 станков, то образуется собственно-случайная выборка. Но можно отбирать изделия отдельно из продукции первого, второго и т. д. станков. Тогда будет образована типическая выборка. Если же членами генеральной совокупности считать цехи и в каждом из цехов образовать повторную или бесповторную выборку, то вся отобранная продукция составит серийную выборку.

1.4. Статистическая совокупность и статистические признаки

Все объекты (элементы), составляющие генеральную совокупность, должны иметь хотя бы один общий признак, позволяющий классифицировать объекты, сравнивать их друг с другом (например, пол, возраст, спортивная квалификация и т.п.). Наличие общего признака является основой для образования статистической совокупности. Таким образом, статистическая совокупность представляет собой результаты описания или измерения общих признаков объектов исследования.

Предметом изучения в статистике естественно являются изменяющиеся (варьирующие) признаки, которые иногда называют статистическими признаками. Они делятся на

качественные	количественные	
признаки, которыми объект обладает либо не обладает. Они не поддаются непосредственному измерению (например, цвет волос, специализация, квалификация, национальность, территориальная принадлежность, образование и т.п.).	результаты подсчета или измерения. В соответствии с этим они делятся на	
	дискретные	непрерывные
	могут принимать лишь отдельные значения из некоторого ряда чисел, например количество прожитых лет, число попаданий и промахов при серии выстрелов.	могут принимать любые значения в определенном интервале. Например, время работы механизма, скорость движения и т. п.

Отдельные числовые значения варьирующего признака называются вариантами. Варианты принято обозначать строчными латинскими буквами из конца алфавита: x , y , z .

1.5. Общая схема выборочного эксперимента

По эмпирическим данным, представляющим собой выборку из некоторой генеральной совокупности, оцениваются параметры, позволяющие описать всю генеральную совокупность; определяется интервал, в котором с заданным уровнем доверия находится истинное значение оцениваемого параметра; а затем проверяются те или иные утверждения и делаются выводы о свойствах всей генеральной совокупности.

2. ЭМПИРИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ

2.1. Введение

Рассмотрим методы построения эмпирических распределений, т.е. распределений элементов выборки по значениям изучаемого признака.

Выборочные данные, полученные в ходе эксперимента, называются соответственно экспериментальными (эмпирическими) данными.

2.2. Вариационные ряды

Вариационный ряд – ранжированный в порядке возрастания или убывания ряд вариантов с соответствующими им весами (частотой, частотой ...). То есть вариационный ряд – двойной числовой ряд, показывающий, каким образом численные значения изучаемого признака связаны с их повторяемостью в выборке. Вариационные ряды имеют большое значение при статистической обработке экспериментальных данных, поскольку дают наглядное представление о характерных особенностях варьирования признака.

Вариационные ряды бывают двух типов: интервальные и безынтервальными.

В интервальном вариационном ряду частоты (или частоты), характеризующие повторяемость вариант

в выборке, распределяются по интервалам группировки. Интервальный вариационный ряд строится, если изучаемый признак варьирует непрерывно, но используется и для дискретно варьирующих признаков в тех случаях, когда признак варьирует в широких пределах.

В безынтервальном вариационном ряду частоты (или частости) распределяются непосредственно по значениям варьирующего признака. Для построения безынтервального вариационного ряда необходимо варианты выборки расположить в порядке возрастания или убывания (проранжировать) и затем подсчитать, сколько раз каждая из них встречается в выборке. Безынтервальный вариационный ряд применяется в тех случаях, когда исследуемый признак варьирует дискретно и слабо.

Пример:

Превышение разрешенной скорости движения (км/ч)	Кол-во нарушений
20-30	10
30-40	20
40-45	15
45-60	10
Больше 60	5

Признак – непрерывный

Зрение (диоптрии)	Кол-во человек
-10:-6	1
-6:-3	5
-3:-1	8
-1:+1	11
+1:+5	3
+5:+10	2

Признак дискретный, сильно варьирующий.

Экзаменационная оценка	Кол-во студентов
5	5
4	8
3	12
2	5

Признак дискретный, слабо варьирующий.

2.3. Табличное представление экспериментальных данных.

Как правило, необработанные (первичные) экспериментальные данные представлены в виде неупорядоченного набора чисел, записанных исследователем в порядке их поступления. Этот набор данных трудно обзрим, и сделать по ним какие-то выводы невозможно. Поэтому первичные данные нуждаются в обработке, которая всегда начинается с их группировки.

Группировка представляет собой процесс систематизации, или упорядочения, первичных данных с целью извлечения содержащейся в них информации. Группировка выполняется различными методами в зависимости от целей исследования, вида изучаемого признака и количества экспериментальных данных (объема выборки), но наиболее часто группировка сводится к представлению данных в виде статистических таблиц.

Группировка заключается в распределении вариантов выборки по группам, или интервалам группировки, каждый из которых содержит некоторый диапазон значений изучаемого признака.

Шаг 1

Первая задача, которую необходимо решить при группировке, состоит в том, чтобы разбить весь диапазон варьирования признака в выборке (между минимальной и максимальной вариантами выборки) на интервалы группировки. Эта задача требует определения числа интервалов группировки и ширины каждого из них. Обычно предпочтительны интервалы одинаковой ширины, а при выборе числа интервалов исходят из следующих соображений.

Группировка производится для того, чтобы построить эмпирическое распределение и сформировать с его помощью предположения о форме распределения изучаемого признака в генеральной совокупности, из которой взята выборка.

При увеличении числа интервалов группировки и, следовательно, при сужении каждого из них уменьшается число экспериментальных данных, попадающих в каждый интервал. Поскольку выборочные значения случайны, они случайным образом распределяются по интервалам группировки, поэтому картина эмпирического распределения будет содержать много случайных деталей, что мешает установить общие закономерности варьирования признака.

И наоборот, при чрезмерно широких интервалах группировки нельзя получить детальной картины распределения, поэтому возникает опасность упустить важные закономерные подробности формы распределения.

Поэтому вопрос о выборе числа и ширины интервалов группировки приходится решать в каждом конкретном случае исходя из целей исследования, объема выборки и степени варьирования признака в выборке. Однако приближенно число интервалов k можно оценить исходя только из объема выборки n .

Делается это одним из следующих способов:

- 1) по формуле Стерджеса $k = 1 + 3.32 \lg n$;
- 2) с помощью табл. 2.1,

Таблица 2.1

Выбор числа интервалов группировки

Объем выборки, n	Число интервалов, k
25—40	5—6
40—60	6—8
60—100	7—10
100—200	8—12
Больше 200	10—15

Шаг 2

Находим ширину каждого из интервалов (при этом все они будут одинаковой ширины) по следующей формуле:

$$h = \frac{x_{\max} - x_{\min}}{k - 1}, \quad (2.1)$$

где h — ширина интервалов; x_{\max} и x_{\min} — максимальная и минимальная варианты выборки (x_{\max} и x_{\min} находятся непосредственно по таблице исходных данных).

Шаг 3

Наметим границы интервалов группировки. Нижняя граница первого интервала выбирается так, чтобы минимальная варианта выборки x_{\min} попадала примерно в середину этого интервала. Обычно нижняя граница первого интервала определяется как

$$x_{H1} = x_{\min} - \frac{h}{2} \quad (2.2)$$

Прибавив к этой величине ширину интервала, найдем нижнюю границу второго интервала $x_{H2} = x_{H1} + h$. Это будет одновременно и верхняя граница x_{B1} предыдущего (первого) интервала.

Аналогично походим $x_{H3} = x_{H2} + h$ и т.д. для всех интервалов.

После того, как намечены границы всех интервалов, остается распределить по этим интервалам выборочные варианты. *Однако при этом возникает следующий вопрос: как поступать в тех случаях, если какая-либо из вариантов попадает точно на границу соседних интервалов группировки, т.е. варианта совпадает с нижней границей одного и верхней границей соседнего с ним интервала. Такие варианты могут быть с одинаковыми основаниями отнесены к любому из соседних интервалов. Этот выбор оставляется на усмотрение экспериментатора.*

Шаг 4

Вычислим срединные значения интервалов группировки x_i , которые отстоят от нижних границ на величину, равную половине ширины интервалов, т. е.

$$x_i = x_{Hi} + \frac{h}{2}, \quad (2.3)$$

где x_{Hi} — нижняя граница i -го интервала.

Шаг 5

Далее на основании первичных данных распределяем варианты выборки по интервалам группировки, то есть, подсчитываем повторяемость вариантов в каждом интервале. Получившиеся числа имеют в статистике определенное название. Числа, показывающие, сколько раз варианты, относящиеся к каждому интервалу группировки, встречаются в выборке, называются частотами интервалов.

Обозначим частоты символом n_i . Общая сумма всех частот всегда равна объему выборки n , что можно использовать для проверки правильности подсчетов.

Шаг 6.1

Вычисляем накопленную частоту интервала — это число, полученное последовательным суммированием частот в направлении от первого интервала к последнему, до того интервала включительно, для которого определяется накопленная частота. Накопленные частоты обозначим n_{xi} .

Шаг 6.2

Вычисляем относительную частоту интервала (отношение частоты к объему выборки). Обозначим частоты символом w_i .

$$w_i = \frac{n_i}{n}. \quad (2.4)$$

Они показывают (выражают) доли (удельные веса) членов совокупности с одинаковым значением признака (для дискретных рядов) или попадающие в один интервал (для непрерывных).

Шаг 6.3

Вычисляем относительные частоты. Накопленной относительной частотой (частотью) называется отношение накопленной частоты к объему выборки.

Обозначив накопленную частоту как F_i , получаем:

$$F_i = \frac{n_{xi}}{n} \quad (2.5)$$

Сумма всех частостей всегда равна 1.

Таблица 2.2

Табличное представление данных о результатах

Номер интервала i	Границы интервалов		Срединные значения x_i	Частоты n_i	Накопл. частоты n_{xi}	Частости w_i	Накопл. относит.ч астоты F_i
	x_{Hi}	x_{Bi}					

2.3. Графическое представление экспериментальных данных

Для повышения наглядности эмпирических распределений, используется их графическое представление. Наиболее распространенными способами графического представления являются гистограмма, полигон частот и полигон накопленных частот (кумулята).

2.3.1. Гистограмма

Гистограмма используется для графического представления распределений непрерывно варьирующих признаков и состоит из примыкающих друг к другу прямоугольников, как показано на рис. 2.1. Основание каждого прямоугольника равно ширине интервала группировки, а высота его такова, что **площадь** прямоугольника пропорциональна частоте (или частости) попадания в данный интервал. Если ряд безинтервальный, то ширина всех столбцов выбирается произвольной, но одинаковой. Таким образом, высоты прямоугольников должны быть пропорциональны величинам

$$p_i = \frac{n_i}{h_i}, \quad (2.6)$$

где n_i — частота i -го интервала группировки; h_i — ширина i -го интервала группировки.

На графике гистограммы основание прямоугольников откладывается по оси абсцисс (x), а высота — по оси ординат (y) прямоугольной системы координат.

Однако в тех случаях, когда ширина всех интервалов группировки одинакова, вид гистограммы не изменится, если по оси ординат откладывать не величины p_i , а частоты интервалов n_i .

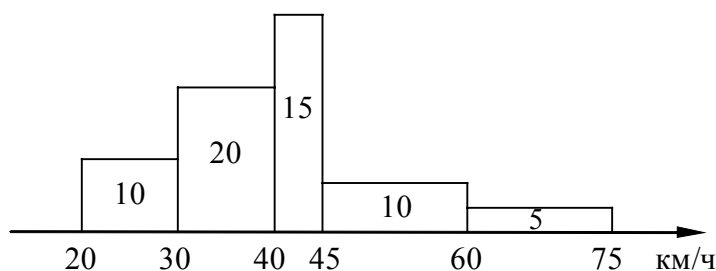


Рис. 2.1. Гистограмма распределения результатов в примере стр. 4 (когда ширина некоторых интервалов группировки неодинакова).

В этом случае чтобы не нарушить принцип построения гистограммы (площади прямоугольников пропорциональны частотам интервалов), по оси ординат уже нельзя откладывать частоты, а надо – высоты прямоугольников (которые должны быть пропорциональны отношениям $\frac{n_i}{h_i}$).

2.3.2. Полигон частот

Еще одним распространенным способом графического представления является полигон частот.

Полигон частот образуется ломаной линией, соединяющей точки, соответствующие срединным значениям интервалов группировки и частотам этих интервалов, срединные значения откладываются по оси x , а частоты – по оси y .

Из сравнения двух рассмотренных способов графического представления эмпирических распределений следует, что для получения полигона частот из построенной гистограммы (для случая моноширинного распределения) нужно середины вершин прямоугольников, образующих гистограмму, соединить отрезками прямых. Пример полигона частот представлен на рис. 2.2.

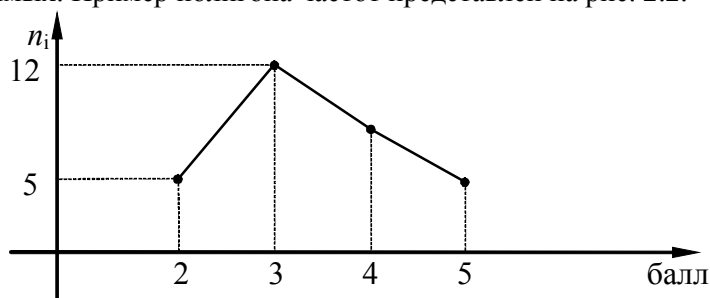


Рис. 2.2. Полигон частот в примере стр. 4 (для дискретного признака).

Полигон частот используется для представления распределений как непрерывных, так и дискретных признаков. В случае непрерывного распределения полигон частот является более предпочтительным способом графического представления, чем гистограмма, если график эмпирического распределения описывается плавной зависимостью.

2.3.3. Полигон накопленных частот

Полигон накопленных частот (кумулята) получается при соединении отрезками прямых точек, координаты которых соответствуют верхним границам интервалов группировки и накопленным частотам. Если по оси ординат откладывать накопленные частоты, то полученный график называется полигоном накопленных частотей. Пример полигона накопленных частот приведен на рис. 2.3.

На практике полигон накопленных частот используется в основном для представления дискретных данных. Ему свойственна более плавная форма, чем у гистограммы или полигона частот. Данное свойство и позволяет иногда отдавать предпочтение этому способу графического представлений эмпирических распределений.

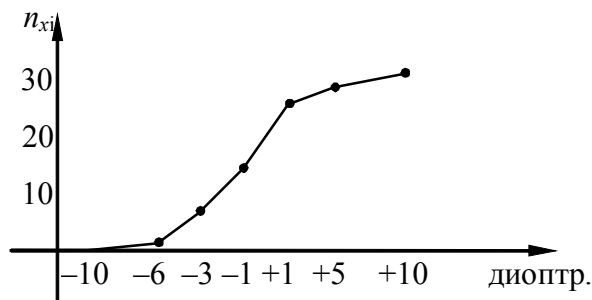


Рис. 2.3. Полигон накопленных частот в примере стр. 4.

2.3.4. Эмпирическая функция распределения

Эмпирической функцией распределения называется функция, вычисляемая для любого значения x по формуле

$$F(x) = \frac{n_x}{n},$$

где n – объем выборки, n_x – количество вариантов, значения которых меньше, чем x .

Свойства $F(x)$:

1. Если $x \leq x_{\min}$, то $F(x) = 0$;
2. Если $x > x_{\max}$, то $F(x) = 1$;
3. Если $x \in \mathbf{R}$, то $0 \leq F(x) \leq 1$;
4. $F(x)$ — функция неубывающая.

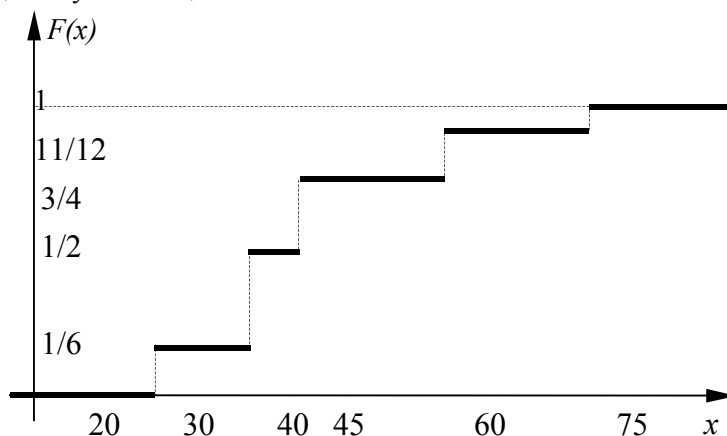


Рис. 2.4 График функции распределения

3. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ

3.1. Введение

Вариационные ряды и графики эмпирических распределений дают наглядное представление о том, как варьирует признак в выборочной совокупности. Но они недостаточны для полной характеристики выборки, поскольку содержат много деталей, охватить которые невозможно без применения обобщающих числовых характеристик.

Числовые характеристики выборки дают количественное представление об эмпирических данных и позволяют сравнивать их между собой. Наибольшее практическое значение имеют характеристики положения, рассеяния и асимметрии эмпирических распределений.

В этой главе рассматриваются характеристики положения и рассеяния, а также практические методы их вычисления. Характеристики асимметрии будут рассмотрены в гл. 6 применительно к проверке гипотез о виде распределения генеральной совокупности.

3.2. Характеристики положения

В этом разделе рассмотрены характеристики положения, определяющие положение центра эмпирического распределения. Чаще всего употребляются такие характеристики положения, как среднее арифметическое, медиана и мода.

3.2.1. Среднее арифметическое

Среднее арифметическое, или просто среднее, — одна из основных характеристик выборки.

Определение. Среднее арифметическое — такое значение признака, сумма отклонений от которого выборочных значений признака равна нулю (с учетом знака отклонения).

Если воспользоваться геометрической интерпретацией, то среднее арифметическое можно определить как точку на оси x , которая является абсциссой центра масс гистограммы.

Среднее принято обозначать той же буквой, что и варианты выборки, с той лишь разницей, что над буквой ставится символ усреднения — черта. Например, если обозначить исследуемый признак через X , а его числовые значения — через x_i , то среднее арифметическое имеет обозначение \bar{x} .

Среднее арифметическое, как и другие числовые характеристики выборки, может вычисляться как по необработанным первичным данным, так и по результатам группировки этих данных.

Для несгруппированных данных среднее арифметическое определяется по следующей формуле:

$$\bar{x}_B = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1)$$

где n — объем выборки; x_i — варианты выборки.

Если данные сгруппированы, то

$$\bar{x}_B = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n} = \frac{1}{n} \sum_{i=1}^k n_i x_i \quad (3.2)$$

где n — объем выборки; k — число интервалов группировки; n_i — частота i -ого интервала; x_i — срединное значение i -ого интервала.

Среднее арифметическое, вычисленное по формуле (3.2), называют также взвешенным средним, подчеркивая этим, что в формуле (3.2) числа x_i , суммируются с коэффициентами (весами), равными частотам попадания в интервалы группировки.

Среднее арифметическое измеряется в тех же единицах, что и значения признаков.

Нахождение среднего арифметического непрерывного вариационного ряда осложняется если крайние интервалы не замкнуты (то есть имеют вид “менее 10” “более 60”). В этом случае считается, что ширина первого интервала равна ширине второго, а ширина последнего — ширине предпоследнего.

3.2.2. Медиана

Определение. Медианой (Me) называется такое значение признака X , когда ровно половина значений экспериментальных данных меньше ее, а вторая половина — больше.

Собственно, этим и ограничивается смысловое значение медианы. Широкое использование этой характеристики на практике объясняется простотой ее вычисления и независимостью от формы распределения эмпирических данных.

Медиана обычно несколько отличается от среднего арифметического. Так бывает всегда, когда имеет место несимметричная форма эмпирического распределения.

Для тех случаев, когда эмпирическое распределение оказывается сильно асимметричным, среднее арифметическое теряет свою практическую ценность, поскольку при этом значительно большая часть значений признака оказывается выше или ниже среднего арифметического. В этой ситуации медиана представляет собой лучшую характеристику центра распределения.

Для дискретного ряда медиана находится по определению.

1. Если данных немного (объем выборки невелик), медиана вычисляется очень просто. Для этого выборку ранжируют, т. е. располагают данные в порядке возрастания или убывания, и в ранжированной выборке, содержащей n членов, ранг R (порядковый номер) медианы определяется как

$$R_{Me} = \frac{n+1}{2}.$$

Пусть, например, имеется ранжированная выборка, содержащая нечетное число членов $n = 9$:

12 14 14 18 20 22 22 26 28.

Тогда ранг медианы

$$R_{Me} = \frac{9+1}{2} = 5,$$

и медиана, обозначаемая символом Me , совпадает с пятым членом ряда: $Me = 20$.

2. Если выборка содержит четное число членов, то медиана не может быть определена столь однозначно. Например, получен ряд из 10 членов:

6 8 10 12 14 16 18 20 22 24.

Ранг медианы оказывается равным

$$R_{Me} = \frac{10+1}{2} = 5,5.$$

Медианой в этом случае может быть любое число между 14 и 16 (5-м и 6-м членами ряда). Для определенности принято считать в качестве медианы среднее арифметическое этих значений, т. е.

$$Me = \frac{14+16}{2} = 15.$$

Если необходимо найти медиану для сгруппированных данных, то поступают следующим образом.

Вначале находят интервал группировки, в котором содержится медиана, путем подсчета накопленных частот или накопленных относительных частот. Медианным будет тот интервал, в котором накопленная частота впервые окажется больше $n/2$ (n — объем выборки) или накопленная относительная частота — больше 0,5. Внутри медианного интервала медиана определяется по следующей формуле:

$$Me = x_{Me_h} + h_{me} \frac{0,5n - n_{x_{Me-1}}}{n_{Me}}, \quad (3.3)$$

где x_{Me_h} — нижняя граница медианного интервала; $0,5n = \frac{n}{2}$ — половина объема выборки; h_{me} — ширина медианного интервала; $n_{x_{Me-1}}$ — накопленная частота интервала, предшествующего медианному, n_{Me} — частота медианного интервала.

3.2.3. Мода

Определение. Мода (Mo) представляет собой значение признака, встречающееся в выборке наиболее часто.

Ряд называется *унимодальным*, если в нем только одно модальное значение и *полимодальным* в противном случае. Для полимодального ряда значение моды не определяют.

1. Для дискретного ряда мода находится по определению. Например, если имеем следующие данные измерений кол-ва :

10, 15, 20, 22, 22, 23, 25, 25, 26, 26, 26, 27, 28, 28, 30, 30, 31

то чаще всего встречается (самым *модным* значением является) 26, значит данный ряд унимодальный и $Mo = 26$.

2. Интервал группировки с наибольшей частотой называется модальным. Например, в вариационных рядах, приведенных в разделе 2.2 (стр 4.) модальными интервалами является 30-40 (для превышение разрешенной скорости движения), -1:+1 (острота зрения).

Для определения моды в интервальном ряду используется следующая формула:

$$Mo = x_{Mon} + h \frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} - n_{Mo+1})}, \quad (3.4)$$

где x_{mon} — нижняя граница модального интервала; h — ширина интервала группировки; n_{Mo} — частота модального интервала; n_{Mo-1} — частота интервала, предшествующего модальному; n_{Mo+1} — частота интервала, следующего за модальным.

3.3. Характеристики рассеяния

Средние значения не дают полной информации о варьирующем признаке. Нетрудно представить себе два эмпирических распределения, у которых средние одинаковы, но при этом у одного из них значения признака рассеяны в узком диапазоне вокруг среднего, а у другого — в широком. Поэтому наряду со средними значениями вычисляют и характеристики рассеяния выборки. Рассмотрим наиболее употребительные из них.

3.3.1. Размах вариации

Определение. Размах вариации — разность между максимальной и минимальной вариантами выборки:

$$R = x_{\max} - x_{\min}$$

Как видим, размах вычисляется очень просто, и в этом его главное и единственное достоинство. Информативность этого показателя невелика. Можно привести очень много распределений, сильно отличающихся по форме, но имеющих одинаковый размах. Размах вариации используется иногда в практических исследованиях при малых (не более 10) объемах выборки. Например, по размаху вариации легко оценить, насколько различаются лучший и худший результаты в группе спортсменов. При больших объемах выборки к его использованию надо откоситься с осторожностью.

3.3.2. Дисперсия и стандартное отклонение

Дисперсия и стандартное отклонение являются важнейшими характеристиками рассеяния.

Определение. Дисперсией называется средний квадрат отклонения значений признака от среднего арифметического. Дисперсия, вычисляемая по выборочным данным, называется выборочной дисперсией и обозначается σ_B^2 .

Выборочную дисперсию вычисляют по приведенным ниже формулам:

Для несгруппированных данных

$$\sigma_B^2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^2}{n}. \quad (3.5)$$

В этой формуле $\sum_{i=1}^n (x_i - \bar{x}_B)^2$ — сумма квадратов отклонений значений признака x_i от среднего арифметического \bar{x} . Для получения среднего квадрата отклонений эта сумма поделена на объем выборки n .

Для сгруппированных в интервальный вариационный ряд данных:

$$\sigma_B^2 = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n}. \quad (3.6)$$

Здесь x_i — срединные значения интервалов группировки; $\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2$ — взвешенная сумма квадратов отклонений.

Размерность дисперсии не совпадает с единицами измерения варьирующего признака. Дисперсия измеряется в единицах измерения признака в квадрате.

Определение. Стандартным отклонением (или средним квадратическим отклонением) называется корень квадратный из дисперсии:

$$\sigma_B = \sqrt{\sigma_B^2}. \quad (3.7)$$

Размерность стандартного отклонения в отличие от размерности дисперсии совпадает с единицами измерения варьирующего признака, поэтому в практической статистике для того, чтобы охарактеризовать рассеяние признака используют обычно стандартное отклонение, а не дисперсию.

3.3.3. Коэффициент вариации

Стандартное отклонение выражается в тех же единицах измерения, что и характеризуемый им признак. Если требуется сравнить между собой степень варьирования признаков, выраженных в разных единицах измерения, возникают определенные неудобства. Пусть, например, результаты в беге на 100 м, показанные группой IX классов, имеют стандартное отклонение 0,9 сек (при среднем времени 14,8 сек), а исследование роста тех же учащихся показывает, что его стандартное отклонение составляет 6 см (при среднем росте 168 см). Какой из признаков варьирует сильнее? Очевидно, что только на основании сравнения стандартных отклонений на этот вопрос ответить нельзя. Требуется сопоставить стандартные отклонения со средними арифметическими этих признаков. Поэтому вводится относительный показатель

$$V = \frac{\sigma_B}{\bar{x}_B}, \quad (3.8)$$

называемый коэффициентом вариации.

Обычно он выражается в процентном отношении:

$$V = \frac{\sigma_B}{\bar{x}_B} \cdot 100\%.$$

Коэффициент вариации является относительной мерой рассеяния признака.

Коэффициент вариации используется и как показатель однородности выборочных наблюдений. Считается, что если коэффициент вариации не превышает 10 %, то выборку можно считать однородной, т. е. полученной из одной генеральной совокупности.

Однако к использованию коэффициента вариации нужно подходить с осторожностью. Продемонстрируем возможные ошибки на следующем примере.

Если на основании многолетних наблюдений среднее арифметическое среднесуточных температур 8 марта составляет в какой-либо местности 0°C, то по формуле (3.8) получим бесконечный коэффициент вариации независимо от разброса температур. Поэтому в данном случае коэффициент вариации не применим! в качестве показателя рассеяния температур, а специфику явления более объективно оценивает стандартное отклонение 5.

Коэффициент вариации можно использовать как относительную меру рассеяния только в тех случаях, когда значения признака измерены в шкале с абсолютным нулем.

Практически коэффициент вариации применяется в основном для сравнения выборок из однотипных генеральных совокупностей.

3.3.4. Коэффициент осцилляции

С целью, аналогичной введению коэффициента вариации, вводится коэффициент осцилляции по

$$K_p = \frac{R}{\bar{x}_B} \cdot 100\%.$$

3.3.5. Коэффициент асимметрии и эксцесса

Выборочный центральный момент s -ого порядка вычисляется следующим образом:
для несгруппированных данных

$$\mu_s^* = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^s}{n}, \quad (3.9)$$

для сгруппированных в интервальный вариационный ряд данных:

$$\mu_s^* = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^s}{n}. \quad (3.10)$$

В частности, при $s=2$ второй центральный момент случайной величины есть дисперсия (сравни с формулами 3.5 и 3.6).

На практике используются третий и четвертый центральные моменты, позволяющие судить о симметричности и остроте вершины кривой распределения случайной величины.

Применяется так называемый коэффициент асимметрии, который является безразмерной величиной и определяется как

$$\gamma_3^* = \frac{\mu_3^*}{\sigma_B^3}. \quad (3.11)$$

Если $\gamma_3^* = 0$, то распределение симметрично относительно математического ожидания, если $\gamma_3^* > 0$, то преобладают положительные отклонения от математического ожидания, если $\gamma_3^* < 0$ — отрицательные.

Об остроте вершины кривой распределения судят по коэффициенту эксцесса:

$$\gamma_4^* = \frac{\mu_4^*}{\sigma_B^4} - 3. \quad (3.12)$$

Если $\gamma_4^* > 0$, то распределение имеет острый пик (по сравнению с нормальным распределением), если $\gamma_4^* < 0$ (минимальное значение $\gamma_4^* = -2$), то распределение имеет плосковершинную форму (по сравнению с нормальным распределением, для которого $\gamma_4 = 0$ см. 4.9.1).

4. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

4.1. Введение

В главах 2-3 были рассмотрены эмпирические распределения и методы вычисления их числовых характеристик. Но обработка экспериментальных данных не ограничивается рассмотренными методами. Обычно исследователь, получив данные эксперимента на одной или нескольких группах испытуемых и определив по ним некоторые обобщающие числовые характеристики (среднее, стандартное отклонение и др.), пытается найти ответ на следующие вопросы: насколько точно полученные результаты можно обобщить для более широкой совокупности (например, на всех жителей данного возраста)? Как хорошо его данные согласуются с данными других исследователей? Насколько достоверно различие экспериментальных данных, полученных в разных группах испытуемых или в одной и той же группе, но в разные промежутки времени? Существует ли связь между различными признаками, изучаемыми в проводимом исследовании, и если да, то насколько она сильна?

В ряде случаев исследователь пытается установить такую экспериментальную зависимость между изучаемыми признаками, чтобы по значениям одного из них, легко поддающегося измерению, установить значение другого, измерить который трудно или невозможно.

Конечно, в зависимости от целей конкретного исследования задачи могут быть различными и не ограничиваются приведенным перечнем.

Методы математической статистики, с помощью которых можно получить ответы на поставленные выше вопросы, рассматриваются в следующих главах. Чаще всего эти методы основаны на использовании тех или иных согласующихся с условиями проводимого эксперимента математических моделей, разработанных теорией вероятностей.

В данной главе рассматриваются некоторые ее элементарные, положения в том минимальном объеме, который необходим для дальнейшего изложения.

4.2. Определение вероятности

4.2.1. Испытание, событие, случайная величина

Под *испытанием* (*опытом*) в теории вероятностей принято понимать наблюдение какого-либо явления при соблюдении определенного комплекса условий, который должен каждый раз строго выполняться при повторении данного испытания. Если то же самое явление наблюдается при другом комплексе условий, то это уже другое испытание.

Когда речь идет о соблюдении комплекса условий данного испытания, имеется в виду постоянство значений всех факторов, контролируемых в данном испытании. Но при этом, как правило, имеет место большое число неконтролируемых факторов, которые трудно или невозможно учесть.

Результаты испытаний можно охарактеризовать качественно и количественно.

Качественная характеристика заключается в регистрации какого-либо явления, которое может наблюдаться или не наблюдаться при данном испытании. Любое из этих явлений называется в теории вероятностей *событием*.

События делятся на:

невозможные (в результате опыта никогда не произойдут),	достоверные (в результате опыта происходят всегда),	случайные (в результате опыта событие может произойти или не произойти).
---	---	--

Теория вероятностей рассматривает именно случайные события. При этом предполагается, что испытание может быть повторено неограниченное (по крайней мере, теоретически) число раз. Например, выполнение штрафного броска в баскетболе есть испытание, а попадание в кольцо — событие.

Другим примером события, часто приводимым в учебниках по теории вероятностей, является выпадение определенного числа очков (от 1 до 6) при бросании игральной кости.

События в теории вероятностей принято обозначать начальными прописными латинскими буквами A, B, C, ...

Случайные события называются *несовместными* если появление одного исключает появление другого. В противном случае они называются *совместными*.

Если в результате опыта произойдет хоть одно из некой группы событий, то они образуют *полную группу*. Появление хотя бы одного события из полной группы – достоверное событие.

Если, по условиям испытания нет никаких оснований предполагать, что один из исходов появляется чаще других, то все исходы являются *равновозможными*.

Два события называются *независимыми*, если появление одного из них не изменяет вероятности другого.

Количественная характеристика испытания состоит в определении значений некоторых величин, которыми интересуются при данном испытании (например, число подтягиваний на перекладине или время на беговой дистанции). В силу действия большого числа неконтролируемых факторов эти величины могут принимать различные значения в результате испытания. Причем до испытания невозможно предсказать значение величины, поэтому она называется *случайной величиной*.

4.2.2. Вероятность событий

Вероятность какого либо события – численное выражение возможности его наступления.

В некоторых простейших случаях вероятности событий могут быть легко определены непосредственно исходя из условий испытаний.

Представим себе общую схему таких испытаний.

Пусть испытание имеет n возможных несовместных исходов, т. е. отдельных событий, могущих появиться в результате данного испытания; причем при каждом повторении испытания возможен один и только один из этих исходов. Кроме того, пусть по условиям испытания, нет никаких оснований предполагать, что один из исходов появляется чаще других, т. е. все исходы являются равновозможными.

Допустим теперь, что при n равновозможных несовместных исходах интерес представляет некоторое событие A , появляющееся при каждом из m исходов и не появляющееся при остальных $n-m$ исходах. Тогда принято говорить, что в данном испытании имеется n случаев, из которых m благоприятствуют появлению события A .

Вероятность события A равна отношению числа исходов, благоприятствующих событию A , к общему числу всех равновозможных несовместных исходов опыта:

$$P(A) = \frac{m}{n}. \quad (4.1)$$

Формула (4.1) представляет собой так называемое *классическое определение вероятности* по Лапласу, пришедшее из области азартных игр, где теория вероятностей применялась для определения перспективы выигрыша.

Статистическое определение вероятности.

Будем фиксировать число испытаний, в результате которых появилось некоторое событие A . Пусть было проведено N испытаний, в результате которых событие A появилось ровно n_N раз. Тогда число n_N называется частотой события, а отношение $\frac{n_N}{N}$ — частотью (относительной частотой) события.

Замечательным экспериментальным фактом является то, что частота события при большом числе повторений испытания начинает мало изменяться и стабилизируется около некоторого определенного значения, в то время как при малом числе повторений она принимает различные, совершенно случайные значения. Поэтому интуитивно ясно, что если при неограниченном повторении испытания частота события будет стремиться к вполне определенному числовому значению, то это значение можно принять и качестве объективной характеристики события A . Такое число $P(A)$, связанное с событием A , называется вероятностью события A .

Математически неограниченное число повторений испытания записывается в виде предела (*lim*) при N , стремящемся к бесконечности (∞):

$$P(A) = \lim_{N \rightarrow \infty} \frac{n_N}{N}.$$

Поскольку n_N никогда не может превзойти N , то вероятность оказывается заключенной в интервале $0 \leq P(A) \leq 1$.

Следует отметить, что приведенное определение вероятности является абстрактным, оно не может быть экспериментально проверено, так как на практике нельзя реализовать бесконечно большое число повторений испытания.

Пусть проводятся независимые испытания, при каждом из которых вероятность события A неизменна. Справедливо утверждение, называемое *законом больших чисел* или *теоремой Бернулли*: если N достаточно

велико, то с вероятностью сколь угодно близкой к единице, отличие $\frac{n_N}{N}$ от $P(A)$ меньше любого наперед заданного положительного числа или, в символьной записи, $P\left(\left|\frac{n_N}{N} - P(A)\right| > \varepsilon\right) = 1$. Т.е. много раз бросая монету, мы “почти наверняка” будем получать примерно равные частоты выпадения герба и цифры.

4.3. Действия над событиями

В этом разделе приводятся основные правила операций над событиями с использованием для наглядности их графического изображения в виде диаграмм.

Вначале введем понятие “поле событий” как совокупности всех случайных событий данного испытания, для которых определены вероятности. На рис. 4.1 поле событий изображено в виде заштрихованного прямоугольника.

1. Сумма (объединение) событий (рис. 4.2) представляет собой сложное событие, состоящее в появлении хотя бы одного из событий A и B . Объединение событий обозначается как $A \cup B$, или $A + B$.

2. Произведением (пересечением) событий A и B называется их совместное появление (рис. 4.3). Обозначается произведение событий как $A \cap B$, или $A \bullet B$.

3. Достоверным событием называется событие, которое обязательно происходит в результате данного испытания (рис. 4.4). Оно обозначается обычно как E .

4. Невозможное событие – событие, которое не может произойти в результате данного испытания. Принятое обозначение – \emptyset .

5. Несовместными называются события, которые в результате данного испытания не могут произойти вместе (рис. 4.5). Примеры несовместных событий: попадание и промах при выстреле, выпадение двух и трех очков при бросании игральной кости. Рис. 4.5 наглядно показывает, что для несовместных событий $A \bullet B = \emptyset$.

6. Противоположным к A событием называется событие, состоящее в не появлении события A (рис. 4.6). Обозначается противоположное событие символом \bar{A} . Примеры противоположных событий: промах и попадание при выстреле, выпадение герба или цифры при одном подбрасывании монеты.



Рис. 4.1. Поле событий

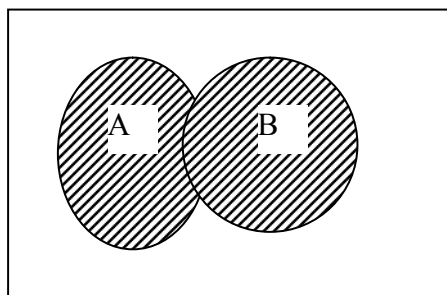


Рис. 4.2. Сумма событий

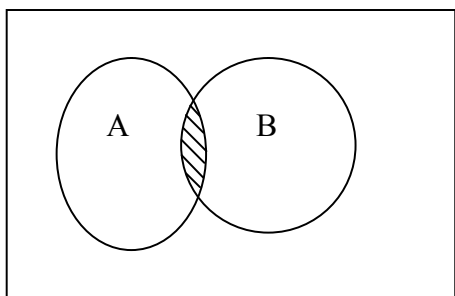


Рис. 4.3. Произведение событий



Рис. 4.4. Достоверное событие

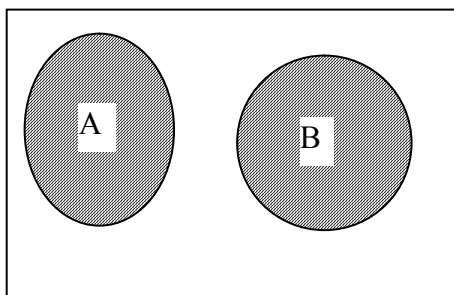


Рис. 4.5. Несовместные события

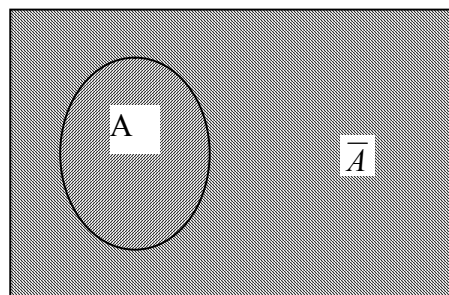


Рис. 4.6. Противоположные события

4.4. Исчисление вероятностей

4.4.1. Примеры непосредственного определения вероятностей

Рассмотрим несколько примеров на вычисление вероятностей по формуле (4.1).

Пример 4.1

Испытание состоит в подбрасывании игральной кости, на каждой из граней которой проставлено число очков (от 1 до 6). Какова вероятность того, что: 1) выпадает 2 очка? 2) выпадает нечетное число очков?

Решение 1: В данном испытании имеется 6 равновероятных случаев (выпадение 1, 2, 3, 4, 5, 6 очков), так как нет оснований предполагать, что появление какого-то определенного числа очков более вероятно (если, конечно, кость симметрична). Поэтому вероятность выпадения любого числа очков, в том числе и 2, при одном подбрасывании равна $1/6$.

Событию A , заключающемуся в появлении нечетного числа очков, благоприятствуют три случая (выпадение 1, 3 и 5), поэтому по формуле (4.1) получаем

$$P(A) = \frac{3}{6} = 0,5.$$

Решение 2: В данном испытании имеется 2 равновероятных исхода (выпадение четного числа очков (т.е. 2, 4, 6) и нечетного), так как кость симметрична, то очевидно, что эти исходы равновероятные.

Событию A , заключающемуся в появлении нечетного числа очков, благоприятствуют 1 случай из двух, поэтому по формуле (4.1) получаем $P(A) = 1/2 = 50\%$.

Отметим, что построенную таким образом пространство элементарных событий непригодно для расчета вероятности того, что выпадает 2 очка, так как этому событию не благоприятствует не один из введенных нами элементарных исходов.

Пример 4.2

В урне 5 белых и 10 черных шаров, не отличающихся по размеру. Шары тщательно перемешивают и затем наугад вынимают 1 шар. Какова вероятность того, что вынутый шар окажется белым?

Решение. В этом примере имеется 15 равновероятных (шары не отличаются по размеру) исходов опыта, причем событию A (появлению белого шара) благоприятствуют 5 из них, поэтому искомая вероятность составит $5/15 = 33.33\%$.

4.4.2. Основные правила вычисления вероятностей сложных событий

Ниже приведены основные правила, позволяющие определить вероятность появления сложного события на основании известных вероятностей составляющих его более простых событий.

1. Вероятность достоверного события равна единице:

$$P(E) = 1. \quad (4.2)$$

2. Вероятность объединения (суммы) несовместных событий равна сумме их вероятностей:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) \quad (4.3)$$

Эти два равенства являются аксиомами теории вероятностей, т. е. принимаются в качестве исходных, но требующих доказательства свойств вероятностей. На их основе строится вся теория вероятностей.

Все остальные, приведенные ниже без доказательств формулы могут быть выведены из принятых аксиом.

3. Вероятность невозможного события равна нулю:

$$P(\emptyset) = 0. \quad (4.4)$$

4. Вероятность события, противоположного событию A , равна

$$P(\bar{A}) = 1 - P(A). \quad (4.5)$$

Формула (4.5) оказывается полезной на практике в тех случаях, когда вычисление вероятности непосредственно события A затруднительно, в то время как вероятность противоположного события находится просто (см. ниже п.9).

5. Теорема сложения вероятностей. Вероятность объединения произвольных событий равна сумме их вероятностей за вычетом вероятности произведения событий:

$$P(A \cup B) = P(A) + P(B) - P(AB). \quad (4.6)$$

Для несовместных событий $P(AB) = 0$ и формула (4.6) переходит в (4.3).

6. Условная вероятность. Если требуется найти вероятность события B при условии, что произошло некоторое другое событие A , то такую ситуацию характеризуют с помощью условной вероятности $P(B|A)$. Условная вероятность равна отношению вероятности произведения событий A и B к вероятности события A :

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (4.7)$$

В тех случаях, когда события A и B несовместны, $P(AB) = 0$ и соответственно $P(B|A) = 0$.

7. Определение условной вероятности в виде (4.7) дает возможность записать следующую формулу для вычисления вероятности произведения событий (*теорема умножения вероятностей*)

$$P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B) \quad (4.8)$$

8. Поскольку вероятность события A (или B) для независимых событий по определению не изменяется при появлении другого события, то условная вероятность $P(A|B)$ совпадает с вероятностью события A , а условная вероятность $P(B|A)$ — с $P(B)$. Вероятности $P(A)$ и $P(B)$ в отличие от условных вероятностей называются безусловными.

$$P(A|B) = P(A), \quad P(B|A) = P(B), \quad (4.9)$$

Теорема умножения вероятностей для независимых событий записывается следующим образом:

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n), \quad (4.10)$$

т. е. вероятность произведения независимых событий равна произведению их вероятностей.

9. Вычислим вероятность появления хотя бы одного события в n испытаниях

A — появление в n испытаниях **хотя бы** один раз интересующего нас события.

\bar{A} — интересующее нас событие не появилось в n испытаниях **ни разу**.

A_1 — интересующее нас событие появилось в первом испытании.

A_2 — интересующее нас событие появилось во втором испытании.

....

A_n — интересующее нас событие появилось в n -ом испытании.

$$P(A) = 1 - P(\bar{A}) = 1 - P(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_n) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \dots \cdot P(\bar{A}_n) \quad (4.11)$$

10. Формула полной вероятности.

Если событие A может произойти только при появлении одного из несовместных событий H_1, H_2, \dots, H_n , то

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_n)P(A|H_n). \quad (4.12)$$

Пример 4.3

В урне 5 белых, 20 красных и 10 черных шаров, не отличающихся по размеру. Шары тщательно перемешивают и затем наугад вынимают 1 шар. Какова вероятность того, что вынутый шар окажется белым или черным?

Решение. Пусть событие A — появление белого или черного шара. Разобьем это событие на более простые. Пусть B_1 — появление белого шара, а B_2 — черного. Тогда, $A = B_1 + B_2$ по определению суммы событий. Следовательно $P(A) = P(B_1 + B_2)$. Так как B_1 и B_2 — несовместные события, то по теореме о вероятности суммы несовместных событий (формула 4.3) $P(B_1 + B_2) = P(B_1) + P(B_2)$.

Вычислим вероятности событий B_1 и B_2 . В этом примере имеется 35 равновозможных (шары не отличаются по размеру) исходов опыта, событию B_1 (появлению белого шара) благоприятствуют 5 из них, поэтому $P(B_1) = \frac{5}{35}$. Аналогично, $P(B_2) = \frac{10}{35}$. Следовательно, $P(A) = \frac{5}{35} + \frac{10}{35} = \frac{15}{35}$.

Пример 4.4

Ведутся поиски двух преступников. Каждый из них независимо от другого может быть обнаружен в 2006/2007 учебный год

течение суток с вероятностью 0,5. Какова вероятность того, что в течение суток будет обнаружен хотя бы один преступник?

Решение. Пусть событие A – “обнаружен хотя бы один преступник”. Разобьем это событие на более простые. Пусть B_1 – обнаружен первый преступник, а B_2 – обнаружен второй преступник. Тогда, $A=B_1+B_2$ по определению суммы событий. Следовательно $P(A)=P(B_1+B_2)$. Так как B_1 и B_2 – совместные события, то по теореме о вероятности суммы событий (формула 4.6)

$$P(B_1+B_2) = P(B_1)+P(B_2)-P(B_1 B_2) = 0,5+0,5 - 0,25=0,75.$$

Можно решать и через обратное событие: $P(A)=1-P(\bar{A})=1-P(\bar{B}_1 \cdot \bar{B}_2)=1-\frac{1}{2} \cdot \frac{1}{2}=1-\frac{1}{4}=\frac{3}{4}$.

Пример 4.5 а)

Преступник имеет 3 ключа. В темноте он открывает дверь выбирая ключ случайным образом. На открытие каждой из дверей он тратит 5 сек. Найти вероятность того, что он откроет все двери за 15 сек.

Решение. Пусть событие A – “открыты все двери”. Разобьем это событие на более простые. Пусть B – “открыта 1-я”, C – “открыта 2-я”, а D – “открыта 3-я”. Тогда, $A=BCD$ по определению произведения событий. Следовательно $P(A)=P(BCD)$. По теореме о вероятности произведения независимых событий (формула 4.10) $P(BCD) = P(B)P(C)P(D)$.

Вычислим вероятности событий B , C и D . В этом примере имеется 3 равновозможных (каждый ключ выбираем из трех) исходов опыта. Каждому из событий B , C и D благоприятствует 1 из них, поэтому

$$P(B)=P(C)=P(D)=\frac{1}{3}. \quad P(A)=\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

Пример 4.5 б)

Изменим задачу: считаем, что преступник – забывчивый человек. Пусть преступник открыв дверь, оставляет ключ в ней. Какова тогда вероятность, что он откроет все двери за 15 сек?

Решение. Событие A – “открыты все двери”. Опять, $A=BCD$ по определению произведения событий. Следовательно $P(A)=P(BCD)$. Но, теперь события B , C и D – зависимы. По теореме о вероятности произведения зависимых событий $P(BCD) = P(B)P(C|B)P(D|BC)$.

Вычислим вероятности: $P(B)=\frac{1}{3}$, $P(C|B)=\frac{1}{2}$ (ключа осталось только два и один из них подходит!),

$$P(D|BC)=\frac{1}{1} \text{ и, значит, } P(A)=\frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{6}.$$

Пример 4.6

Ведутся поиски двух преступников. Каждый из них независимо от другого может быть обнаружен в течение суток с вероятностью 0,5. После поимки одно из них, в связи с увеличением количества сотрудников, занятых в поисках, вероятность найти второго возрастает до 0,7. Какова вероятность того, что в течение суток будут обнаружены оба преступника.

Решение. Пусть событие A – “обнаружены два преступника”. Разобьем это событие на более простые. Пусть B_1 – обнаружен первый преступник, а B_2 – обнаружен второй преступник, после того, как пойман первый. Тогда, $A=B_1B_2$ по определению произведения событий. Следовательно $P(A)=P(B_1B_2)$. Так как B_1 и B_2 – зависимые события, то по теореме о вероятности произведения зависимых событий (формула 4.8) $P(B_1B_2) = P(B_1)P(B_2|B_1) = 0,5 \cdot 0,7=0,35$.

Пример 4.7

Найти вероятность того, что при подбрасывании монеты 10 раз герб выпадет хотя бы 1 раз.

Решение. Пусть событие A – “герб выпадет хотя бы 1 раз”. Рассмотрим обратное событие: \bar{A} – “герб не выпадет ни разу”. Очевидно, что обратное событие легче чем исходное разбить на более простые. Пусть A_1 – герб не выпал при первом броске, A_2 – герб не выпал при втором броске, ... A_{10} – герб не выпал при 10-м броске. Все события $A_1 \dots A_{10}$ независимы, следовательно, (формула 4.11)

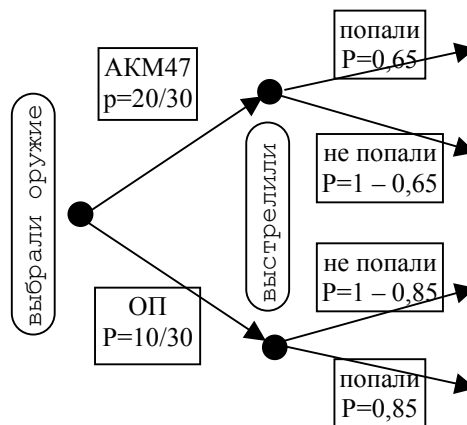
$$P(A)=1-P(\bar{A})=1-P(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_{10})=1-P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \dots \cdot P(\bar{A}_{10})=1-\frac{1}{2} \cdot \frac{1}{2} \cdot \dots \cdot \frac{1}{2}=1-\frac{1}{1024}=\frac{1023}{1024}.$$

Пример 4.8

В проведении операции по освобождению заложников участвуют 2 группы снайперов: 10 человек с винтовкой ОП21 и 20 человек с АКМ47. Вероятность поражения из ОП21 – 0,85, а АКМ47 – 0,65. Найти вероятность того, что при одном выстреле произвольного снайпера преступник будет поражен.

Решение. Пусть событие A – “преступник поражен”. Разобьем это событие на более простые. Преступник может быть поражен либо из ОП21, либо из АКМ47. Вероятность того, что произвольный снайпер вооружен ОП21 (событие H_1) равна 10/30. Вероятность того, что произвольный снайпер вооружен АКМ47 (событие H_2) равна 20/30.

В подобных задачах полезно изобразить “дерево возможностей”:



Вероятность того, что преступник поражен равна (формула 4.12)

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) = \frac{10}{30} \cdot 0,85 + \frac{20}{30} \cdot 0,65.$$

4.4.3. Комбинаторика

Правило сложения. Если первое действие можно выполнить n различными способами, а второе — m способами, то выполнить первое **ИЛИ** второе действие можно $m + n$ способами.

Правило умножения. Если первое действие можно выполнить n различными способами, а второе — m способами, то выполнить первое **И** второе действие (в таком порядке) можно nm способами.

Эти правила можно обобщить на случай 3-х, 4-х и более действий.

Пример 4.9. Рекламный плакат мебельной фабрики утверждает, что возможно составить 100000 различных вариантов расстановки производимых ей шкафов если купить хотя бы 5 шкафов. Верно ли это, если выпускается всего 10 различных типов шкафов?

Решение. Вычислим, сколькими способами можно расставить 5 шкафов рядом друг с другом. Первую позицию можно заполнить 10-ю различными способами (принцип сложения), вторую — также 10-ю (никто не мешает купить и второй шкаф той же модели, что и первый), третью — опять 10-ю и т.д. Вообще имеем (принцип умножения) $10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 100000$ различных вариантов расстановки пяти шкафов рядом друг с другом. Если же купить шкафов больше, чем 5, то, очевидно, вариантов расстановки будет еще больше.

Вывод: реклама является добросовестной.

Пример 4.10. Из тщательно перемешанной колоды в 52 карты вытягивают 3 карты. Сколько существует различных вариантов карт на руках у игрока?

Решение. В данном опыте производится 3 действия: вытягивание 1-й карты, 2-й карты и 3-й карты.

Вычислим, сколькими способами можно вытянуть 1-ую карту. Так как всего в колоде 52 карты, то имеем 52 различных способа. (Здесь мы применили принцип сложения: карта может быть двойка пик **ИЛИ** тройка пик **ИЛИ** ... **ИЛИ** туз червей. Значит, всего имеем $1+1+\dots+1=52$ способа.)

Вычислим, сколькими способами можно вытянуть 2-ую карту. Так как в колоде осталось 51 карта, то, значит, второе действие можно выполнить 51-м способом.

Аналогично рассуждая, находим, что 3-е действие можно осуществить 50-ю способами.

Всего различных вариантов **расположения** карт на руках у игрока будет $52 \cdot 51 \cdot 50 = 132600$ способов. Для ответа осталось разделить это число на $3 \cdot 2 \cdot 1$ — это кол-во способов перетасовать эти 3 розданные карты.

Ответ: 22100.

В общем виде:

Если стоит задача вычислить сколькими способами можно расположить “в ряд” (т.е. важен порядок их следования) вытянутые m предметов из коробки содержащей различных n предметов, то имеем так называемую ситуацию “перестановок”.

Вычислим это количество: первую позицию можно заполнить n способами, вторую — $n - 1$ способом, третью — $n - 2$ способом, и т.д. Искомое количество способов заполнить **все n позиций** равно (по принципу умножения)

$$n(n-1)(n-2)(n-3)\dots(n-m+1)$$

Если стоит задача вычислить сколькими способами можно расположить “в ряд” (т.е. важен порядок их следования) вытянутые m предметов из коробки содержащей различных n предметов, то имеем так называемую ситуацию “перестановок”.

Вычислим это количество: первую позицию можно заполнить n способами, вторую – $n - 1$ способом, третью – $n - 2$ способом, и т.д. Искомое количество способов равно (по принципу умножения)

$$A_n^m = n(n-1)(n-2)(n-3)\dots(n-m+1).$$

Но, поскольку нам не важно какой именно элемент стоит на каком месте, то необходимо A_n^m разделить на количество способов по разному переставлять уже выбранные элементы. А это количество равно $A_n^m = n(n-1)(n-2)(n-3)\dots 3 \cdot 2 \cdot 1 = n!$ (читается n факториал).

Искомое количество способов заполнить **все n позиций** равно $A_n^m / n!$ и обозначается C_n^m .

Пример 4.11. В совбесе ООН 11 членов: 5 постоянных и 6 так называемые “малые нации”. Для принятия решения, надо, чтобы было 7 голосов “ЗА”, причем следующим образом: все постоянные + как минимум 2 временных. Сколько всего вариантов голосования? Сколько всего можно организовать выигрышных коалиций? (Выигрышной коалицией называется такая, когда как бы не голосовали противники решение все равно будет принято.)

Решение. Так, как голосуют 11 делегаций и у них есть 2 выбора (“за”, “против”), то по принципу умножения имеем $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^{11} = 2048$ – вариантов голосования. Так как все постоянные члены должны проголосовать “за”, то выигрышная коалиция определяется только временными членами, а кол-во – количеством способов выбрать 2 или 3 или 4 или 5 или 6 временных членов, голосующих “за”.

$$\text{Имеем } 1 \cdot \left(C_6^2 + C_6^3 + C_6^4 + C_6^5 + C_6^6 \right) = 15 + 20 + 15 + 6 + 1 = 57 \text{ способов, причем } 15 - \text{число так называемых}$$

минимальных выигрышных коалиций.

4.4.4. Схема Бернулли

Пусть производится n одинаковых независимых опытов. В каждом испытании некоторое событие A может произойти с вероятностью p (а, значит, не произойти с вероятностью $q = 1 - p$).

Вычислим вероятность того, что событие произойдет **ровно k раз** в проведенных n опытах.

$$P_n(0) = \underbrace{q \cdot q \cdot \dots \cdot q}_n = q^n - \text{вероятность того, что во всех опытах событие не произойдет (см. также пример}$$

4.7)

$$P_n(1) = \underbrace{p \cdot q \cdot \dots \cdot q}_n + \underbrace{q \cdot p \cdot \dots \cdot q}_n + \dots + \underbrace{q \cdot q \cdot \dots \cdot p}_n = n \cdot p \cdot q^{n-1} - \text{вероятность того, что событие произойдет ровно}$$

в одном опыте (из проведенных n опытов).

$$P_n(2) = \frac{n(n-1)}{2} p^2 q^{n-2} - \text{вероятность того, что событие произойдет ровно 2 раза в } n \text{ опытах.}$$

$$P_n(k) = C_n^k p^k q^{n-k} - \text{вероятность того, что событие произойдет ровно } k \text{ раз в } n \text{ опытах.}$$

$$P_n(n) = \underbrace{p \cdot p \cdot \dots \cdot p}_n = p^n - \text{вероятность того, что событие произойдет во всех опытах.}$$

4.5. Случайные величины

Случайные величины (С.В.) – численное значение появляющееся в результате опыта, и принимающее произвольное значение из заранее определенного множества.

Существует два типа случайных величин: дискретные и непрерывные.

Дискретные случайные величины принимают в результате испытания одно из изолированного дискретного множества значений. Они хорошо подходят для описания результатов измерений, связанных с подсчетом и выражаемых целыми числами.

Примеры дискретных случайных величин: число зарегистрированных правонарушений, совершенных сделок, подтягиваний на перекладине и т. п.

Вероятность принятия дискретной случайной величиной каждого из возможных ее значений больше нуля. Эта вероятность может быть записана как

$$P(\{X = x_i\}) = p_i,$$

где $i = \dots -1, 0, 1 \dots$

Здесь X — обозначение случайной величины; x_i — конкретные числовые значения, принимаемые дискретной случайной величиной; p_i — вероятности этих значений.

Индекс i может в общем случае пробегать значения от $-\infty$ до ∞ .

Функция $P(\{X = x_i\})$, связывающая значения дискретной случайной величины с их вероятностями, называется ее **распределением (законом распределения)**. Обычно закон распределения записывается в виде таблицы вида

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

Пример 4.12. Пусть X — число очков выпавшее на игральной кости при одном броске. Тогда, эта с.в. распределена по закону

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

Непрерывные случайные величины в результате испытания могут принимать любые значения из некоторого интервала.

Примеры непрерывных случайных величин: спортивный результат в беге или прыжках, рост и масса тела человека, сила мышц и др.

Поскольку число возможных значений непрерывной случайной величины бесконечно велико и чаще всего нет оснований предположить, что одни значения появляются существенно чаще других, то вероятность принятия непрерывной случайной величиной каждого отдельного значения оказывается равной нулю. По этой причине нельзя описать распределение непрерывной случайной величины в виде вероятностей ее отдельных значений, как в случае дискретных случайных величин. Здесь необходимы другие подходы, которые рассмотрены в разделах 4.6 и 4.7.

4.6. Функция распределения

Рассмотрим вероятность того, что случайная величина X окажется меньше или равной некоторому заданному **числу** x , т. е.

$$F(x) = P(\{X \leq x\}). \quad (4.11)$$

Эта вероятность, рассматриваемая как функция переменной x , называется функцией распределения случайной величины X . Она используется для записи распределений как дискретных, так и непрерывных случайных величин. Для дискретной случайной величины функция распределения имеет скачки. Функция распределения непрерывной случайной величины будет непрерывной функцией (рис. 4.8).

Как было сказано ранее, вероятность принятия непрерывной случайной величиной какого-либо конкретного значения равна 0. Но для непрерывной случайной величины обычно и надо определить вероятность попадания ее в заданный интервал $[x_1; x_2]$, которая по известной функции распределения находится как

$$P(\{x_1 \leq X < x_2\}) = F(x_2) - F(x_1) \quad (4.12)$$

В этом выражении совершенно не обязательно записывать интервал таким образом. Можно было бы записать $x_1 < X < x_2$, $x_1 \leq X \leq x_2$ или $x_1 < X \leq x_2$, при этом вероятность попадания случайной величины в интервал не изменится. Это связано с тем, что, как уже отмечалось, функция распределения случайной непрерывной величины не имеет скачков ни при каких значениях x .

Свойства функции распределения совпадают со свойствами эмпирической функции распределения (см. 2.3.4)

1. $F(x)$ неубывающая функция.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$
3. $\lim_{x \rightarrow +\infty} F(x) = 1$

График функция распределения представляет собой теоретический аналог полигона накопленных частот, рассмотренного в разделе 2.3.3.

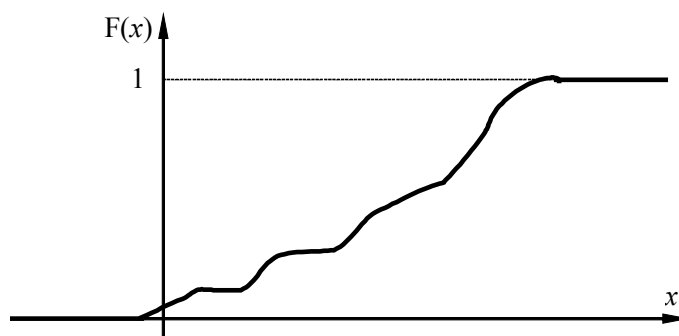


Рис. 4.8. Функция распределения непрерывной случайной величины

4.7. Плотность распределения вероятностей

Для непрерывных случайных величин вводится понятие плотность распределения вероятностей, или “плотность вероятностей”, играющее исключительно важную роль при их описании.

Плотность вероятностей — это производная от функции распределения непрерывной случайной величины, т.е.

$$f(x) = \frac{dF(x)}{dx} \quad (4.14)$$

Более подробно при рассмотрении конкретных непрерывных распределений об этой функции рассказано в разделе 4.9. Типичный вид графика плотности вероятностей показан на рис. 4.9.

Вероятность попадания непрерывной случайной величины в интервал между значениями x_1 и x_2 пропорциональна площади под кривой плотности вероятностей, заключенной между точками x_1 и x_2 . Эта вероятность математически записывается в виде интеграла от $f(x)$ в пределах x_1 и x_2 .

$$P(\{x_1 \leq X < x_2\}) = \int_{x_1}^{x_2} f(x) dx. \quad (4.15)$$

Свойства $f(x)$:

1. $f(x) \geq 0$;
2. $f(x) = 0$ при $x < x_{\min}$;
3. $f(x) = 0$ при $x > x_{\max}$;
4. $\int_{-\infty}^{+\infty} f(x) dx = 1$.

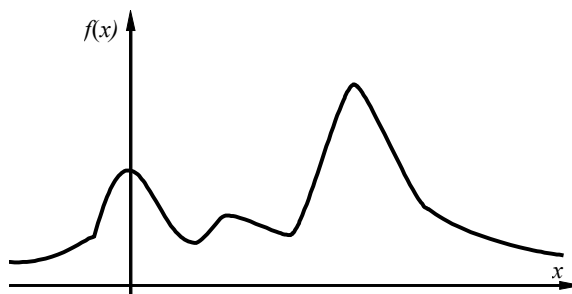


Рис. 4.9. Вид графика плотности распределения вероятностей

График плотности вероятностей является теоретическим аналогом полигона частот и гистограммы, (2.3.2 и 2.3.4).

4.8. Числовые характеристики случайных величин

Распределение случайной величины, заданное в виде функции распределения или плотности вероятностей, полностью ее характеризует. Однако такая исчерпывающая характеристика случайной величины сложна и далеко не всегда необходима. Для решения многих практических задач не нужно знать распределение случайной величины, а достаточно иметь лишь некоторые обобщающие числовые характеристики этого распределения.

4.8.1. Математическое ожидание

Для более наглядного определения математического ожидания (среднего значения) случайной величины рассмотрим подход к этому понятию на конкретном примере.

Пусть имеется дискретная случайная величина X с возможными значениями x_1, x_2, \dots, x_k и вероятностями этих значений p_1, p_2, \dots, p_k . В качестве примера X рассмотрим случайную величину — количество правонарушений за сутки. Каждое из значений x_1, x_2, \dots, x_n (отмечено 0; 1; ... нарушений) будет наблюдаться некоторое число раз. Обозначим эти числа через n_1, n_2, \dots, n_k . Очевидно, что сумма $n_1 + n_2 + \dots + n_k = n$.

Таким образом, имеем n наблюдений за случайной величиной X , т. е. выборку объема n . Определим выборочное среднее арифметическое: $\bar{x}_{Bn} = \frac{1}{n} \sum_{i=1}^k n_i x_i = \sum_{i=1}^k x_i \frac{n_i}{n}$.

Здесь индекс n при \bar{x}_B обозначает, что среднее арифметическое вычислено по n наблюдениям.

Теперь представим, что испытание, состоящее в регистрации количества правонарушений, повторяется неограниченное число раз. Здесь, абстрагируясь от физической реализуемости такого эксперимента, будем считать, что наблюдению доступна вся теоретически бесконечная генеральная совокупность значений случайной величины X .

Согласно статистическому определению вероятности, данному в разделе 4.2.2, относительные частоты событий стремятся к их вероятностям при неограниченном повторении испытания.

Поэтому в пределе при $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \bar{x}_{Bn} = \lim_{n \rightarrow \infty} \sum_{i=1}^k x_i \frac{n_i}{n} = \sum_{i=1}^k x_i \lim_{n \rightarrow \infty} \frac{n_i}{n} = \sum_{i=1}^k x_i p_i.$$

Таким образом, выборочное среднее арифметическое случайной величины X стремится при неограниченном повторении испытания (при неограниченном увеличении объема выборки) к некоторому постоянному числу, так как в последней сумме x_i и p_i — постоянные числа. Это число носит название математического ожидания (среднего значения) случайной величины.

Математическое ожидание обозначает как $M(X)$ или m_x .

Математическое ожидание дискретной случайной величины равно сумме всех ее возможных значений, умноженных на вероятности этих значений:

$$M(X) = m_x = \sum_i x_i P(X = x_i) = \sum_i x_i p_i. \quad (4.16)$$

В этой записи \sum_i означает, что суммирование производится по всем возможным i .

Только что рассмотренный пример показывает, что математическое ожидание — абстрактное понятие. Оно является теоретическим аналогом выборочного среднего арифметического.

Математическое ожидание равно среднему значению генеральной совокупности.

Для непрерывных случайных величин математическое ожидание определяется с помощью плотности вероятностей по формуле:

$$M(X) = m_x = \int_{-\infty}^{\infty} x f(x) dx. \quad (4.17)$$

Свойства математического ожидания

1. $M(C) = C$,
2. $M(CX) = CM(X)$.
3. $M(X+Y) = M(X) + M(Y)$,
4. $M(X \cdot Y) = M(X)M(Y)$, если X и Y — независимые с.в.

4.8.2. Дисперсия и стандартное отклонение

Точно так же, как математическое ожидание, являющееся теоретическим аналогом среднего арифметического, можно ввести теоретические аналоги всех числовых характеристик выборки, рассмотренных в гл. 3. Для этого нужно в соответствующих формулах для выборочных характеристик

заменить все средние арифметические на математические ожидания.

Дисперсией случайной величины X называется математическое ожидание квадрата отклонений случайной величины от ее математического ожидания (сравните с определением п. 3.4.2). Дисперсия обозначается как $D(X)$, или σ^2 .

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

Для дискретных случайных величин

$$D(X) = \sum_i (x_i - M(X))^2 p_i, \quad (4.18)$$

т. е. дисперсия дискретной случайной величины равна сумме квадратов отклонений отдельных значений случайной величины от ее математического ожидания, умноженных на вероятности этих значений.

Для непрерывных случайных величин

$$D(X) = \int_{-\infty}^{\infty} (x - M(X))^2 f(x) dx. \quad (4.19)$$

Положительный корень квадратный из дисперсии называется средним квадратическим (стандартным) отклонением случайной величины.

Эта величина обозначается, как σ_x

$$\sigma_x = \sqrt{D}. \quad (4.20)$$

Дисперсия и стандартное отклонение характеризуют изменчивость (вариативность) случайной величины. Чем сильнее случайная величина отклоняется от своего математического ожидания, тем больше величины $D(X)$ и σ_x . Последнюю (σ_x) использовать удобнее, так как его размерность совпадает с размерностью случайной величины (Если С.В. X – количество выигранных в лотерею \$, то D измеряется в долларах “в квадрате”, что не имеет никакого прикладного смысла, а σ_x – это доллары!).

Свойства дисперсии

1. $D(C) = 0$,
2. $D(CX) = C^2 \cdot D(X)$.
3. $D(X + Y) = D(X) + D(Y)$, если X и Y – независимые с.в.

Обобщение 4.8.1 и 4.8.2.

Пусть X_1, \dots, X_n – с.в., а C_1, \dots, C_n – константы. Тогда $M(C_1 X_1 + \dots + C_n X_n) = C_1 M(X_1) + \dots + C_n M(X_n)$. В частности $M\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{M(X_1) + \dots + M(X_n)}{n}$. Если более того X_1, \dots, X_n – независимые, то

$$D(C_1 X_1 + \dots + C_n X_n) = C_1^2 D(X_1) + \dots + C_n^2 D(X_n). \text{ В частности } D\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{D(X_1) + \dots + D(X_n)}{n^2}.$$

Пример 4.13.

Играем в следующую игру: один раз бросаем игральную кость и получаем столько \$, сколько выпало очков. Цена игры: 4\$. Выгодно ли играть?

Пусть с.в. X – количество очков, выпавшее при броске игральной кости. (см. пример 4.12)

Вычислим $M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3,5$ – именно столько очков (а, значит, и \$) ”в

среднем” мы будем получать если играть достаточно долго. Значит, игра невыгодна для нас. Мы ”в среднем” теряем 0.5\$ в каждой игре. Для вычисления $D(X)$ обычно пользуются формулой

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

С.В. X^2 имеет следующее распределение

X^2	1	$2^2=4$	$3^2=9$	$4^2=16$	$5^2=25$	$6^2=36$
P	1/6	1/6	1/6	1/6	1/6	1/6

Вычислив $M(X^2)=91/6$ находим $D(X) = 91/6 - 441/36 = 105/36 = 2\frac{11}{12}$.

4.8.3. Моменты

Математическое ожидание и дисперсия представляют собой частные случаи общих числовых

характеристик случайной величины, называемых моментами.

Ниже кратко рассматриваются лишь так называемые центральные моменты случайной величины.

S -ым центральным моментом случайной величины X называется математическое ожидание s -й степени отклонения случайной величины от ее математического ожидания:

$$\mu_s = M((X - M(X))^s).$$

В частности, при $s = 2$ второй центральный момент случайной величины есть дисперсия.

На практике часто используются также третий и четвертый центральные моменты, позволяющие судить о симметричности и остроте вершины кривой распределения случайной величины.

Если $\mu_3 = 0$, то распределение симметрично относительно математического ожидания, если $\mu_3 > 0$, то преобладают положительные отклонения от математического ожидания, если $\mu_3 < 0$ — отрицательные. Для удобства применяется так называемый коэффициент асимметрии, который является безразмерной величиной и определяется как

$$\gamma_3 = \frac{\mu_3}{\sigma^3} \quad (4.21)$$

Об остроте вершины кривой распределения судят по коэффициенту эксцесса:

$$\gamma_4 = \frac{\mu_4}{\sigma^4} - 3 \quad (4.22)$$

Если $\gamma_4 > 0$, то распределение имеет острый пик, если $\gamma_4 < 0$ (минимальное значение $\gamma_4 = -2$), то распределение имеет плосковершинную форму по сравнению с рассмотренным ниже нормальным распределением, для которого $\gamma_4 = 0$.

4.9. Биномиальное распределение. Распределение Бернулли и Пуассона.

Пусть проводятся n испытаний по схеме Бернулли (см. 4.4.3.). Событие A может произойти в результате этой серии опытов 0 раз, 1 раз, ... n раз. Рассмотрим случайную величину — число испытаний в которых событие A произошло. Имеем дискретную с.в. с законом распределения

X	0	1	...	k	...	n
P	$P_n(0) = q^n$	$P_n(1) = n \cdot p \cdot q^{n-1}$...	$P_n(k) = C_n^k p^k q^{n-k}$...	$P_n(n) = p^n$

Также говорят, что с.в. X распределена по **биномиальному** закону с параметрами n и p и пишут $X \sim B(n; p)$.

Если $n = 1$, то говорят, что с.в. X имеет **распределение Бернулли** параметром p .

Теорема. Пусть X_1, X_2, \dots, X_n — независимые с.в. распределенные по Бернулли с одинаковым параметром p . Пусть $S_n = X_1 + X_2 + \dots + X_n$. Тогда $S_n \sim B(n, p)$.

Числовые характеристики биномиального закона.

$$M(X) = n \cdot p,$$

$$D(X) = n \cdot p \cdot q.$$

Если n — велико, а p — мало, то вычисления вероятности по формуле $P_n(k) = C_n^k p^k q^{n-k}$ на практике невозможно. При этих условиях используется **формула Пуассона** для вычисления вероятности маловероятных событий в массовых испытаниях: $P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, где $\lambda = np$. Соответствующая с.в. распределена по закону Пуассона.

Пример 4.14.

Вероятность встретить блондина в равна 20%. Опрашивается 1000 человек. Как распределено X — число блондинов в этой выборке? Найти вероятность того, что среди опрошенных 100 человек будет 20 блондинов.

Поскольку имеем 100 испытаний по схеме Бернулли (с вероятностью успеха $p = 20\% = 0,2$), то $X \sim B(1000; 0,2)$. Искомая вероятность вычисляется как $P_{100}(20) = C_{100}^{20} 0,2^{20} 0,8^{80}$, и довести это выражение до конкретного числа сложная задача даже с использованием компьютера. Но, поскольку имеем право

использовать формулу Пуассона, то приблизительно имеем $\lambda = 100 \cdot 0.2 = 20$, $P_{100}(20) = \frac{20^{20}}{20!} e^{-20}$ а это значение уже легко вычислить даже в Excel и получить $0.08883531739208530000=8,88\%$.

4.10. Нормальное распределение

4.10.1. Определение и значение

Большинство экспериментальных исследований, в том числе и в области правоведения, связано с измерениями, результаты которых могут принимать практически любые значения в заданном интервале и, как уже было отмечено, описываются моделью непрерывных случайных величин. Поэтому в дальнейшем будут рассматриваться в основном непрерывные случайные величины и связанные с ними непрерывные распределения.

Одним из непрерывных распределений, имеющим основополагающую роль в математической статистике, является нормальное, или гауссово*, распределение.

Нормальное распределение является самым важным в статистике. Это объясняется целым рядом причин.

1. Прежде всего, многие экспериментальные наблюдения можно успешно описать с помощью нормального распределения. Следует сразу же отметить, что не существует распределений эмпирических данных, которые были бы в точности нормальными, поскольку (как будет показано ниже) нормально распределенная случайная величина находится в пределах от $-\infty$ до $+\infty$, чего никогда не бывает на практике. Однако нормальное распределение очень часто хорошо подходит как приближение.

Проводятся ли измерения IQ, роста и других физиологических параметров — везде на результаты оказывает влияние очень большое число случайных факторов (естественные причины и ошибки измерения). Причем, как правило, действие каждого из этих, факторов незначительно. Опыт показывает, что результаты именно в таких случаях будут распределены приближенно нормально.

2. Нормальное распределение хорошо подходит в качестве аппроксимации (приближенного описания) других распределений (например, биномиального).

3. Многие распределения, связанные со случайной выборкой, при увеличении объема последней переходят в нормальное.

4. Нормальное распределение обладает рядом благоприятных математических свойств, во многом обеспечивших его широкое применение в статистике.

В то же время следует отметить, что в природе встречается много экспериментальных распределений, для описания которых модель нормального распределения малопримодна. Для этого в математической статистике разработан ряд методов, некоторые из которых приводятся в следующих главах.

Говорят, что с.в. распределена по нормальному закону с параметрами m и σ и записывать $X \rightarrow N(m, \sigma^2)$ если ее плотность вероятностей задается следующим образом

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad -\infty < x < \infty. \quad (4.23)$$

График плотности (нормальная кривая) представлен на рис. 4.10.

Укажем основные свойства нормального распределения $N(m, \sigma^2)$.

1. Нормальная кривая имеет колоколообразную форму, симметричную относительно точки $x = m$, с точками перегиба, абсциссы которых отстоят от m на $\pm \sigma$.

2. Для нормального распределения математическое ожидание $M(X) = m$, дисперсия равна σ^2 и, следовательно, стандартное отклонение равно σ .

3. Как видно из выражения (4.23), нормальное распределение полностью определяется двумя параметрами: m и σ — математическим ожиданием и стандартным отклонением.

График плотности вероятности нормального распределения показывает, что для нормально распределенной случайной величины вероятность отклонения от среднего значения m быстро уменьшается с ростом величины отклонения.

* Гаусс Карл Фридрих (1777—1855) — немецкий ученый в области практической и прикладной математики. Распределение названо по его имени, так как он ввел его одним из первых.

4. Медиана и мода нормального распределения совпадают и равны математическому ожиданию m .

5. Коэффициенты асимметрии и эксцесса нормального распределения равны нулю ($\gamma_3 = 0$, $\gamma_4 = 0$).

Последнее свойство (5) используется для проверки предположения о нормальности распределения генеральной совокупности (гл. 6).

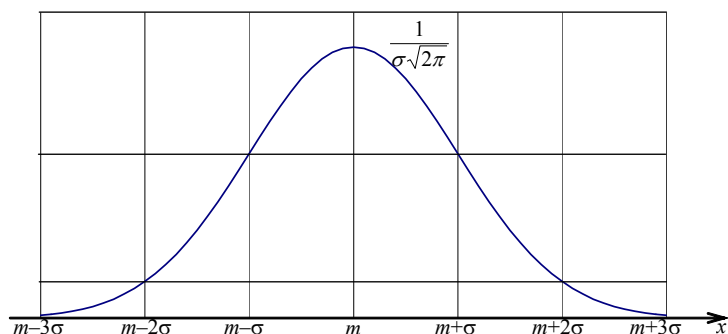


Рис. 4.10. Плотность вероятностей нормального распределения

4.10.2. Нормированное нормальное распределение

Формула (4.23) описывает целое семейство нормальных кривых, зависящих, как было сказано выше, от двух параметров — m и σ , которые могут принимать любые значения, поэтому возможно бесконечно много нормально распределенных совокупностей.

Чтобы избежать неудобств, связанных с расчетами для каждого конкретного случая по достаточно сложной формуле (4.23), используют так называемое нормированное (или стандартное) нормальное распределение $N(0;1)$, для которого составлены подробные таблицы.

Нормированное нормальное распределение имеет параметры $m=0$ и $\sigma=1$. Это распределение получается, если **пронормировать** нормально распределенную величину X по формуле:

$$U = \frac{X - m}{\sigma} \quad (4.24)$$

Плотность распределения вероятностей нормированного нормального распределения записывается в виде:

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, -\infty < u < \infty.$$

На кривой нормированного нормального распределения (рис. 4.11) указаны в процентах доли площадей соответствующих отмеченным значениям нормированного отклонения u , по отношению к общей площади под кривой, равной 1 (100%). Эти площади определяют вероятности попадания случайной величины в соответствующие интервалы.

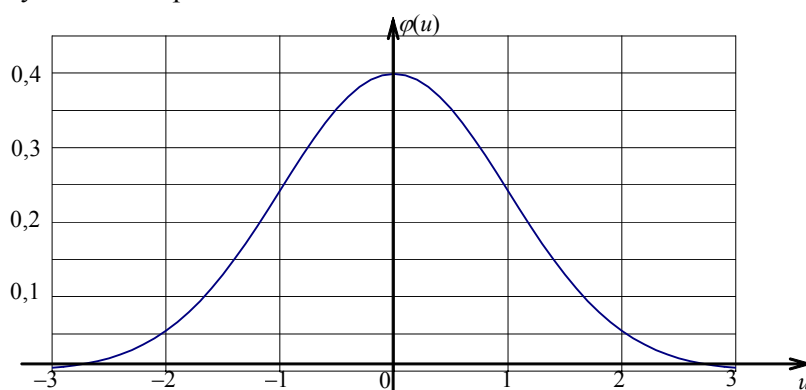


Рис. 4.11. Нормированное нормальное распределение

Таблица значений — ординат нормальной кривой приведена в специальных таблицах. Значения $\varphi(u)$ для некоторых характерных нормированных отклонений представлены в табл. 4.1.

Таблица 4.1
Ординаты нормальной кривой

Нормированное отклонение, u	0	$\pm 0,5$	$\pm 1,0$	$\pm 2,0$	$\pm 3,0$
Ордината нормальной кривой, $\varphi(u)$	0,399	0,352	0,242	0,054	0,004

4.10.3. Вероятность попадания в заданный интервал

Очень часто исследователя интересует вопрос: какова вероятность того, что изучаемый признак генеральной совокупности находится в заданных границах (например, вероятность того, что результат измерения IQ для группы испытуемых окажется в пределах 115 — 125)? Если предполагается нормальное распределение признака в генеральной совокупности, то получить ответ на этот вопрос очень просто.

Как говорилось ранее, вероятность попадания нормально распределенной случайной величины в заданный интервал $[x_1; x_2]$ можно определить по функции распределения: $P(x_1 < X < x_2) = F(x_2) - F(x_1)$ или

с помощью функции плотности вероятностей: $P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx$.

Итак, вероятность попадания с.в. $U \rightarrow N(0;1)$ в заданный интервал:

$$P(a < U < b) = P(U < b) - P(U < a) = \Phi(b) - \Phi(a),$$

где Φ — принятое обозначение для функции нормированного нормального распределения которое имеет следующий вид:

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt, \quad (4.25)$$

при этом $\Phi(-u) = 1 - \Phi(u)$.

Часто представляет интерес вероятность попадания с.в. $U \rightarrow N(0;1)$ в симметричный интервал. Тогда

$$P(-a < U < a) = \Phi(a) - \Phi(-a) = \Phi(a) - (1 - \Phi(a)) = 2\Phi(a) - 1 \quad (4.25^*)$$

Учитывая свойства функции Лапласа, получаем:

Интеграл, входящий в выражение (4.25), не выражается в элементарных функциях, поэтому для вычисления функции $\Phi(u)$ используют вспомогательную функцию — функцию Лапласа (интеграл вероятностей):

$$\Phi_0(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-t^2/2} dt \quad (4.26)$$

который табулируется. Функция Лапласа является нечетной, т.е. $\Phi_0(-u) = -\Phi_0(u)$.

В книгах по теории вероятности приведена либо таблица значений функции Лапласа $\Phi_0(u)$, либо $\Phi(u)$.

Чтобы найти вероятность попадания нормально распределенной случайной величины $X \rightarrow N(m, \sigma^2)$ в заданный интервал $[x_1; x_2]$ с помощью функции Лапласа, сначала с.в. X нормализуется (см. 4.24), а затем используется следующая формула:

$$\begin{aligned} P(x_1 < X < x_2) &= P(x_1 < m + \sigma U < x_2) = P(x_1 - m < \sigma U < x_2 - m) = P\left(\frac{x_1 - m}{\sigma} < U < \frac{x_2 - m}{\sigma}\right) = \\ &= \Phi_0\left(\frac{x_2 - m}{\sigma}\right) - \Phi_0\left(\frac{x_1 - m}{\sigma}\right) \end{aligned} \quad (4.27)$$

Пример 4.14. Вычислить $P(0,74 < X < 3,26)$ если $X \rightarrow N(2,9)$.

Решение.

$$P(0,74 < X < 3,26) = P(0,74 < 2 + \sqrt{9}U < 3,26) = P(-1,26 < 3U < 1,26) = P(-0,42 < U < 0,42) = 2P(0,42) - 1 = 2 \cdot 0,6628 - 1 = 0,3256$$

4.10.4. Правило трех сигм

В табл. 4.2 приведены полученные по формуле (4.28) вероятности того, что нормально распределенная случайная величина отклонится от своего среднего значения m , не более, чем на $\pm 0,5\sigma$, $\pm\sigma$, $\pm 2\sigma$, $\pm 3\sigma$.

Таблица 4.2

Вероятности попадания нормально распределенной случайной величины в заданный интервал

Границы интервала, $m \pm x$	$m \pm 0.5\sigma$	$m \pm \sigma$	$m \pm 2\sigma$	$m \pm 3\sigma$
Вероятность попадания в интервал	0,3829	0,6827	0,9545	0,9973

Из табл. 4.2 следует, что $P[-3\sigma < (X - m) < 3\sigma] = 0,9973$.

Это выражение известно в статистике как “правило трех сигм”. Оно означает, что с вероятностью 0,9973 (практически с единичной) нормально распределенная случайная величина окажется в пределах $\pm 3\sigma$ от среднего значения. Иначе говоря, отклонения от среднего больше чем на $+3\sigma$ можно ожидать примерно в 1 случае из 370 испытаний.

4.11. Применение нормального распределения.

Теорема Муавра-Лапласа.

Пусть X — с.в. распределенная по биномиальному закону с параметрами n и p (т.е. $X \rightarrow B(n, p)$).

Если $n \geq 30$, $np(1-p) \geq 5$, то можно приблизительно описать закон распределения с.в. с помощью $N(np, np(1-p))$.

Тогда вероятность того, что дискретная с.в. X примет значение k (т.е. событие произойдет ровно k раз в проведенных n опытах) вычисляют так: $P(X = k) = P\left(k - \frac{1}{2} < X < k + \frac{1}{2}\right) = P\left(-\frac{1}{2} < np + \sqrt{npq} \cdot U < k + \frac{1}{2}\right)$, где U — нормированная нормальная с.в.

Пример 4.15. Вычислить вероятность того, что при 144 бросках игральной кости двойка выпадет от 28 до 30 раз.

Пусть X — кол-во выпадений двойки в этих 144 бросках. Очевидно, что X — с.в. распределенная по биномиальному закону $B\left(144, \frac{1}{6}\right)$. Вычислить $P(27 < X < 31) = P_{144}(28) + P_{144}(29) + P_{144}(30)$ по формуле

Бернулли не представляется возможным. Но т.к. $n = 144 \geq 30$ и $np(1-p) = 144 \cdot \frac{1}{6} \cdot \frac{5}{6} = 120 \geq 5$, то применим

теорему Муавра-Лапласа: $X \rightarrow B\left(144, \frac{1}{6}\right) \approx N\left(144 \cdot \frac{1}{6}, 144 \cdot \frac{1}{6} \cdot \frac{5}{6}\right) = N(24, 20)$, и, следовательно

$$P(27 < X < 31) = P\left(27 < 24 + \sqrt{24} \cdot U < 31\right) = P\left(27 - 24 < \sqrt{24} \cdot U < 31 - 24\right) = P\left(\frac{3}{\sqrt{24}} < U < \frac{7}{\sqrt{24}}\right) =$$

$$= P(0,61 < U < 1,43) = P(U < 1,43) - P(U < 0,61) = 0,9236 - 0,7291 = 0,1945$$

Теорема (центральная предельная в упрощенной формулировке).

Если случайные величины X_1, X_2, \dots, X_n — независимы, одинаково распределены (все равно по какому закону!) и имеют конечные мат.ожидания $M[X]$ и дисперсии $D[X]$, то при $n \geq 30$ их среднее арифметическое $\bar{X}_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$ распределено близко к $N\left(M[X], \frac{D[X]}{n}\right)$ или, что то же самое, их сумма S_n распределена близко к $N(n \cdot M[X], n \cdot D[X])$.

Пример 4.16. Игровая кость, используемая мошенниками, подверглась испытаниям в результате которых были установлены вероятности выпадения граней (т.е. построен закон распределения с.в. X — кол-во выпавших очков):

X	1	2	3	4	5	6
P	0,1	0,1	0,1	0,1	0,2	0,4

Вычислить вероятность того, что при 50 бросках среднее кол-во очков будет между 3 и 4 (т.е. вычислить $P\left(3 < \frac{X_1 + X_2 + \dots + X_{50}}{50} < 4\right)$, где X_i — число полученное при i -ом броске).

Решение.

Очевидно, что X_1, X_2, \dots, X_{50} — независимы, одинаково распределены, следовательно центральная предельная теорема можно аппроксимировать закон распределения с.в. $\bar{X}_{50} = \frac{X_1 + X_2 + \dots + X_{50}}{50}$ с помощью $N\left(M[X], \frac{D[X]}{50}\right)$.

Вычисли мат. ожидание и дисперсию для с.в. X : $M[X]=1\cdot 0,1+2\cdot 0,1+3\cdot 0,1+4\cdot 0,1+5\cdot 0,2+6\cdot 0,4=4,4$ очка, $D[X]=3,04$.

Итак, с.в. \bar{X}_{50} распределена приблизительно по закону $N\left(4,4; \frac{3,04}{50}\right)$. Вычислим теперь

$$P(3 < \bar{X}_{50} < 4) = P\left(3 < 4,4 + \sqrt{\frac{3,04}{50}} \cdot U < 4\right) = P(-1,6 < 0,25 \cdot U < -0,4) = P(-1,62) - P(-6,48) = (1 - 0,94738) - 0 = 0,05262$$

Итак, получить в среднем между 3 и 4 очками маловероятно (вероятность лишь около 5% в то время как для стандартной игральной кости эта вероятность будет более 96%. Получите этот результат самостоятельно).

4.12. Некоторые специальные непрерывные распределения

Нормальное распределение широко применяется как математическая модель для описания экспериментальных данных. В этом разделе будут рассмотрены три распределения, которые играют очень важную роль при обработке результатов, связанных со случайной выборкой объема n , и составляют основу применения критериев значимости и проверки статистических гипотез. Примеры использования этих распределений приводятся в гл. 6, посвященной указанным статистическим методам.

4.12.1. χ^2 -распределение

Если U_1, U_2, \dots, U_ν независимые случайные величины, каждая из которых имеет нормированное нормальное распределение с параметрами $m=0$ и $\sigma=1$, то сумма квадратов этих величин

$$\chi^2 = U_1^2 + \dots + U_\nu^2$$

имеет так называемое χ^2 (хи-квадрат)-распределение. Его плотность вероятностей представлена на рис. 4.12 и зависит от единственного параметра — числа степеней свободы ν .

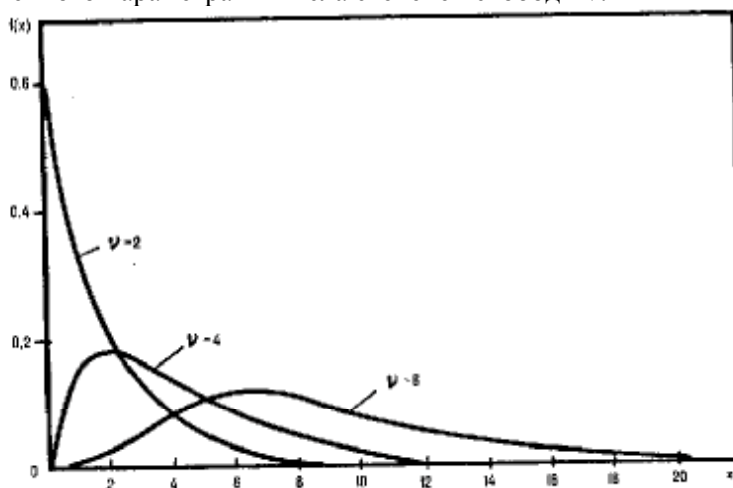


Рис. 4.12. χ^2 -распределение

Кривая χ^2 -распределения имеет положительную асимметрию. С ростом числа степеней свободы ν она становится все более симметричной и при $\nu \geq 30$ переходит в нормальное.

Таблицы χ^2 -распределения приводятся в приложениях к учебникам по ТВ. В этих таблицах обычно содержатся значения x , соответствующие вероятностям $P = 1 - \alpha$, при α , равном 0,05; 0,01 и 0,001 для различного числа степеней свободы ν .

4.12.2. t-распределение Стьюдента

Вторым из широко используемых специальных распределений является t-распределение Стьюдента, или просто t-распределение. Это распределение случайной величины:

$$T = \frac{U}{\sqrt{V/\nu}},$$

где U — случайная величина, имеющая нормированное нормальное распределение; V — случайная величина с распределением χ^2 с ν степенями свободы, t -распределение применяется при проверке статистических гипотез при малом объеме выборки. Эти вопросы рассмотрены в гл. 6. Форма t -распределения полностью определяется одним параметром — числом степеней свободы ν .

Вид кривой плотности t -распределения показан на рис. 4.13. t -распределение симметрично при любом ν и при $\nu \geq 30$ переходит в нормальное с параметрами $\mu = 0$ и $\sigma = \sqrt{\nu/(\nu-2)}$.

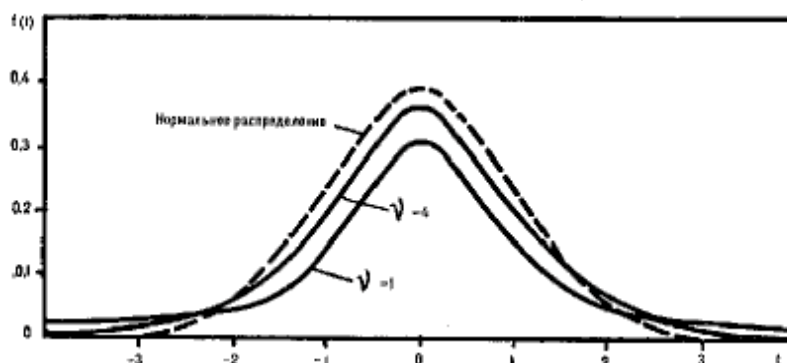


Рис. 4.13. t -распределение Стюдента

Таблицы t -распределения обычно приводятся в приложениях к учебникам по ТВ.

4.12.3. F-распределение

Если случайные величины U и V независимы и каждая из них распределена как χ^2 с ν_1 и ν_2 степенями свободы соответственно, то величина

$$F = \frac{U / \nu_1}{V / \nu_2}$$

подчиняется так называемому F-распределению, которое зависит от двух параметров — ν_1 и ν_2 , называемых числами степеней свободы, F-распределение применяется в основном в задачах, связанных с дисперсиями. Эти задачи также рассмотрены в гл. 6. Таблицы χ^2 -распределения приводятся в приложениях к учебникам по ТВ.

5. ОЦЕНКА ГЕНЕРАЛЬНЫХ ПАРАМЕТРОВ

5.1. Введение

Материал, содержащийся в предыдущих главах, можно рассматривать как минимальный набор сведений, необходимых для использования основных статистических методов, объединенных в теории статистических выводов.

Перейдем теперь к рассмотрению этих методов. Для этого необходимо определить их место в рамках единого подхода к решению конкретных задач статистических исследований в области правоведения.

Основная задача, решаемая с помощью методов математической статистики, — получение информации о закономерностях изменения изучаемого признака для большой совокупности объектов исследования, объединенных по этому признаку. В терминах математической статистики это означает, что делаются выводы о свойствах генеральной совокупности.

В гл. 4 определено, что для описания генеральной совокупности используются математические модели теории вероятностей. Исчерпывающую информацию о генеральной совокупности дает распределение вероятностей. Чаще всего используется модель нормально распределенной генеральной совокупности. И в этом случае распределение полностью определено всего двумя параметрами: средним значением (математическим ожиданием) и стандартным отклонением.

Следовательно, чтобы полностью описать нормальную генеральную совокупность, нужно знать значения двух генеральных параметров: среднего значения и стандартного отклонения. Так, если интерес вызывают данные о количестве правонарушений, то это среднее количество всех правонарушений и стандартное отклонение. Эти параметры неизвестны и предположительно находятся в каких-то пределах. Единственное, что можно сделать, чтобы их определить — это провести эксперимент. Эксперимент для всей генеральной совокупности нереализуем или неоправдан, поэтому применяется выборочный метод.

На основании данных, полученных по выборке, делается вывод относительно всей генеральной совокупности. Используемые для этого методы теории статистических выводов обычно подразделяются на два класса: оценка параметров и проверка гипотез.

Задача оценки параметров состоит в получении наилучших в определенном смысле оценок параметров распределения генеральной совокупности на основании выборочных данных.

Проверка гипотез охватывает методы использования выборочных данных для проверки предположений относительно распределения и параметров распределения генеральной совокупности, которые делаются до получения выборочных данных.

В данной главе будут рассмотрены основные положения теории оценок.

5.2. Случайная выборка из генеральной совокупности

Чтобы по выборке можно было делать выводы о свойствах всей генеральной совокупности, она должна быть представительной (репрезентативной). Это обеспечивается в тех ситуациях, когда выборка является случайной. Модель случайной выборки предъявляет к ней следующие требования:

- 1) каждый из объектов, составляющих генеральную совокупность, должен иметь одинаковую вероятность быть представленным в выборке;
- 2) все n измерений, образующих выборку, должны быть независимыми, т. е. результаты каждого измерения не должны зависеть от предыдущих измерений.

Существует два основных метода отбора объектов из генеральной совокупности в выборку: повторный и бесповторный.

При повторном отборе каждый объект после измерения значения признака возвращается в генеральную совокупность. При этом состояние генеральной совокупности перед каждым новым измерением восстанавливается и требование независимости всегда выполняется.

При бесповторном отборе после измерения объект не возвращается в генеральную совокупность. В этом случае соотношение значений признака в оставшейся части генеральной совокупности меняется, и, следовательно, проводимые измерения не являются независимыми, т. е. бесповоротный отбор не является случайным. На практике бесповоротный отбор используется чаще. Когда проводится измерение каких-то признаков, относящихся, например, к преступникам, выборка составляется таким образом, что после того, как очередной человек принял участие в измерениях, он уже не участвует в следующих измерениях.

Но, как правило, можно считать, что объем генеральной совокупности настолько велик, что при исключении из нее относительно малого числа единиц, составляющих выборку, состояние генеральной совокупности практически не меняется. При бесконечной генеральной совокупности различие между

повторным и бесповторным отбором исчезает.

На практике используется несколько способов получения случайных выборок (см. 1.3.):

1. собственно случайная,
2. механический отбор.
3. типический отбор.
4. серийный отбор.

При проведении выборочных исследований предполагается, что выборка является однородной. Это означает, что она получена из одной генеральной совокупности, т. е. в исходной совокупности отсутствуют объекты, резко выделяющиеся по значениям изучаемого признака. Предположение об однородности выборки на практике обычно основывается на предварительном изучении условий эксперимента. Так, обычно есть уверенность в том, что полученные выборочные данные о количестве правонарушений представляют собой результаты измерений для одинаковых по численности городов.

5.3. Точечные оценки

Определения

Под термином “оценка” в теории оценок понимаются как сами значения параметров генеральной совокупности, полученные по выборке, так и процесс получения этих значений, т. е. правило, по которому они получены.

Оценки подразделяются на два класса

точечные	интервальные
конкретные значения параметров генеральной совокупности, полученные по выборочным данным. Естественно, что эти значения должны быть максимально близки к значениям соответствующих параметров генеральной совокупности, которые являются истинными значениями оцениваемых параметров.	определяют границы интервалов, между которыми с некоторой вероятностью p (так называемая доверительная вероятность) находится истинное значение оцениваемого параметра.

5.3.1. Требования к точечным оценкам

Начнем с точечных оценок и рассмотрим оценку произвольного параметра (среднего, дисперсии или какого-то другого) генеральной совокупности, который обозначим α . Оценивая параметр α по выборке, находим такую величину α_B , которую принимаем за точечную оценку параметра α . Естественно, при этом стремимся, чтобы оценка была в *определенном смысле* наилучшей, поэтому к ней предъявляется ряд требований:

1. **Состоятельность.** Точечная оценка α_B называется состоятельной, если при неограниченном увеличении объема выборки ($n \rightarrow \infty$) она стремится к истинному значению параметра α .

В математической статистике показывается, что состоятельной оценкой генерального среднего значения μ , является выборочное среднее арифметическое \bar{x}_B , а состоятельной оценкой генеральной дисперсии σ^2 — выборочная дисперсия σ_B^2 . Методы вычисления этих выборочных характеристик были рассмотрены в гл. 3.

2. **Несмещенность.** Оценка α_B называется несмещенной, если она не содержит систематической ошибки, т. е. среднее значение оценки, определенное по многократно повторенной выборке объема n из одной и той же генеральной совокупности, стремится к истинному значению соответствующего генерального параметра α .

Выборочное среднее арифметическое \bar{x}_B является несмещенной оценкой генерального среднего μ .

Несмещенной оценкой генеральной дисперсии σ^2 является *исправленная выборочная дисперсия*, вычисляемая по формуле:

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^2}{n-1} \quad \text{для несгруппированных данных,}$$

$$S^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^2 n_i}{n-1} \quad \text{для сгруппированных данных,}$$

Замечание 1

Одним из свойств выборочного среднего арифметического является то, что сумма квадратов отклонений значений признака от среднего арифметического меньше, чем сумма квадратов отклонений от любой другой величины (в том числе и от генерального среднего μ), т.е. $\sum (x_i - \bar{x}_B)^2 < \sum (x_i - \mu)^2$ для любой выборки. Поэтому вычисление оценки дисперсии по формуле $\sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2$ будет содержать систематическую ошибку, и такая оценка будет смещенной.

Можно показать, что если использовать $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_B)^2$, то она будет несмещенной, т.е. при неограниченном повторении выборки из генеральной совокупности и усреднении выборочной дисперсии, полученной на основании этой формулы, по всем выборкам получается истинное значение генеральной дисперсии.

3. Эффективность. Несмещенная оценка является эффективной, если она имеет наименьшую дисперсию по сравнению с другими несмещенными оценками того же параметра генеральной совокупности.

Это надо понимать так: полученные по выборке оценки \bar{x}_B и S^2 — случайные величины, так как случайны сами выборочные значения. Поэтому можно говорить о математическом ожидании и дисперсии оценок \bar{x}_B и S^2 . Эффективность этих оценок означает, что их дисперсии $D(\bar{x}_B)$ и $D(S^2)$ меньше дисперсий любых других несмещенных оценок среднего значения и дисперсии генеральной совокупности.

Итак, наилучшими в указанном смысле оценками генерального среднего значения и генеральной дисперсии являются выборочные характеристики \bar{x}_B , S^2 .

5.3.2. Стандартная ошибка среднего арифметического

Оценки \bar{x}_B и S^2 , полученные по выборке, естественно не совпадают с истинными значениями параметров μ и σ^2 генеральной совокупности. Экспериментально проверить это утверждение невозможно, поскольку не известны истинные значения этих параметров. Но если брать повторные выборки из одной и той же генеральной совокупности с параметрами μ и σ^2 и каждый раз вычислять их оценки \bar{x}_B и S^2 , то окажется, что эти оценки для разных выборок не совпадают, хотя все это из одних и тех же генеральных параметров.

Отклонения оценок генеральных параметров от истинных значений этих параметров называются статистическими ошибками, или ошибками репрезентативности. Их происхождение не имеет ничего общего с ошибками измерения, а возникают они только потому, что не все объекты генеральной совокупности представлены в выборке.

Величины статистических ошибок оценивают по среднему квадратическому (стандартному) отклонению выборочных характеристик. Здесь рассматривается только стандартное отклонение выборочного среднего арифметического.

Если взять очень много независимых выборок объема n из одной и той же генеральной совокупности и определить для каждой из них среднее арифметическое, то окажется, что полученные средние арифметические варьируют вокруг своего среднего значения (равного μ) в \sqrt{n} раз меньше, чем отдельные варианты выборки. Т.е. стандартное отклонение выборочного среднего арифметического будет равно

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

где σ — стандартное отклонение генеральной совокупности.

В качестве оценки стандартного отклонения выборочного среднего используется величина

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} \quad (5.1)$$

называемая стандартной ошибкой среднего арифметического. В формуле (5.1) S — выборочное стандартное отклонение $S = \sqrt{S^2}$.

Величина $S_{\bar{x}}$ показывает, какая ошибка в среднем допускается, если использовать вместо генерального среднего μ его выборочную оценку \bar{x}_B . Поэтому вычисленное среднее арифметическое

часто указывают в виде

$$\bar{x}_B \pm S_{\bar{x}}$$

чтобы оценить точность оценки x .

Из формулы (5.1) видно, как зависит стандартная ошибка $S_{\bar{x}}$ от объема выборки n : с увеличением объема выборки n стандартная ошибка $S_{\bar{x}}$ уменьшается пропорционально корню квадратному из n .

Замечание 2.

Теперь можно вернуться к вопросу, который был оставлен открытым при вычислении выборочных характеристик: “с такой точностью нужно вычислять выборочные характеристики?”

Как мы только что убедились, при ограниченном объеме выборки n истинное значение генерального среднего μ не может быть определено сколь угодно точно, поэтому при вычислении \bar{x}_B оставлять большое число значащих цифр не имеет смысла. Существует эмпирическое правило, согласно которому в окончательном результате положение последней значащей цифры должно соответствовать положению первой значащей цифры в величине $S_{\bar{x}}/3$. Чтобы избежать накопления ошибок, связанных с округлением, промежуточные результаты нужно вычислять с точностью на один порядок больше, чем точность окончательных результатов.

5.4. Интервальные оценки

По известной величине выборочной характеристики (\bar{x}_B или S^2 и др.) можно определить интервал, в котором с той или иной вероятностью определяется значение параметра генеральной совокупности, оцениваемого по этой выборочной характеристике.

Вероятности, признанные достаточными для того, чтобы уверенно судить о генеральных параметрах на основании выборочных характеристик, называются *доверительными*.

Обычно в качестве доверительных вероятностей выбирают значения 0,95, 0,99 или 0,999 (их принято выражать в процентах). Перечисленным значениям соответствуют 95%, 99% и 99,9 %. Выбор той или иной доверительной вероятности производится исследователем исходя из практических соображений о той ответственности, с какой делаются выводы о генеральных параметрах.

Замечание 3

Как правило, в научных исследованиях в области правоведения считается достаточной доверительная вероятность 0,95 (95%). В некоторых случаях, когда уточняются результаты предыдущих исследований или когда выводы, сделанные в данном исследовании, связаны с большой ответственностью (например, предлагается в корне пересмотреть результаты следствия), применяются более высокие уровни доверительной вероятности: 99 или 99,9 %.

Интервал, в котором с заданной доверительной вероятностью находится оцениваемый генеральный параметр, называется *доверительным интервалом*.

В соответствии с доверительными вероятностями на практике используются 95%, 99%, 99,9%-процентные доверительные интервалы.

В литературе по математической статистике обычно говорят о $100 \cdot (1 - \alpha)\%$ -процентном доверительном интервале, где $(1 - \alpha)$ — *доверительная вероятность*, а α — некоторое малое число ($\alpha = 0,05; 0,01; 0,001$), задающее вероятность того, что оцениваемый генеральный параметр выходит за границы доверительного интервала.

Теперь рассмотрим формирование доверительного интервала для среднего (математического ожидания) μ , нормально распределенной генеральной совокупности. Пронормируем значение среднего арифметического \bar{x}_B , найденного по выборке объема n из этой генеральной совокупности, по формуле:

$$t = \frac{\bar{x}_B - \mu}{S_{\bar{x}}}$$

где μ — оцениваемый параметр — среднее значение генеральной совокупности; $S_{\bar{x}}$ — стандартная ошибка выборочного среднего арифметического.

Величина t имеет Т-распределение Стьюдента (см. 4.12.2) с $\nu = n - 1$ степенями свободы.

Необходимо определить доверительный интервал, в котором с доверительной вероятностью $100(1 - \alpha)\%$ находится истинное значение оцениваемого параметра μ . Для этого задается значение α (например, 0,05). Доверительная вероятность будет соответствовать площади под кривой Т-распределения Стьюдента, заключенной между точками $-t_\alpha$ и t_α (рис. 5.1). Следовательно, доверительный интервал можно записать как

$$-t_{\alpha} \leq \frac{\bar{x}_B - \mu}{S_{\bar{x}}} \leq t_{\alpha}$$

Преобразуем это выражение к виду

$$\bar{x}_B - t_{\alpha} S_{\bar{x}} \leq \mu \leq \bar{x}_B + t_{\alpha} S_{\bar{x}} \quad (5.2)$$

Это и есть стандартная форма записи доверительного интервала.

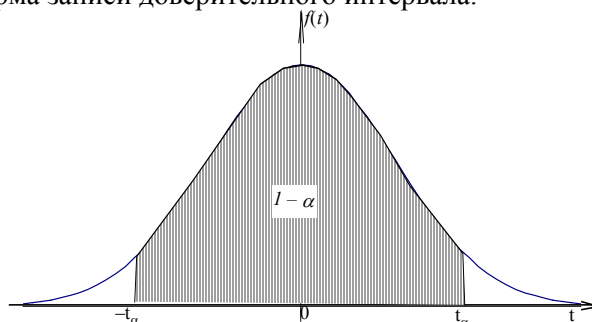


Рис. 5.1. К определению доверительного интервала

Учитывая формулу (5.1) приходим к окончательному выражению:

$$\bar{x}_B - t_{\alpha} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x}_B + t_{\alpha} \frac{S}{\sqrt{n}} \quad (5.3)$$

т. е. истинное значение μ с вероятностью $100(1 - \alpha)\%$ лежит в границах $\bar{x}_B - t_{\alpha} \frac{S}{\sqrt{n}}$ и $\bar{x}_B + t_{\alpha} \frac{S}{\sqrt{n}}$.

Значения t_{α} для стандартных значений α (0,05, 0,01 и 0,001) и различных значений параметра ν t-распределения ($\nu = n - 1$) приведены в специальных таблицах.

Чтобы найти границы доверительного интервала для среднего значения генеральной совокупности, действуем в следующем порядке:

- 1) по полученной выборке объема n вычисляем среднее арифметическое \bar{x}_B и стандартное отклонение S . Методы вычислений рассмотрены в гл. 3;
- 2) задаемся доверительной вероятностью $1 - \alpha$ (например, 0,95) исходя из целей исследования;
- 3) по таблице Т-распределения Стьюдента находим граничные значения t_{α} . В силу симметричности Т-распределения достаточно знать только положительное значение t_{α} . Например, если объем выборки $n=12$, то число степеней свободы Т-распределения $\nu = 12 - 1 = 11$, и по таблице Т-распределения определяем для $\alpha = 0,05$ значение $t_{0,05} = 2,20$;
- 4) находим границы доверительного интервала по формуле (5.3). Для $\alpha = 0,05$ и $n = 12$:

$$\bar{x}_B - 2,2 \frac{S}{\sqrt{12}} \leq \mu \leq \bar{x}_B + 2,2 \frac{S}{\sqrt{12}}$$

Как было отмечено в 4.12.2, при больших объемах выборки (практически при $n \geq 30$) t-распределение Стьюдента переходит в нормальное. Поэтому для определения границ доверительного интервала для μ при больших объемах выборки можно пользоваться таблицами нормированного нормального распределения (табл. 1 Приложения).

Доверительный интервал для μ при $n \geq 30$ записывается в следующем виде:

$$\bar{x}_B - u_{\alpha} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x}_B + u_{\alpha} \frac{S}{\sqrt{n}} \quad (5.4)$$

где u_{α} — процентные точки нормированного нормального распределения, определяемые по табл. 1 Приложения.

Для стандартных доверительных вероятностей (95%, 99%, 99,9%) значения u_{α} приведены в таблице 5.2.

Таблица 5.2

Значения u_α для стандартных доверительных вероятностей		
α	$1 - \alpha$	u_α
0,05	0,95	1,96
0,01	0,99	2,58
0,001	0,999	3,28

Чтобы найти доверительный интервал для среднего значения генеральной совокупности при больших объемах выборки ($n \geq 30$), поступаем следующим образом:

1. По выборочным данным находим среднее арифметическое \bar{x}_B и стандартное отклонение S , как показано в гл. 3.

2. Выбираем доверительную вероятность $1 - \alpha$ (например, 0,95).

3. По табл. 5.2 находим значение u_α , соответствующее заданной доверительной вероятности ($u_{0,05} = 1,96$).

4. Определяем границы доверительного интервала по формуле (5.4). Для $\alpha = 0,05$ получаем:

$$\bar{x}_B - 1,96 \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x}_B + 1,96 \frac{S}{\sqrt{n}}$$

Как видно из сравнения найденного доверительного интервала с доверительным интервалом, полученный выше по Т-распределению, при малых объемах выборки границы первого интервала шире ($t_{0,05} = 2,20$, а $u_{0,05} = 1,96$). Это понятно из простых физических соображений: при малом объеме выборки получается меньше информации о свойствах генеральной совокупности.

Сделаем еще одно замечание по поводу доверительных интервалов.

Среднее значение (μ , генеральной совокупности является хотя и неизвестным, но фиксированным параметром, а границы доверительного интервала, полученные по случайной выборке объема n , будут также случайными величинами. Когда говорится о 95-процентной доверительной вероятности, это означает, что примерно в 95 % случаев фиксированное, но неизвестное значение μ , окажется в границах доверительного интервала.

Образная трактовка доверительных интервалов приведена в книге «Статистика и планирование эксперимента в технике и науке»*. «Доверительный интервал и связанные с ним понятия похожи на то, с чем мы сталкиваемся при игре с набрасыванием подковы на кол. Кол здесь играет роль оцениваемого параметра (его положение никогда не изменяется)... Подкова выступает в роли доверительного интервала. Если при 100 набрасываниях подковы удастся в среднем 90 раз набросить ее на кол, то имеется 90 %-ная гарантия (или уровень доверия) набросить подкову на кол. Доверительный интервал, подобно подкове, меняет свое положение. При любом броске (или при построении некоторой интервальной оценки) кол (или параметр) может как попасть внутрь подковы (интервала), так и оказаться вне ее. Таким образом, делается вероятностное утверждение относительно переменных величин, характеризующих положение подковы».

Оценку параметра μ , найденную в форме доверительного интервала, часто записывают в виде $\bar{x}_B \pm t_\alpha S / \sqrt{n}$. Чтобы избежать неоднозначности в толковании результатов (перепутывания с записью результата как $\bar{x}_B \pm S_{\bar{x}}$), запись доверительного интервала необходимо сопровождать пояснением. Например 95%-ный доверительный интервал для среднего роста ($174\text{см} \pm 1,3\text{ см}$).

5.5. Определение необходимого объема выборки для получения оценок заданной точности

Обычно исследователя интересует вопрос: какой минимальный объем выборки необходим для того, чтобы оценка (чаще всего выборочное среднее арифметическое \bar{x}) отличалась от истинного значения среднего значения генеральной совокупности не более чем на заданную величину?

Ответить на этот вопрос можно, если ввести доверительную вероятность и выбрать объем выборки n

* Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке. — М.: Мир, 1980, с. 228—229.

таким образом, чтобы доверительный интервал имел заданный размер.

Если генеральная совокупность предполагается нормально распределенной и ее дисперсия σ^2 известна, то доверительный интервал для среднего значения μ записывается следующим образом:

$$\bar{x}_B - u_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x}_B + u_\alpha \frac{\sigma}{\sqrt{n}}$$

где u_α для стандартных доверительных вероятностей определены в табл. 5.2.

Пусть требуется, чтобы выборочное среднее \bar{x}_B отличалось от генерального μ , не более чем на заданную величину d . Это означает, что половина ширины доверительного интервала должна быть равна d , т. е. половина от

$$\left(\bar{x}_B + u_\alpha \frac{\sigma}{\sqrt{n}} \right) - \left(\bar{x}_B - u_\alpha \frac{\sigma}{\sqrt{n}} \right) = 2u_\alpha \frac{\sigma}{\sqrt{n}}$$

должна равняться d :

$$u_\alpha \frac{\sigma}{\sqrt{n}} = d$$

Отсюда требуемый объем выборки определяется следующим образом:

$$n = \left(\frac{u_\alpha \sigma}{d} \right)^2 \quad (5.5)$$

Истинное значение параметра σ генеральной совокупности обычно неизвестно, но при больших объемах выборки ($n \geq 30$) можно использовать его выборочную оценку S . Тогда

$$n = \left(\frac{u_\alpha S}{d} \right)^2 \quad (5.6)$$

6. КРИТЕРИИ ЗНАЧИМОСТИ И ПРОВЕРКА ГИПОТЕЗ

6.1. Введение

В этой главе рассматривается группа статистических методов, которые получили наибольшее распространение в статистических исследованиях, связанных с правоведением. Эти методы применяются всегда, когда предстоит проверить какие-то теоретические предположения, связанные с эффективностью мероприятий, направленных на совершенствование какого-либо процесса. Исследователь выдвигает предположения исходя из анализа конкретного явления с позиций педагогики, физиологии, медицины, психологии или другой области знаний, представителем которой он является. Затем справедливость предположений проверяется на основании данных соответствующего эксперимента, условия которого контролируются.

6.2. Основные понятия и определения

6.2.1. Нулевая гипотеза (нуль-гипотеза) и альтернатива (альтернативная гипотеза)

Статистической гипотезой (или просто гипотезой) называется утверждение о распределении генеральной совокупности, соответствующее некоторым представлениям об изучаемом явлении. В частном случае это может быть утверждение о значениях параметров (σ и μ) нормально распределенной генеральной совокупности.

Предположим, что в эксперименте участвуют две группы заключенных. Одна из них (контрольная) содержится и перевоспитывается по традиционной программе, а для второй (экспериментальная) используются новые методы. Действенность нового комплекса оценивается по различию уровня агрессивности, показанных в этих группах после определенного срока. По полученным данным необходимо проверить следующие утверждения:

1. Среднее значение уровня агрессивности не изменилось, т. е.

$$\mu_1 = \mu_2.$$

Здесь μ_1 и μ_2 – средние значения соответствующих генеральных совокупностей (уровни агрессивности всех аналогичных заключенных, которые могли бы перевоспитываться по традиционной (μ_1) и новой (μ_2) программам).

2. Вариативность агрессивности возросла:

$$\sigma_2 > \sigma_1.$$

Здесь σ_2 и σ_1 — так же, как и в п. 1, значения соответствующих генеральных параметров.

3. Средняя агрессивность возросла на 3 единицы:

$$\mu_1 - \mu_2 = 3$$

Это три различные статистические гипотезы. Конечно, возможные утверждения не ограничиваются приведенным списком. Гипотезы предстоит проверить с помощью какого-то метода — критерия.

Статистические гипотезы обычно рассматривают, генеральные совокупности, одна из которых может представлять собой теоретическую модель (например, нормальное распределение), а о второй судят по выборке из нее. В других случаях обе генеральные совокупности представлены выборками.

При проверке статистических гипотез принят следующий подход. Считается, что получение в результате эксперимента любых новых данных об изучаемом явлении, не согласующихся с данными, имеющимися до проведения эксперимента, — маловероятное событие. В то же время, если взять две выборки, представляющие собой результаты измерения одного и того же признака, и сравнить между собой их характеристики (среднее арифметическое, стандартное отклонение и др.), то окажется, что они практически всегда различаются. Это различие можно рассматривать как обусловленное только действием случайностей. Поэтому первоначально гипотезу всегда можно сформулировать таким образом: между двумя генеральными совокупностями нет ожидаемого различия.

Такая гипотеза называется *нулевой гипотезой*, или нуль-гипотезой. Обратное ей утверждение о том, что в действительности между генеральными совокупностями есть различие, называется *альтернативной гипотезой*, или альтернативой.

Нулевую гипотезу принято обозначать, как H_0 , а альтернативную — H_1 .

Итак, вначале выдвигается нулевая гипотеза о том, что различие между генеральными совокупностями равно нулю. Затем получают выборку или несколько выборок, и если выборочные данные не противоречат нулевой гипотезе, т. е. различие можно объяснить только случайностью выборки, то нулевая гипотеза

сохраняется (принимается). Если же полученные результаты не удается объяснить только действием случайных факторов, то нулевая гипотеза отвергается, а принимается альтернативная гипотеза.

Пусть, например, оценивается эффективность нового метода перевоспитания для заключенных по среднему значению агрессивности в контрольной и экспериментальной группах. Тогда нулевую гипотезу H_0 можно сформулировать так: среднее значение результатов не изменилось, т.е. $\mu_1 = \mu_2$. Для краткости это записывается так: $H_0: \mu_1 = \mu_2$.

Если заранее нельзя сказать, к чему приведет новый метод — к увеличению или уменьшению агрессивности, то альтернативная гипотеза H_1 будет состоять в том, что средние значения генеральных совокупностей неодинаковы: $H_1: \mu_1 \neq \mu_2$.

6.2.2. Ошибки при проверке гипотез

Ошибки, допускаемые при проверке гипотез, удобно разделить на два типа: 1) отклонение гипотезы H_0 , когда она верна, — ошибка первого рода; 2) принятие гипотезы H_0 , когда в действительности верна какая-то другая гипотеза, — ошибка второго рода.

Вероятность ошибки первого рода обозначается α . Величина α называется **уровнем значимости** критерия, по которому проверяется справедливость гипотезы H_0 .

Вероятность ошибки второго рода обозначается β . Ее величина зависит от альтернативной гипотезы H_1 . Рассмотрим для приведенного выше примера следующие две ситуации: 1) в действительности средняя агрессивность возросла на 3 единицы, 2) средняя агрессивность увеличилась на 30 единиц. Ясно, что для одних и тех же условий эксперимента и одинакового уровня значимости α вероятность ошибки второго рода β (принять гипотезу об отсутствии различия) для второй из альтернатив будет меньше.

Вероятности α и β удобно представить, как это сделано в табл. 6.1.

Таблица 6.1

Ошибки при проверке гипотез

	Решение	
	Принять H_0	Принять H_1
Справедлива H_0	Правильное с вероятностью $1 - \alpha$	Ошибочное с вероятностью α
Справедлива H_1	Ошибочное с вероятностью β	Правильное с вероятностью $1 - \beta$

Наглядным способом интерпретации ошибок является их графическое представление.

Предположим, что проверяется гипотеза $H_0: \mu_1 = \mu_0$ о равенстве среднего значения генеральной совокупности заданной величине μ_0 (известной, например, из предыдущих экспериментов).

Для этого берется выборка объема n , находится ее среднее арифметическое \bar{x}_B и по его величине судят о справедливости гипотезы H_0 .

Распределение среднего арифметического \bar{x}_B при условии, что верна гипотеза H_0 , будет $f(\bar{x}_B / \mu_0)$. Это распределение чисто качественно представлено на рис. 6.1.

Распределение среднего арифметического \bar{x}_B при условии, что верна альтернативная гипотеза $H_1: \mu_1 \neq \mu_0$, будет уже другим — $f(\bar{x}_B / \mu_1)$.

Будем считать, что гипотеза H_0 отвергается, если выборочное среднее арифметическое \bar{x}_B окажется больше некоторого значения $K_{\text{критич}}$, т. е. $\bar{x}_B > K_{\text{критич}}$, как показано на рис. 6.1.

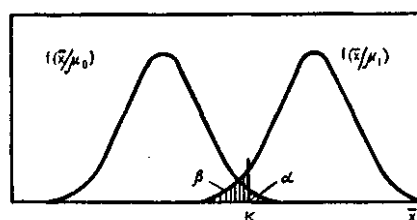


Рис. 6.1. Ошибки первого и второго рода

Область непринятия гипотезы H_0 называется **критической областью** критерия. Она показана на рис. 6.1 наклонной штриховкой. Уровень значимости α будет соответствовать площади критической области.

Вероятность ошибки второго рода β будет равна площади под кривой распределения $f(\bar{x}_B / \mu_1)$, показанной на рис. 6.1. вертикальной штриховкой.

Величина $1 - \beta$ называется **мощностью критерия**.

Следует особо подчеркнуть, что любая гипотеза должна формулироваться, а уровень значимости α задаваться исследователем, всегда до получения экспериментальных данных, по которым эта гипотеза будет проверяться.

При выборе уровня значимости α исследователь исходит из практических соображений, отвечая на вопрос: какую вероятность ошибки он считает допустимой для его конкретной задачи?

Обычно считают достаточным $\alpha = 0,05$ (5%), иногда $\alpha = 0,01$, редко $\alpha = 0,001$.

Между стандартными статистическими критериями и стандартными доверительными интервалами существует тесная связь: если принимается гипотеза о том, что значение параметра (μ , σ) нормально распределенной генеральной совокупности равно фиксированному значению (μ_0 , σ_0) с уровнем значимости α , то это эквивалентно заданию $100(1 - \alpha)\%$ -ного доверительного интервала для данного параметра нормального распределения. Поэтому оба подхода — доверительные интервалы и критерии значимости — в данном случае равноценны. Преимущество доверительных интервалов в том, что они дают представление об истинном значении параметра генеральной совокупности, а недостаток в том, что их трудно построить в более сложных случаях, например при анализе дисперсий (стандартных отклонений).

6.2.3. Критерии значимости

В рассмотренном выше примере (см. п. 6.2.2) при проверке гипотезы об отсутствии различия среднего уровня агрессивности в контрольной и экспериментальной группах можно было бы поступить следующим образом: вычислить средние арифметические результаты в группах и сравнить их между собой. Если окажется, что различие средних арифметических больше, например, 50 единиц, то можно утверждать, что новый комплекс оказался эффективным. Но при этом неизвестно, какие ошибки допускаются при таком утверждении, поэтому невозможно точно доказать наличие или отсутствие различий.

Методы, которые для каждой выборки формально точно определяются, удовлетворяют выборочные данные нулевой гипотезы или нет, называются *критериями значимости*.

Общая схема проверки гипотез

Процедура проверки гипотез обычно проводится по следующей схеме:

Формулируются гипотезы H_0 и H_1 .

Выбирается уровень значимости критерия α .

По выборочным данным вычисляется значение $K_{набл}$ некоторой случайной величины K , называемой статистикой критерия, или просто критерием, который имеет известное стандартное распределение (нормальное, Т-распределение Стьюдента и т.п.)

Вычисляется критическая область и область принятия гипотезы. То есть находится критическое (граничное) значение критерия $K_{критич}$ при уровне значимости α , взятым из соответствующих таблиц.

Найденное значение $K_{набл}$ критерия сравнивается с $K_{критич}$ и по результатам сравнения делается вывод: принять гипотезу или отвергнуть.

♦ Если вычисленное по выборке значение критерия $K_{набл}$ меньше чем $K_{критич}$, то гипотеза H_0 принимается на заданном уровне значимости α .

(В этом случае наблюдаемое по экспериментальным данным различие генеральных совокупностей можно объяснить только случайностью выборки. Однако принятие гипотезы H_0 совсем не означает доказательства равенства параметров генеральных совокупностей. Просто имеющийся в распоряжении статистический материал не дает оснований для отклонения гипотезы о том, что эти параметры одинаковы. Возможно, появится другой экспериментальный материал, на основании которого эта гипотеза будет отклонена.)

♦ Если вычисленное значение критерия $K_{набл}$ больше $K_{критич}$, то гипотеза H_0 отклоняется в пользу гипотезы H_1 при данном уровне значимости α .

(В этом случае наблюдаемое различие генеральных совокупностей уже нельзя объяснить только случайностями и говорят, что наблюдаемое различие значимо (статистически значимо) на уровне значимости α .)

Следует подчеркнуть разницу между статистической значимостью и практической значимостью. Заключение о практической значимости всегда делается человеком, изучающим данное явление. И здесь истинным критерием является опыт и интуиция исследователя, а статистические критерии значимости — лишь формально точный инструмент, используемый в исследовании. Чем больше исследователь знает об изучаемом явлении, тем точнее будет сформулированная им гипотеза и тем точнее будут выводы,

сделанные с помощью критериев значимости.

Замечание 1

Ранее уже подчеркивалось, что уровень значимости и должен выбираться исследователем до получения экспериментальных данных, по которым будет проверяться гипотеза. Но часто с предварительным выбором возникают затруднения. Обычно говорят, что для научных исследований (в том числе и в правоведении) достаточен уровень значимости $\alpha=0,05$, но если выводы, которые предстоит сделать по результатам проверки гипотез, связаны с большой ответственностью, то рекомендуется выбирать $\alpha=0,01$ или $\alpha=0,001$.

Как установить ответственность в трактовке результатов эксперимента и тот риск, который связан с выбором уровня значимости α ? Чтобы не давать прямых ответов на эти непростые вопросы, часто поступают следующим образом: уровень значимости до эксперимента точно не устанавливается, а по экспериментальным данным вычисляется вероятность P того, что критерий (статистика критерия) выйдет за пределы значения, рассчитанного по выборке. Таким образом, P — это экспериментальный уровень значимости. Точное значение P обычно не указывают, а окончательные результаты приводят в следующем виде: 1) если вычисленное значение критерия не превосходит критического значения на уровне значимости $\alpha=0,05$, то различие считается статистически незначимым; 2) если вычисленное по выборке значение критерия превышает критические значения при $\alpha=0,05$, $\alpha=0,01$ или $\alpha=0,001$, то записывают $P<0,05$, $P<0,01$ или $P<0,001$. Это означает, что наблюдаемые различия статистически значимы на уровнях значимости 0,05, 0,01 или 0,001.

Критерии значимости подразделяются на три типа:

1. Критерии значимости, которые служат для проверки гипотез о параметрах распределений генеральной совокупности (чаще всего нормального распределения). Эти критерии называются параметрическими.

2. Критерии, которые для проверки гипотез не используют предположений о распределении генеральной совокупности. Эти критерии не требуют знания параметров распределений, поэтому называются непараметрическими.

3. Особую группу критериев составляют критерии согласия, служащие для проверки гипотез о согласии распределения генеральной совокупности, из которой получена выборка, с ранее принятой теоретической моделью (чаще всего нормальным распределением).

6.2.4. Односторонние и двусторонние критерии

Если цель исследования том, чтобы выявить различие параметров двух генеральных совокупностей, которые соответствуют различным ее естественным условиям (условия жизни, возраст испытуемых и т. п.), то часто неизвестно, какой из этих параметров будет больше, а какой меньше. Например, если интересуются вариативностью результатов в контрольной и экспериментальной группах, то, как правило, нет уверенности в знаке различия дисперсий или стандартных отклонений результатов, по которым оценивается вариативность. В этом случае нулевая гипотеза состоит в том, что дисперсии равны между собой ($H_0 : \sigma_1^2 = \sigma_2^2$), а цель исследования — доказать обратное ($H_1 : \sigma_1^2 \neq \sigma_2^2$), т.е. наличие различия между дисперсиями. При этом допускается, что различие может быть любого знака. Такие гипотезы называются двусторонними.

Но иногда задача состоит в том, чтобы доказать увеличение или уменьшение параметра; например, средний результат в экспериментальной группе выше, чем контрольной. При этом уже не допускается, что различие может быть другого знака. Тогда альтернативная гипотеза $H_1 : \mu_2 > \mu_1$ (или $H_1 : \mu_2 < \mu_1$ а обратное ей утверждение $H_0 : \mu_2 \leq \mu_1$ (или $H_0 : \mu_2 \geq \mu_1$). Такие гипотезы называются односторонними.

Критерии значимости, служащие для проверки двусторонних гипотез, называются двусторонними, а для односторонних — односторонними.

Возникает вопрос о том, какой из критериев следует выбирать в том или ином случае. Ответ на этот вопрос находится за пределами формальных статистических методов и полностью зависит от целей исследования. Ни в коем случае нельзя выбирать тот или иной критерий после проведения эксперимента на основе анализа экспериментальных данных, поскольку это может привести к неверным выводам. Если до проведения эксперимента допускается, что различие сравниваемых параметров может быть как положительным, так и отрицательным, то следует использовать двусторонний критерий. Если же есть дополнительная информация, например, из предшествующих экспериментов, на основании которой можно сделать предположение, что один из параметров больше или меньше другого, то используется односторонний критерий. Когда имеются основания для применения одностороннего критерия, его следует предпочесть двустороннему, потому что односторонний критерий полнее использует информацию об

изучаемом явлении и поэтому чаще дает правильные результаты.

Например, необходимо доказать различие средних значений генеральных совокупностей (средних значений некоторого результата исследований) при двух различных методиках применяемых в контрольной и экспериментальной группах. Если есть данные, что экспериментальная группа покажет в среднем лучший результат, то нужно выдвинуть нулевую гипотезу $H_0: \mu_2 = \mu_1$ против двусторонней альтернативы $H_1: \mu_2 \neq \mu_1$. Различие доказывается по разности средних арифметических результатов в контрольной и экспериментальной группах $(\bar{x}_2 - \bar{x}_1)$. Распределение разности $\bar{x}_2 - \bar{x}_1$ при условии, что верна нулевая гипотеза H_0 схематично представлено на рис. 6.2, а.

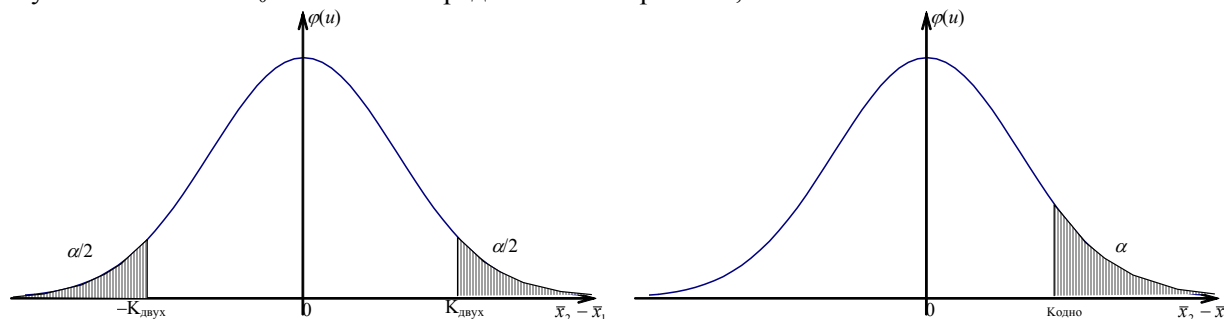


Рис. 6.2. Уровни значимости при двустороннем (а) и одностороннем (б) критериях

Решение об отклонении гипотезы H_0 принимается в том случае, если разность $\bar{x}_2 - \bar{x}_1$ выходит за пределы некоторого значения $K_{двух}$ (допустимы отклонения в обе стороны от нуля). Ошибка, которая при этом допускается, равна, как известно, уровню значимости α . Но поскольку отклонения возможны в обе стороны, то при симметричном распределении вероятности отклонений, больших $K_{двух}$ и меньших $K_{двух}$, будут одинаковы и составят $\alpha/2$.

Если предположить, что в экспериментальной группе будут показаны в среднем более высокие результаты, то можно выдвинуть одностороннюю альтернативу $H_1: \mu_2 > \mu_1$. В этом случае при той же нулевой гипотезе $H_0: \mu_2 = \mu_1$ распределение разности $\bar{x}_2 - \bar{x}_1$ будет таким же, как и для двустороннего критерия (см. рис. 6.2, б). Но теперь представляют интерес только положительные значения разности $\bar{x}_2 - \bar{x}_1$. Решение об отклонении H_0 принимается, когда $\bar{x}_2 - \bar{x}_1$ окажется больше некоторого $K_{одно}$. При том же уровне значимости α $K_{одно}$ будет всегда меньше $K_{двух}$, поэтому нулевая гипотеза будет при одностороннем критерии отклоняться чаще.

Таким образом, двусторонние критерии оказываются более консервативными, чем односторонние.

В этом нет никакого противоречия или доказательства несостоятельности статистических методов. Просто в первом случае, используя двустороннюю гипотезу, мы допускали и отрицательный эффект новой программы. В такой ситуации выводы должны быть более осторожными, чем в случае односторонней гипотезы, когда имеется дополнительная информация, позволяющая сделать предположение о положительном эффекте новой программы, что, естественно, дает возможность сделать более точный вывод. Правда, следует отметить, что если превышение критического значения в каком либо исследовании незначительно, то в достоверности вывода о наличии положительного эффекта можно усомниться. В такой ситуации следует провести дополнительные исследования.

6.3. Критерии согласия

Все рассмотренные ниже критерии значимости являются оптимальными, т. е. обеспечивают наивысшую достоверность статистических выводов только в тех случаях, когда выборки получены из *нормально распределенной генеральной совокупности*. При отклонениях от нормального распределения точность оптимальных критериев существенно падает, поэтому, чтобы уверенно применять оптимальные критерии, необходимо проверить предположение о нормальном распределении генеральной совокупности. Для этого используются критерии согласия. Здесь нулевая гипотеза H_0 представляет собой утверждение о том, что распределение генеральной совокупности, из которой получена выборка, не отличается от нормального. Существует несколько разновидностей критериев согласия. Рассмотрим те из них, которые получили наибольшее распространение на практике.

6.3.1. Предварительная проверка соответствия нормальному распределению

Критерии согласия требуют достаточно большой вычислительной работы, поэтому целесообразно перед тем, как их использовать, проверить с помощью более простых методов соответствие имеющихся экспериментальных данных нормальному распределению. Эти методы, естественно, обладают меньшей

мощностью и позволяют установить только значительные расхождения с нормальным распределением, но если такие расхождения будут установлены, то необходимость в применении более точных, но более сложных критериев, как правило, отпадает.

Для предварительной проверки эмпирического распределения на нормальность можно использовать основные свойства нормального распределения, изложенные ранее. При этом эмпирическое распределение представляется в виде вариационного ряда или гистограммы. Если в качестве параметров μ и σ нормального распределения принять их выборочные оценки \bar{x}_B и S , то для проверки можно использовать следующие свойства нормального распределения: 1) практически все отклонения от среднего значения (99,7 %) должны быть меньше $\pm 3S$; 2) примерно 2/3 всех отклонений (68,3 %) должны быть меньше $\pm S$; 3) половина всех отклонений от среднего значения должна быть меньше $\pm 0,657S$; 4) можно использовать такое свойство нормального распределения, что его коэффициенты асимметрии γ_3 и эксцесса γ_4 равны нулю.

Значения коэффициентов A_s и E_x сравниваются с критическими значениями на уровне значимости α , и если критические значения превышены, то делается вывод о том, что распределение генеральной совокупности, из которой получена выборка, не согласуется с нормальным. В противном случае модель нормального распределения может быть принята. Таблица критических значений A_s и E_x содержится в различных книгах. Здесь не будем подробно останавливаться на этих приближенных критериях. Отметим лишь еще раз, что они могут использоваться только совместно с более точными критериями, рассмотренными ниже.

6.3.2. Критерий согласия χ^2 (хи-квадрат)

Критерий согласия χ^2 разработан лучше других критериев и чаще других используется. Он основан на сравнении эмпирических частот интервалов группировки с теоретическими (ожидаемыми) частотами, рассчитываемыми по формулам нормального распределения.

Условия применения: объем выборки $n \geq 40$, выборочные данные сгруппированы в интервальный вариационный ряд с числом интервалов не менее 7, ожидаемые (теоретические) частоты интервалов не должны быть меньше 5.

Гипотеза H_0 : $f(x) = f^{теор}(x)$ — плотность распределения $f(x)$ генеральной совокупности, из которой взята выборка, соответствует теоретической модели $f^{теор}(x)$ нормального распределения.

Альтернатива H_1 : $f(x) \neq f^{теор}(x)$.

Уровень значимости: α .

Порядок, применения:

1. Формулируется гипотеза, выбирается уровень значимости α .
2. Получается выборка объема $n \geq 40$ независимых наблюдений и представляется эмпирическое распределение в виде интервального вариационного ряда.
3. Рассчитываются выборочные характеристики \bar{x}_B и S . Их используют в качестве генеральных параметров μ и σ нормального распределения, с которым предстоит сравнить эмпирическое распределение.
4. Вычисляются значения теоретических частот $n_i^{теор}$ попадания в i -й интервал группировки. Для этого необходимо вычислить:

$$n_i^{теор} = n \left[\Phi_0 \left(\frac{x_{vi} - \bar{x}_B}{S} \right) - \Phi_0 \left(\frac{x_{ni} - \bar{x}_B}{S} \right) \right]$$

где $\Phi_0(u)$ — функции Лапласа, x_{vi} и x_{ni} — верхняя и нижняя границы i -го интервала группировки.

Если окажется, что вычисленные ожидаемые частоты $n_i^{теор}$ некоторых интервалов группировки меньше 5, то соседние интервалы объединяются так, чтобы сумма их ожидаемых частот была больше или равна 5. Соответственно складываются и эмпирические частоты объединяемых интервалов.

5. Значение χ^2 -критерия рассчитывается по формуле:

$$\chi_{набл}^2 = \sum_{i=1}^k \frac{(n_i - n_i^{теор})^2}{n_i^{теор}}$$

где n_i — эмпирические частоты; $n_i^{теор}$ — ожидаемые (теоретические) частоты; k — число интервалов группировки после объединения.

6. Из таблиц распределения χ^2 находится критическое значение $\chi_{критич}^2$ критерия для уровня

значимости α и числа степеней свободы $\nu=n-3$.

7. Вывод: если $\chi^2_{\text{набл}} \geq \chi^2_{\text{критич}}$ то эмпирическое распределение не соответствует нормальному распределению на уровне значимости α , в противном случае нет оснований отрицать это соответствие.

6.4. Критерии, основанные на нормальном распределении

6.4.1. Сравнение выборочного среднего арифметического с известным средним значением генеральной совокупности

Рассмотрим, как с помощью статистических критериев решить вопрос: значимо ли отличие выборочного среднего значения от среднего значения генеральной совокупности, из которой предположительно взята выборка, или наблюдаемое различие является случайным? Такая постановка вопроса типична для выборочного контроля качества продукции в промышленности, но и при исследованиях в правоведении такой вопрос часто возникает, когда предстоит решить, значимо ли отличается среднее значение признака, полученное по выборке, от среднего значения, известного по результатам многочисленных предыдущих экспериментов.

Применяемый для этих целей t-критерий Стьюдента основан на предположении о нормальности распределения генеральной совокупности, но результаты проверки гипотез удовлетворяют по точности и при небольших отклонениях от нормальности распределения.

Условия применения t-критерия: выборка получена из генеральной совокупности, имеющей приближенно нормальное распределение с параметрами μ и σ .

Гипотеза $H_0: \mu = \mu_0$ – среднее значение μ генеральной совокупности, из которой получена выборка, равно данному значению μ_0 (известному, например, из предыдущих экспериментов).

Альтернатива $H_0: \mu \neq \mu_0$ (двусторонний критерий применяется тогда, когда допускаются отклонения в обе стороны от μ_0).

Уровень значимости: α .

Порядок применения Т-критерия:

1. Принимается предположение о нормальности, формулируются гипотезы H_0 и H_1 задается уровень значимости α .
2. Получают выборку объема n .
3. Вычисляется выборочное среднее арифметическое \bar{x}_B и исправленная выборочная дисперсия S^2 .
4. Определяется значение t-критерия по формуле:

$$t_{\text{набл}} = \frac{|\bar{x}_B - \mu_0|}{S / \sqrt{n}}$$

Величина t имеет при справедливости гипотезы H_0 t-распределение Стьюдента (определенное в гл. 4) с $\nu=n-1$ степенями свободы.

5. По таблицам находится $t_{\text{критич}}$ – критическое значение t-критерия при уровне значимости α и числе степеней свободы $\nu=n-1$. (Таблицы обычно содержат критические значения $t_{\text{критич}}$ для двустороннего критерия.)

6. Делается вывод: если $t_{\text{набл}} \geq t_{\text{критич}}$, то выборочное среднее значимо отличается от μ_0 на уровне значимости α , и в этой ситуации отклоняется гипотеза H_0 , т. е. считается, что выборка взята из другой генеральной совокупности, для которой $\mu \neq \mu_0$. Если $t_{\text{набл}} < t_{\text{критич}}$, то на заданном уровне различие незначимо и сохраняется гипотеза H_0 .

Замечание 2.

При больших объемах выборки ($n \geq 30$), как указано и гл. 4, t-распределение переходит в нормированное нормальное распределение, поэтому при проверке гипотезы вместо t-критерия можно использовать U-критерий, основанный на нормированном нормальном распределении статистики критерия.

В этом случае вычисляют величину

$$U_{\text{набл}} = \frac{|\bar{x}_B - \mu_0|}{S / \sqrt{n}}$$

и сравнивают ее с критическими значениями u_α нормированного нормального распределения. Для стандартных уровней значимости значения u_α приведены в табл. 6.2.

Замечание 1.

Если перед проведением эксперимента известно не только среднее значение μ генеральной совокупности, из которой получена выборка, но и его дисперсия σ , то нет необходимости в вычислении S . Кроме того, при проверке гипотезы вместо t -критерия можно использовать Z -критерий, основанный на нормированном нормальном распределении статистики критерия.

В этом случае вычисляют величину

$$Z_{\text{набл}} = \frac{|\bar{x}_B - \mu_0|}{\sigma / \sqrt{n}}$$

и сравнивают ее с критическими значениями z нормированного нормального распределения. Для стандартных уровней значимости значения z приведены в табл. 6.2.

Таблица 6.2.

Критические значения U -критерия

Уровень значимости, α	Критическое значение, U_α	
	Двусторон ний критерий	Односторо нный критерий
0,05	1,96	1,64
0,01	2,58	2,33
0,001	3,28	3,09

6.4.2. Сравнение двух выборочных средних значений для связанных выборок

Существует много практических задач, в которых две сравниваемые выборки взаимосвязаны в силу особенностей организации эксперимента или просто потому, что этой взаимосвязи нельзя избежать.

В практике медицинских, биологических и педагогических исследований часто используются так называемые парные сравнения. Один из методов таких сравнений заключается в том, что измерения проводятся для одной и той же группы испытуемых до и после применения интересующих исследователя воздействий. Результаты парных сравнений всегда точнее, чем сравнения на независимых группах, и объясняется это тем, что разброс результатов внутри группы испытуемых всегда больше, чем разброс разностей результатов, полученных при повторных измерениях для одних и тех же индивидуумов. Это можно пояснить на следующем простом примере. Допустим, необходимо по измеренному уровню IQ установить влияние на студентов какого-то вида учебной нагрузки. Конечно, можно было бы провести такой эксперимент на двух независимых однородных группах: в одной из них определить среднее значение IQ до чтения курса лекций, а в другой – после. Но и без точных математических доказательств ясно, что выводы будут точнее, если измерения IQ провести у одних и тех же студентов до и после чтения курса лекций. Поэтому парные сравнения всегда выгодно использовать, конечно, если удастся организовать эксперимент так, что будет устранено влияние мешающих факторов (усталость, эффект обучения и т. п.).

При парных сравнениях нельзя использовать рассмотренные выше методы для независимых выборок, поскольку это приведет к большим ошибкам.

Для сравнения средних значений здесь используется модификация T -критерия для связанных выборок. Особенность его в том, что гипотеза формулируется в отношении разностей d , сопряженных пар наблюдений.

Условия применения: $d_i = x_i - y_i$ – разность связанных пар результатов измерения. Делается предположение о нормальном распределении этих разностей в генеральной совокупности с параметрами μ_d, σ_d .

Гипотеза H_0 : $\mu_d = 0$.

Альтернатива H_1 : $\mu_d \neq 0$ (для двустороннего критерия). Можно сформулировать и одностороннюю альтернативу, например, H_1 : $\mu_d > 0$.

Уровень значимости: α .

Порядок применения:

1. Делается предположение о нормальном распределении разностей d_i , формулируется гипотеза H_0 и альтернатива H_1 выбирается уровень значимости α .

2. Получают две выборки объема n , представляющие собой ряды связанных пар наблюдений.

3. Вычисляются среднее арифметическое \bar{d}_B и исправленная выборочная дисперсия S_d^2 .

4. Определяется значение t-критерия:

$$t_{\text{набл}} = \frac{\bar{d}_B}{S_d / \sqrt{n}}. \quad (6.1)$$

5. Из таблицы t-распределения Стьюдента находится $t_{\text{критич}}$ – критическое значение t-критерия при заданном уровне значимости α и числе степеней свободы $\nu = n - 1$.

6. Делается вывод: если $t_{\text{набл}} \geq t_{\text{критич}}$, то наблюдаемое различие значимо на уровне значимости α , в противном случае различие статистически незначимо.

При больших выборках (для $n > 30$) вместо t-критерия можно использовать u-критерий. В этом случае вычисленное по формуле (6.1) значение $t_{\text{набл}}$ сравнивается с критическим значением $u_{\text{критич}}$ нормированного нормального распределения (см. табл. 6.2).

6.4.3. Сравнение двух выборочных дисперсий из нормальных совокупностей

Условия применения F-критерия: обе выборки независимы и получены из нормально распределенных генеральных совокупностей с параметрами μ_x, σ_x и μ_y, σ_y .

Гипотеза $H_0: \sigma_x^2 = \sigma_y^2$.

Альтернатива $H_1: \sigma_x^2 \neq \sigma_y^2$.

(Это двусторонняя гипотеза, поэтому следует применять двусторонний критерий. Если же предположить, что одна из генеральных совокупностей имеет большую дисперсию (обозначим ее σ_1^2), чем другая (σ_2^2), то можно сформулировать одностороннюю гипотезу $H_1: \sigma_x^2 > \sigma_y^2$, и тогда применяется односторонний F-критерий.)

Уровень значимости критерия задается α .

Порядок применения F-критерия:

1. Принимается предположение о нормальности распределения генеральных совокупностей, формулируется гипотеза и альтернатива, назначается уровень значимости α , как указано выше.

2. Получают две независимые выборки из совокупностей X и Y объемом n_x и n_y соответственно.

3. Рассчитываются значения исправленных выборочных дисперсий S_x^2 и S_y^2 . Большую из дисперсий (S_x^2 или S_y^2) обозначают S_1^2 , меньшую — S_2^2 .

4. Вычисляется значение F-критерия по формуле:

$$F_{\text{набл}} = \frac{S_1^2}{S_2^2}$$

5. Сравнивается вычисленное значение $F_{\text{набл}}$ с критическим значением $F_{\text{критич}}$ при заданном уровне значимости α и числе степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$.

(Критические значения F при уровнях значимости α , равных 0,05, 0,01, 0,001 приведены в таблицах).

Отметим, что в обычно в таблице приведены критические значения *одностороннего* F-критерия. Поэтому если цель исследования доказать, что одна дисперсия больше другой ($H_1: \sigma_1^2 > \sigma_2^2$), то критические значения берутся непосредственно из этой таблицы. Если же применяется двусторонний критерий ($H_1: \sigma_1^2 \neq \sigma_2^2$), то критические значения, взятые из таблиц, соответствуют удвоенным уровням значимости: 0,01, 0,02 и 0,002.

6. Делается вывод: если вычисленное значение $F_{\text{набл}}$ больше или равно критическому $F_{\text{критич}}$, то дисперсии различаются значимо на заданном уровне значимости. К противном случае нет оснований для отклонения нулевой гипотезы о равенстве двух дисперсий.

Следует отметить, что F-критерий очень чувствителен к отклонениям от нормальности распределения генеральной совокупности. Если предположение о нормальном распределении не может быть принято, то F-критерий применять не следует. В этом случае используются непараметрические методы.

F-критерий используется для малых и средних объемов выборки ($n < 100$). Для больших объемов выборки ($n > 100$) при проверке гипотезы о равенстве дисперсий удобнее применять U-критерий. В этом случае вычисляется величина

$$U = \frac{|S_x - S_y|}{\sqrt{\frac{S_x^2}{2n_x} - \frac{S_y^2}{2n_y}}}$$

и сравнивается с критическими значениями U , взятыми из таблиц нормированного нормального распределения. Для стандартных уровней значимости значения u приведены в табл. 6.2.

6.4.4. Сравнение двух выборочных средних значений для независимых выборок

В этом разделе рассматривается очень важный для практики спорта критерий математической статистики, позволяющий получить ответ на вопрос: значимо ли различаются средние значения, полученные по двум независимым выборкам (например, по результатам в контрольной и экспериментальной группах)? Здесь также применяется t -критерий Стьюдента, основанный на предположении, что выборки получены из генеральных совокупностей, имеющих приблизительно нормальное распределение. Кроме того, применение t -критерия отличается при различных предположениях относительно дисперсий этих генеральных совокупностей. В математической статистике обычно рассматриваются случаи известных и неизвестных генеральных дисперсий, но, поскольку на практике генеральные дисперсии, как правило, неизвестны, здесь описывается *только случай неизвестных дисперсий*. При этом возможны следующие варианты предположений:

- 1) обе дисперсии неизвестны, но предполагается, что они равны между собой;
- 2) обе дисперсии неизвестны, и предположение о их равенстве не делается.

Как выбрать подходящий вариант? Конечно, если нет уверенности в равенстве дисперсий, нужно использовать второй вариант, потому что в этом случае требуется меньше знаний о распределении генеральных совокупностей, но всегда платой за это является меньшая точность выводов.

Поэтому обычно поступают следующим образом: вначале по имеющимся выборочным данным проверяют гипотезу о равенстве дисперсий, используя F -критерий, а затем уже выбирают тот или иной вариант t -критерия. Строго говоря, это некорректно с точки зрения математической статистики, поскольку, как уже неоднократно подчеркивалось, критерий должен выбираться до получения экспериментальных данных, и правильнее было бы выбрать предположение о равенстве или неравенстве дисперсий по другим, предварительно полученным экспериментальным данным.

При описанном выше подходе t -критерий применяется следующим образом.

Условия применения: обе выборки независимы и получены из генеральных совокупностей X и Y , имеющих нормальное распределение с параметрами μ_x, σ_x и μ_y, σ_y .

Гипотеза H_0 : $\mu_x = \mu_y$.

Альтернатива H_1 : $\mu_x \neq \mu_y$ или H_1 : $\mu_x > \mu_y$ ($\mu_x < \mu_y$) в зависимости от того, что требуется доказать: простое различие средних значений или то, что одно из них больше другого.

Уровень значимости: α .

Порядок применения:

1. Принимается предположение о нормальности, формулируются гипотеза H_0 и альтернатива H_1 , задается уровень значимости α .

2. Получают две независимые выборки из совокупностей X и Y объемом n_x и n_y .

3. Вычисляются выборочные характеристики S_x и S_y методами, рассмотренными в гл. 3.

4. Используется F -критерий для проверки гипотезы о равенстве генеральных дисперсий, как показано в разделе 6.3.1.

5. По результатам применения F -критерия принимается или не принимается предположение о равенстве дисперсий.

6. Вычисляются значение t -критерия и число степеней свободы v . Применяемые для этого формулы приведены в табл. 6.3, они различаются в зависимости от предположения о дисперсиях и соотношения между объемами выборок n_x и n_y .

7. Из таблицы t -распределения Стьюдента находится $t_{\text{критич}}$ – критическое значение t -критерия при заданном уровне значимости α и числе степеней свободы v .

8. Делается вывод: если $t_{\text{набл}} \geq t_{\text{критич}}$, то выборочные средние значимо различаются на уровне значимости α (вероятность ошибки меньше α). В противном случае различие статистически незначимо.

Формулы для вычисления значения t-критерия

№	Предположения о дисперсиях σ_1^2 и σ_2^2	Объемы выборок n_x и n_y	Формула t-критерия	Стандартная ошибка разности $S_{\bar{x}-\bar{y}}$	Число степеней свободы ν
1	$\sigma_1^2 = \sigma_2^2$	$n_x = n_y = n$	$t_{набл} = \frac{ \bar{x} - \bar{y} }{S_{\bar{x}-\bar{y}}}$	$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2 + S_y^2}{n}}$	$\nu = 2n - 2$
2	$\sigma_1^2 = \sigma_2^2$	$n_x \neq n_y$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{n_x + n_y}{n_x n_y}} \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$	$\nu = n_x + n_y - 2$
3	$\sigma_1^2 \neq \sigma_2^2$	$n_x = n_y = n$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2 + S_y^2}{n}}$	$\nu = (n - 1) \frac{(S_x^2 + S_y^2)^2}{S_x^4 + S_y^4}$
4	$\sigma_1^2 \neq \sigma_2^2$	$n_x \neq n_y$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}$	$\nu = \frac{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}\right)^2}{\frac{S_x^4}{n_x^2(n_x - 1)} + \frac{S_y^4}{n_y^2(n_y - 1)}}$