РАЗДЕЛ III. ПОНЯТИЯ И ПРИНЦИПЫ СТАТИСТИЧЕСКОГО ВЫВОДА КАК ОСНОВА АНАЛИЗА СОЦИОЛОГИЧЕСКИХ ДАННЫХ

3.1. Кривые распределения, законы распределения

Понятие нормального частотного распределения. Можно сказать, что основной целью анализа вариационных рядов является выявление закономерности распределения (исключая при этом влияние случайных для данного распределения факторов). Если попробовать каждое значение признака отобразить в виде точки на плоскости координат и соединить все эти значения плавной линией, мы получим некую кривую, которая для полигона частот будет являться некоторым пределом. Эту линию называют кривой распределения [1, с. 115−119; 10].

Кривая распределения есть графическое изображение в виде непрерывной линии изменения частот в вариационном ряду, которое функционально связано с изменением вариант. Кривая распределения отражает закономерность изменения частот при отсутствии случайных факторов. Графическое изображение облегчает анализ рядов распределения [1, с. 138−144].

Любое эмпирические распределения сравнивают с неким эталоном − идеальным распределением. Это сравнение необходимо для:

1) возможности спрогнозировать дальнейшее поведение и развитие того или иного феномена, в случае если различия между эмпирическими и теоретическими распределениями невелики;

2) выявления причин, влияющих на проявление отличий между теоретическими и эмпирическими распределениями, если таковые наблюдаются.

Известно достаточно много форм кривых распределения, по которым может выравниваться вариационный ряд, но в практике статистических исследований наиболее часто используются такие формы, как *нормальное распределение* и *распределение Пуассона.*

Нормальное распределение зависит от двух параметров: среднего арифметического *M[X]* и среднего квадратического отклонения σ. Можно даже сказать, что одно из наиболее важных применений последнего и состоит в описании нормального распределения, то есть «нормальности» распределения, если можно так выразиться.

Применение количественных методов, вошедших в практику социологических исследований, в той или иной степени опирается на предположение, что изучаемый признак (или совокупность признаков) подчиняется определенному статистическому закону распределения. Таким наиболее часто встречающимся распределением и является нормальный (гауссовский) закон распределения. Нормальное распределение признака наблюдается в тех случаях, когда на величину его значений влияет множество случайных независимых или слабо зависимых факторов, каждый из которых играет в общей сумме примерно одинаковую и малую роль (то есть отсутствуют доминирующие факторы).

Функция плотности нормального распределения имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e \frac{1}{2}\left(\frac{x - M[X]}{\sigma}\right)^2,$$

где ● *M[X]* или – среднее значение признака (математическое ожидание);
● σ – среднее квадратическое отклонение.

Несколько фактов из истории. Закон нормального распределения впервые был выведен в 1783 г. Абрахамом де Муавром, английским математиком французского происхождения. Изначально наиболее активно нормальное распределение использовалось в астрономии, для описания распределения ошибок измерений вокруг «истинных» значений (точных координат) расположения того или иного небесного тела. Поэтому по сегодняшний день его часто называют кривой ошибок. В отношении к социальным, психологическим и антропометрическим данным закон нормального распределения был впервые применен Адольфом Кетле, бельгийским статистиком, математиком, астрономом, метеорологом, социологом, в 1830 году.

По мнению А. Кетле[15], социальные и моральные данные стремятся расположиться на нормальной кривой, которая является проявлением всеобщего закона природы. Изображение этой кривой приведено на графике, представленном чуть ниже (см. *Рис. 3.1*).

Смысл графика заключается в том, что все характеристики живой и неживой материи в этом мире распределяются по этому закону. Если, например, глубоко задуматься о том, как все люди в мире распределены по росту, взглянув на график, сразу же можно это увидеть. Становится понятным, что в мире существует три-пять процентов самых высоких людей и три-пять процентов самых низкорослых, рост всех остальных людей равномерно распределяется между этими двумя противоположностями, формируя понятие среднего.

То же самое касается и всех остальных количественных и качественных характеристик «человеческой» совокупности. Например, способности человека, те или иные особенности его поведения, склонности, черты характера, «в идеале» должны быть подчинены закону нормального распределения. Причем считается, что этот закон начинает действовать в группе людей, численность которой составляет более двух человек. Однако чем больше людей, тем нагляднее закон себя проявляет.

Вполне правдоподобным представляется тот факт, например, что люди по уровню одаренности талантом распределены согласно нормальному закону: в

[15] **Адо́льф Кетле́** (*Ламбе́р Адо́льф Жак Кетеле́*; 22 февраля 1796, Гент–17 февраля 1874, Брюссель) − бельгийский математик, астроном, метеоролог, социолог. Один из родоначальников научной статистики.

мире существует три-пять процентов талантливых людей и три-пять процентов «бездарей», которые, так сказать, друг друга «уравновешивают», а все остальные люди располагаются где-то между ними. Однако, если речь идет непосредственно о социологических данных, то было бы ошибочным считать распределения, отличающиеся от нормального, необычными или «незаконными». Взять, к примеру, картину распределения населения Украины по уровню материального положения: было бы крайне оптимистичным полагать, что в нашем государстве всего лишь 3−5% тех, кто находится у черты бедности. По данным Института демографии и социальных исследований им. Птухи НАН Украины, после экономического кризиса 2008–2009 годов в Украине стало больше людей, ощущающих себя нищими. Согласно социологическим исследованиям, проведенным Центром Разумкова[16] в последние годы, уровень нищеты (то есть людей, которые едва сводят концы с концами и их дохода не хватает даже на питание) увеличился с 12,5% в начале 2008 года до 13,8% в октябре 2011 года. Количество бедных людей, то есть тех, доходов которых хватает исключительно на питание и вещи самой первой необходимости, соответственно также увеличилось с 34% до 41% . Это данные – основанные на самооценках граждан Украины. По объективным показателям дохода и имущественного признака мы имеем 10% самых бедных украинцев и 2–5% самых богатых. Причем разница в уровне их материальной обеспеченности составляет (на момент конца 2001 – начала 2012 гг.) от 40 до 60 раз (с учетом того, что «нормальный» разрыв между «бедняками» и «богачами» должен быть не более, чем в 4−6 раз) [17].

Распределение необработанных эмпирических данных, полученных в результате социологического исследования, может подчиняться любому случайному распределению. Тем не менее как статистическая модель вариаций нормальная кривая наиболее популярна, несмотря на то, что она не обладает универсальностью, приписанной ей А. Кетле.

Характеристики нормальной кривой. По определению, нормальная кривая ●*состоит из бесконечного числа точек,* ● *унимодальна,* ●*симметрична* и ●*не ограничена* в обоих направлениях. Следовательно, ●*мода, медиана и среднее арифметическое равны* по величине и делят распределение на две равные части. Графическое изображение этого распределения представляет собой ровную, колоколообразную кривую с характерным крылом, которое нигде не касается базовой линии. Начиная с максимума, кривая спадает все сильнее до точек перегиба, а затем становится более пологой, простираясь бесконечно в обоих направлениях. Точки перегиба находятся на расстоянии, равном одной сигме от максимума, делящего кривую на две симметричные части. Графически сигма равна линейному расстоянию вдоль базовой линии от центра до координаты, определяющей точки перегиба.

[16] Украинский негосударственный аналитический центр, который проводит исследования государственной политики в различных сферах деятельности (внутренняя и внешняя политика, экономическая политика, социальная политика, государственное управление и др.)

Каждому значению сигмы – расстоянию от среднего – строго соответствует определенный процент площади или частоты; следовательно, «сигма-точки» служат удобной мерой положения. Интервал переменной в одну сигму, отсчитываемый от среднего, соответствует 34%. Следующий интервал между средним и точкой, соответствующей двум сигмам (2σ), имеет относительную частоту не 68%, а лишь 48 из-за изменения наклона кривой при увеличении расстояния от среднего. Так как распределение симметрично, то приблизительно две трети случаев (68,26%) заключены внутри центрального интервала, который простирается от $-\sigma$ до $+\sigma$. 95% случаев попадают внутрь интервала размером в две сигмы по обе стороны от среднего и практически 100% (99,72%) попадают внутрь интервала размером в три сигмы по обе стороны от среднего.

Аналогично можно определить долю случаев между средним и любым другим значением интервала на основной линии, выраженном в единицах сигмы. Для этой цели составлены справочные таблицы. Наиболее известная из них – таблица нормальных площадей, заключенных под кривой между средним и выбранным значением переменной в единицах сигм – которая содержит долю полной площади. Такая таблица в сокращенном виде приводится в конце учебника (см. Приложение 2, *табл. Е*). Она описывает лишь одну половину распределения, так как последнее симметрично.

Изучение таблицы показывает, что приблизительно 0,4953 случаев расположены в пределах интервала, длиной $2{,}6\sigma$, отсчитываемого от среднего, то есть практически 99% лежат внутри интервала в $2{,}6\sigma$ по обе стороны от среднего. Следовательно, для всех практических целей в случае нормального распределения достаточно иметь интервал переменной длиной в 6σ.

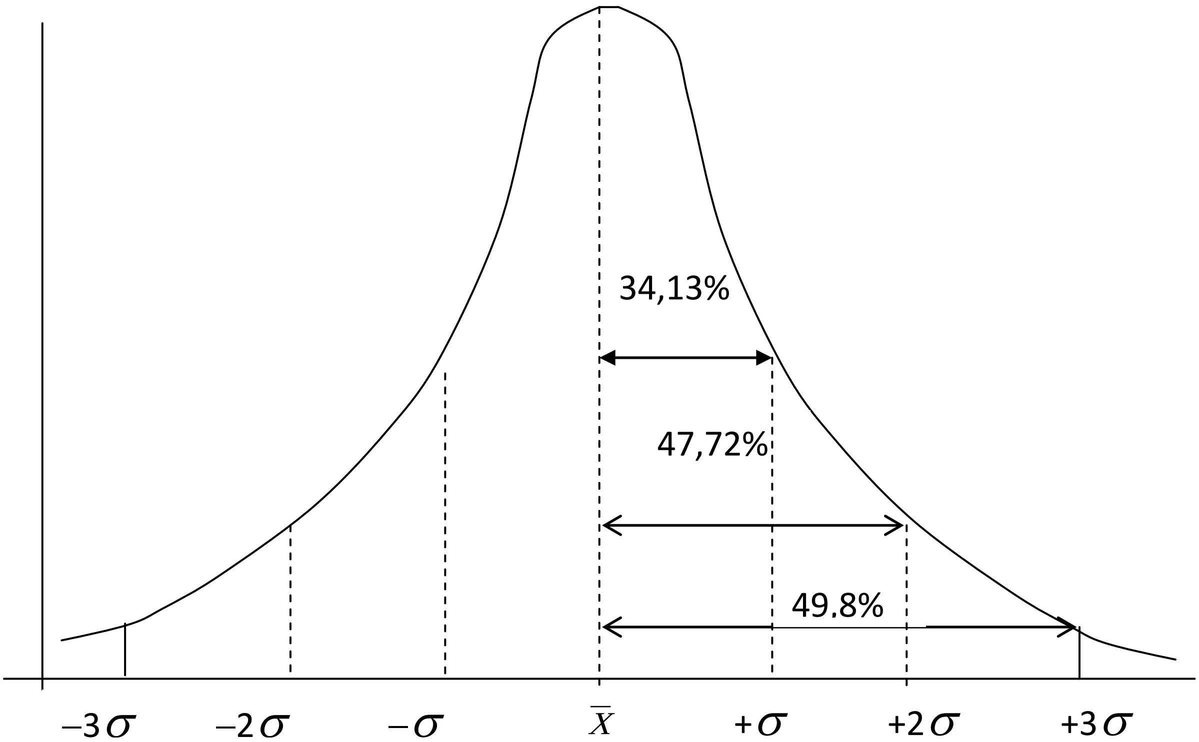


Рис. 3.1. Кривая нормального распределения

Другие законы «идеальных» распределений. К другим типам распределений, реже встречающимся и поэтому не рассматриваемым подробно в данном издании, можно отнести:

1. Закон распределения Пуассона. Определить его можно в случае, если значение среднего арифметического близко к значению среднего квадратического отклонения. По закону Пуассона развивается процесс спроса и предложения.
2. Биномиальный закон распределения, который применяется для альтернативных, качественных признаков.

Стандартное и нормированное отклонение. Выбор сигмы в качестве единицы измерения на базовой линии нормального распределения дает единицу, независимую не только от различных систем измерения, но также и от самих конкретных значений. Для нее безразлично, имеем ли дело с доходами в сотню или миллион долларов, с длительностью в десять секунд или десять лет, или с изменяющейся напряженностью установки.

Однако первичные данные всегда поступают к исследователю в необработанном виде, и поэтому нельзя, например, быстро сравнить оклады учителей и их стаж работы, даже если обе переменные имеют нормальные распределения. Решение этой проблемы лежит, конечно, в переводе первичных данных в сигма-единицы, то есть в сравнимые показатели.

С этой целью отклонения от соответствующего среднего представляются как частные от деления интервалов на сигму. Полученные показатели называются нормированными отклонениями. Их принято обозначать через Z.

$$Z = \frac{x_i - M[X]}{\sigma}.$$

где • x_i значение признака;

• $M[X]$ −среднее значение;

• σ − среднее квадратическое отклонение.

Нанеся абсолютные Z и σ шкалы на базовую линию, можно наглядно показать эквивалентность первичных данных нормированными отклонениями. Более того, с помощью наложения шкал можно обнаружить скрытое первоначально тождество между нормально распределенными переменными.

Для превращения любых последовательностей значения в стандартные отклонения, прежде всего, необходимо вычислить среднее и σ. Перевод в сигма-единицы позволяет установить относительное положение каждого события в совокупности путем использования таблицы нормальных площадей или частот. Из этой таблицы, например, определяется, что точка 1,5σ, отсчитанная от среднего, соответствует приблизительно 43% полной частоты и т. д. Необходимо подчеркнуть, что, хотя величина сигм может быть вычислена для любого распределения, вышеприведенный перевод в сигма-единицы справедлив лишь для нормального распределения.

3.2. Проверка нормальности распределения

Так уж устроены мы, люди, что нами «все познается в сравнении». Не зная, что такое «плохо», мы не в состоянии определить, что такое «хорошо». Наиболее часто в качестве эталона для сравнения мы выбираем некий идеал, как правило, не нами придуманный, существующий объективно. Для того чтобы осуществить анализ данных, полученных в ходе социологического исследования, для того чтобы эти данные превратить в достоверную информацию, социолог также должен прибегнуть к процедуре сравнения. За эталон сравнения в данном случае принимается «идеальное» нормальное распределение. Самая важная социологическая информация получается именно в результате подобного сравнения, на основе определения того, насколько сильно и по каким параметрам эмпирическое распределение значений того или иного признака отличается от этого идеала.

В арсенале математической статистики существует множество методов сравнения эмпирических распределений с «идеальной» теоретической моделью нормального распределения. Вообще, принадлежность наблюдаемых данных нормальному закону является необходимой предпосылкой для корректного применения большинства классических методов математической статистики, используемых в процедуре измерения. Поэтому проверка на отклонение от нормального закона является частой процедурой и при анализе социологических данных [14].

Начинается эта процедура с выдвижения статистической гипотезы, продолжается выбором критерия проверки этой гипотезы и завершается вычислением коэффициента, который свидетельствует о наличии либо отсутствии, силе и/или направлении связи.

Статистический вывод. Проверка гипотез. Так как во многих практических задачах точный закон распределения неизвестен, он представляет собой гипотезу, которая требует статистической проверки. Статистические гипотезы могут иметь различные формулировки, но по своей сути они являются гипотезами о распределении той или иной случайной величины. На практике часто приходится делать некоторые выводы по имеющемуся у нас небольшому объему данных (выборки) о свойствах всей генеральной совокупности. Эти выводы осуществляются с помощью определенных статистик и поэтому называются статистическими. Теория статистического вывода занимает центральное место в статистике. Основным способом, с помощью которого делаются статистические выводы, является проверка гипотез [3, с. 31−33; 5; 9].

Существует два вида гипотез: 1) научные и 2) статистические. *Научная гипотеза* – это предполагаемое решение некоторой проблемы. Она обычно формулируется в виде теоремы. ***Статистическая гипотеза*** – некоторое утверждение относительно неизвестного параметра или какой-либо характеристики. Например, среднее значение генеральной совокупности равно

125 ($M[X]$=125) или коэффициент корреляции равен 0 (r =0).

Для проверки статистических гипотез используются статистические критерии, которые представляют собой некоторое правило, по которому мы делаем вывод о правильности или неправильности рассматриваемой статистической гипотезы [3, с. 31-32 ; 10]. Подробнее о статистических критериях будет сказано в подразделах **3.3** и **3.4**.

При использовании этих критериев возможны четыре исхода проверки гипотез:

1. Гипотеза H_0 верна и принимается, согласно критерию.
2. Гипотеза H_0 неверна и отвергается, согласно критерию.
3. Гипотеза H_0 верна, но отвергается, согласно критерию (*ошибка первого рода*).
4. Гипотеза H_0 не верна, но принимается, согласно критерию (*ошибка второго рода*).

Как видим, при проверке статистических гипотез всегда существует вероятность допустить ошибку, приняв неверную или опровергнув верную гипотезу. При этом *разграничивают ошибки первого или второго рода*.

Ошибкой первого рода называется ошибка, состоящая в опровержении верной гипотезы, когда делается неверное заключение о том, что гипотеза не соответствует истине, когда в действительности такое соответствие есть. При этом, согласно критерию, отвергается гипотеза H^0, которая на самом деле верна. Вероятность такой ошибки не выше определенного уровня значимости. ***Уровнем значимости α*** называется вероятность совершения ошибки первого рода. Значение уровня значимости обычно задается близким к нулю (например, 0,05; 0,01; 0,02 и т. д.), потому что, чем меньше это значение, тем меньше вероятность совершения ошибки первого рода, состоящей в опровержении верной гипотезы H^0 .

Ошибкой второго рода называется ошибка, состоящая в принятии ложной гипотезы, когда делается неверное заключение, о том, что гипотеза H_0 соответствует истине, когда в действительности такого соответствия нет. Вероятность совершения ошибки второго рода, т. е. принятия ложной гипотезы, обозначается как ***β***.

При проверке нулевой гипотезы возможно возникновение следующих ситуаций:

H_0	*Верная*	*Ложная*
Отклоняется	Ошибка первого рода	Решение верное
Не отклоняется	Решение верное	Ошибка второго рода

По сути, проверка статистической гипотезы означает проверку согласования исходных выборочных данных с выдвинутой основной гипотезой. ***Общая схема проверки статистической гипотезы*** состоит из пяти этапов:

I – *выдвигаем две статистические гипотезы*:
1) основную − нулевую (H_0),
2) конкурирующую − альтернативную (H_1).

Например, H_0 среднее значение ГС = 125.

H_1 среднее значение ГС ≠ 125.

Математически это можно записать:

H_0: $\boldsymbol{M[X] = 125}$.

H_1: $\boldsymbol{M[X] \neq 125}$ ($\boldsymbol{M[X]}$ < 125; $\boldsymbol{M[X]}$ > 125).

II – *задаемся уровнем значимости*. Статистический вывод никогда не может быть сделан со стопроцентной уверенностью. Всегда допускается риск принятия неправильного решения. При проверке статистических гипотез мерой такого риска и выступает уровень значимости, который обычно обозначается. Фактически уровень значимости представляет собой долю и процент ошибок, которые мы можем себе позволить при статистических выводах. Чаще всего используют следующие три значения уровня значимости: α = 0,1 или 10%; α = 0,05 или 5%; α = 0,01 или 1%. Наиболее популярным из них является α = 0,05 или 5% (допускается 5% ошибка на 100% выборки).

III – по исходным данным, то есть по выборке, *вычисляем наблюдаемое значение статистического критерия*. В общем случае обозначим его как $K_{набл}$.

IV – по специальным статистическим таблицам (которые в данном издании приводятся в приложении) *находим критическое значение критерия* ($K_{кр}$), в зависимости от заданного уровня значимости (**α**) и числа степеней свободы (***df***).

V – *путем сравнения* найденных наблюдаемых значений критерия ($K_{набл}$) с теоретическим (табличным) критическим значением ($K_{кр}$) делаем вывод о правильности/неправильности, принятии/отвержении гипотезы.

● Если найденное для выборки $K_{набл}$ меньше табличного $K_{кр}$, то гипотеза H_0 принимается на заданном уровне значимости α. В этом случае наблюдаемое различие двух сравниваемых нами совокупностей объясняется только случайностью выборки. Хотя принятие гипотезы H_0 совсем не означает, что параметры этих совокупностей полностью одинаковы. Считается, что имеющийся в распоряжении статистический материал просто не дает оснований для отклонения данной гипотезы.

● Если же найденное для выборки $K_{набл}$ больше табличного $K_{кр}$, то гипотеза H_0 отклоняется в пользу гипотезы H_1 при заданном уровне значимости α. В этом случае наблюдаемое отличие сравниваемых совокупностей нельзя объяснить только лишь случайностью, скорее всего, имеет место наличие одного или нескольких весомых факторов, влияющих на распределение.

В определенном смысле задача проверки статистических гипотез логически предшествует задаче оценивания. Например, при статистическом исследовании имеющейся разницы между средними двух нормальных совокупностей исследователю приходится установить, является ли данная разница статистически значимой, не является ли она результатом выборочной случайной флуктуации, то есть проверить гипотезу о равенстве средних. Если эта гипотеза не подтвердится, исследователю, возможно, придется перейти к следующему этапу − оцениванию величины разницы между средними.

Необходимые для проверки различных гипотез вычисления могут быть сделаны с помощью специально разработанных компьютерных статистических программ. Однако важно понимать, что истинным критерием проверки статистических гипотез все-таки остается опыт самого исследователя. По этому поводу согласимся с учеными, утверждающими, что статистические критерии значимости – лишь формально точный инструмент. Чем больше социолог знает об исследуемом им явлении, тем точнее будет сформулирована гипотеза и выводы, сделанные на основе критериев значимости [11; 12; 15].

3.3. Общие представления о критериях «нормальности» распределения

Как уже говорилось, тестирование данных на нормальность является необходимым для их полноценного, полного анализа, так как большое количество статистических методов исходит именно из предположения нормальности распределения изучаемых данных. Существует группа статистических критериев, предназначенных для проверки нормальности распределения.

Статистический критерий *(К)*− специально подобранная случайная величина, точное либо приближенное распределение которой известно и обычно сведено в таблицы.

Множество всех возможных значений выбранного статистического критерия делится на два непересекающихся подмножества. Первое подмножество включает в себя те значения критерия, при которых основная гипотеза отвергается, а второе подмножество – те значения критерия, при которых основная гипотеза принимается.

Критической областью называется множество возможных значений статистического критерия, при которых основная гипотеза отвергается. Областью принятия гипотезы или областью допустимых значений называется множество возможных значений статистического критерия, при которых основная гипотеза принимается. Если наблюдаемое значение статистического критерия, рассчитанное по данным выборочной совокупности, принадлежит критической области, то основная гипотеза отвергается. Если наблюдаемое значение статистического критерия принадлежит области принятия гипотезы, то основная гипотеза принимается.

Критическими точками или квантилями ($K_{кр}$) называются точки, разграничивающие критическую область и область принятия гипотезы. Критические области могут быть как односторонними (право- или левосторонними), так и двусторонними.

Правосторонняя критическая область характеризуется неравенством вида:

$$K_{набл} > K_{кр},$$

где •$K_{набл}$ – наблюдаемое значение статистического критерия, вычисленное по данным выборки; •$K_{кр}$ – *положительное* критическое (теоретическое) значение статистического критерия, определяемое по таблице распределения данного критерия. Следовательно, для определения правосторонней критической области необходимо рассчитать положительное значение статистического критерия $K_{кр}$.

Левосторонняя критическая область характеризуется неравенством вида:

$$K_{набл} < K_{кр} \ ,$$

где •$K_{набл}$ – эмпирическое (наблюдаемое) значение статистического критерия, вычисленное по данным выборки; •$K_{кр}$ – отрицательное критическое (теоретическое) значение статистического критерия, определяемое по таблице распределения данного критерия. Следовательно, для определения правосторонней критической области необходимо рассчитать отрицательное значение статистического критерия $K_{кр}$.

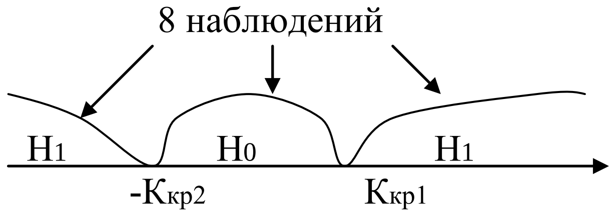
Двусторонняя критическая область характеризуется двумя неравенствами вида:

$$K_{набл} > K_{кр1} \quad \text{и} \quad K_{набл} < K_{кр2} \ , \textit{при этом} \quad K_{кр1} > \ K_{кр2} \ ,$$

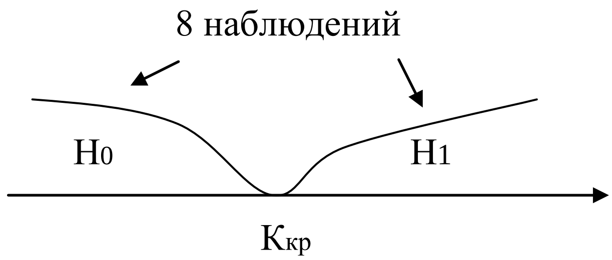
где •$K_{набл}$ – наблюдаемое значение статистического критерия, вычисленное по данным выборки; •$K_{кр1}$ – положительное значение статистического критерия, определяемое по таблице распределения данного критерия; •$K_{кр2}$ – отрицательное значение статистического критерия, определяемое по таблице распределения данного критерия.

Проиллюстрируем наглядно наиболее часто встречавшиеся примеры:

А) ***Для двусторонней критической области***



Б) ***Для односторонней критической области***



С учетом того, что в случае принятия нулевой гипотезы дисперсии сравниваемых распределений будут равны, то есть H_o: $D\ (X) = D\ (Y)$, выбор критической области осуществляется исходя из вида конкурирующей гипотезы *H1*. При этом:

1) *правосторонняя критическая область* выбирается в том случае, если H1: $D(X) > D(Y)$;

2) *левосторонняя критическая область* выбирается в том случае, если H1: $D(X) < D(Y)$;

3) *двусторонняя критическая область* выбирается в том случае, если H1: $D(X) \neq D(Y)$.

Основной вопрос, возникающий на основе сказанного, заключается в определении того, ***какой из критериев, односторонний или двусторонний, выбрать в том или ином случае***. Ответ на этот вопрос следует искать за пределами математико-статистического формализма, он в полной мере зависит от целей и задач социологического исследования. Если логикой исследования допускается, что различия сравниваемых параметров могут принимать как положительные, так и отрицательные значения, то следует применять двусторонний критерий. Если же есть дополнительная информация (основанная, например, на опыте предыдущих исследований) о том, что один параметр может принимать значения либо больше, либо меньше значений другого параметра, то используется односторонний критерий. Причем, если имеются основания, достаточные для применения одностороннего критерия, то следует именно его (а не двусторонний) и предпочесть, так как односторонний критерий полнее использует информацию об изучаемом явлении либо процессе и поэтому чаще дает правильные и более точные результаты [12].

Мощностью статистического критерия называется вероятность попадания данного критерия в критическую область, при условии, что справедлива конкурирующая гипотеза *H1*, то есть выражение $\boldsymbol{1-\beta}$ является мощностью критерия. Ясно, что мощность может принимать любые значения от 0 до 1. Чем ближе мощность к единице, тем эффективнее критерий.

Критическую область следует строить так, чтобы мощность критерия была максимальной. Повторимся, что для выполнения этого требования наиболее часто задается уровень значимости $\alpha = 0{,}05$ или 5% (т.е. допускается 5% ошибки на 100% выборки). Но если выводы, которые предстоит сделать по результатам проверки гипотез, связаны с большой ответственностью, то следует выбирать $\alpha = 0{,}01$, или $\alpha = 0{,}001$. Это обеспечивает минимальную ошибку второго рода, состоящую в том, что будет принята неправильная гипотеза. Однако вполне резонным является вопрос о том, ***как трактовать результаты исследования, связанные с установлением того или иного уровня значимости***. Часто поступают следующим образом: уровень значимости не устанавливается точно, а по экспериментальным данным вычисляется вероятность (*P*) того, что критерий $K_{набл}$ выйдет за пределы $K_{кр}$. Окончательные результаты обычно приводят в следующем виде: 1) если вычисленное значение критерия не превосходит критического (табличного) значения на уровне $\alpha = 0{,}05$, то различие считается статистически незначимым;

2) если вычисленное по выборке значение превышает критические значения $\alpha = 0{,}05$, $\alpha = 0{,}01$и $\alpha = 0{,}001$, то результаты записываются следующим образом: соответственно, $p<0{,}05$; $p<0{,}01$; $p<0{,}001$. Это значит, что наблюдаемые различия исследуемых совокупностей статистически значимы на уровнях значимости 0,05 или 0,01, или 0,001.

Критерии значимости подразделяются на три типа:

1. ***Параметрические критерии*** значимости, то есть те, которые служат для проверки гипотез о параметрах распределения генеральной совокупности (чаще всего – нормального распределения). *Параметрические критерии включают в формулу расчета параметры распределения, то есть средние и дисперсии.*

2. ***Непараметрические критерии*** значимости, то есть те, которые для проверки гипотез не используют предположений о параметрах распределения (например, в случаях, когда эти параметры неизвестны). *Непараметрические критерии не включают в формулу расчета параметров распределения и основаны на оперировании частотами или рангами.*

3. ***Критерии согласия*** – как отдельная, особая группа, которую составляют критерии, служащие для поверки гипотез о согласованности генеральной совокупности, из которой получена выборка, с ранее принятой теоретической моделью (чаще всего – моделью нормального распределения). Именно поэтому критерии согласия часто ассоциируют с так называемыми критериями «нормальности».

Вообще же, критерии нормальности являются лишь частным случаем критериев согласия. Однако именно к этим критериям наиболее часто прибегают социологи при осуществлении анализа эмпирических данных.

В данном издании обратимся к подробному рассмотрению некоторых из этих критериев, представляющих особую важность для исследователя-социолога. Речь идет, прежде всего, о таких критериях, как: ●*критерий асимметрии* (параметрический и/или согласия), ● *критерий Пирсона Хи-квадрат* (непараметрический и/или согласия), ● *критерий Стьюдента* (параметрический), ●*критерий Фишера* (параметрический). О них и пойдет речь в следующем, заключительном подразделе этого раздела.

3.4. Критерии «нормальности»: принципы нахождения и расчетные формулы

Предварительная проверка нормальности распределения: критерий асимметрии (*As*). Все рассматриваемые нами в данном издании критерии значимости обеспечивают наивысшую степень достоверности статистических выводов только в тех случаях, если выборки получены из нормального распределения генеральной совокупности. При отклонении от нормального распределения эта степень существенно снижается. Следовательно, чтобы уверенно применять эти критерии, необходимо изначально осуществить проверку предположения о нормальном распределении генеральной совокупности. Для этого, в общем-то, и существуют критерии согласия. Но, применение большинства этих критериев – довольно трудоемкая вычислительная работа, занимающая немало времени. Поэтому целесообразным представляется осуществление предварительной проверки на нормальность с использованием более простых методов. Они, конечно, обладают сравнительно меньшей мощностью, помогают установить только довольно большие расхождения с нормальным распределением, но и этого может быть достаточно, чтобы исследователь мог избавиться от необходимости осуществлять сложные вычислительные процедуры, сэкономив тем самым и драгоценное время, и энергию.

Такую упрощенную предварительную проверку можно осуществить с помощью критерия асимметрии. При этом H_0: *случайная величина имеет распределение, отличное от нормального*. Если распределение нормально, то его коэффициент асимметрии (*As*) будет равен нулю. Так как нулевые значения коэффициента могут иметь место и для распределений, отличных от нормального, то этот критерий следует воспринимать как критерий установления отклонения от нормальности распределения, но не установления нормальности.

Различают несколько вариантов эмпирических распределений при сравнении их с теоретическими, а именно: симметрические и скошенные, примеры их полигонов были рассмотрены в параграфе **2.6** (см. *Рис. 2.10,* с. 80). *Симметрическое распределение* – это распределение, где частоты равномерно удалены от среднего арифметического (*M[X])*. Однако при анализе реально существующих социальных процессов и явлений чаще встречаются *скошенные (асимметрические) распределения*. Поэтому при характеристике эмпирических рядов распределений рассматривают ●*левостороннюю,* ●*правостороннюю,* ●*умеренно скошенную и* ●*крайне скошенную асимметрии*.

При симметрическом распределении, имеет место уменьшение частот от максимального, модального значения (Mo) одинаково в обе стороны. *Умеренно-асимметрическое* распределение характеризуется тем, что большая часть частот ответов респондентов приходится на одну сторону от максимального значения. Именно такие распределения чаще всего встречаются в социологических исследованиях. В некотором роде аномальными видами распределений являются:

– *Т-образное или крайне скошенное* – это распределение, в котором наибольшая численность групп находится на одном конце амплитуды колебаний;

– *V-образное распределение* – это распределение, при котором наибольшая численность групп находится на концах амплитуды колебаний, а его минимум приходится на центральный интервал. При появлении такого распределения чаще всего его разбивают на две части и рассматривают два независимых процесса. Разбиение производится таким образом, чтобы полученные распределения приблизительно соответствовали требованию $\bar{x} = Mo = Me$.

Для того чтобы определить величину скошенности или асимметрии, вычисляют коэффициент *As* по следующей формуле:

$$As = \frac{M[X] - \mathrm{M}o}{\sigma},$$

где • $M[X]$ – среднее арифметическое значение признака;

• Mo – модальное значение;

• σ - среднее квадратическое отклонение.

Свойства асимметрии:

● коэффициент As всегда изменяется только в пределах от –3 до +3. Чем ближе As к граничным значениям (–3, +3), тем больше скошенность;

● чем больше разница между средним арифметическим и модой (медианой), тем больше асимметрия ряда;

● если значение As положительно, то говорят, что распределение вправо скошено, если отрицательно, то – влево скошено, если $As = 0$ – то асимметрии нет, то есть распределение симметрично и $\bar{x} = Mo = Me$;

● если $A_s > 0$, то кривая распределения имеет длинный правый «хвост», то есть налицо правосторонняя асимметрия. При этом выполняется соотношение $Mo < Me < M\,[X]$;

● если $A_s < 0$, то асимметрия левосторонняя, кривая распределения имеет длинный левый «хвост». При этом $M\,[X] > Me > Mo$;

● на практике асимметрия считается значительной, если коэффициент асимметрии превышает по модулю 0,25.

Критерий Пирсона *Хи*-квадрат (χ^2). Наиболее часто употребляемым критерием оценки того, в соответствии с каким законом распределены значения признака, является критерий *Хи*-квадрат (далее по тексту – χ^2). Популярность χ^2 обусловлена главным образом тем, что применение его не требует предварительного знания закона распределения изучаемого признака. Кроме того, признак может принимать как непрерывные, так и дискретные значения, причем измеренные хотя бы на номинальном уровне.

Критерий χ^2 – статистический критерий для проверки гипотезы H_0: *наблюдаемая случайная величина подчиняется некому теоретическому закону (нормального распределения).* Следовательно, подтверждение данной гипотезы непременно требует сравнения эмпирических частот (то есть тех, которые были получены в результате социологического исследования) и теоретических частот (то есть тех, которые могли бы быть получены, если бы данное распределение было «идеально» нормальным).

По своей сути, критерий χ^2 отвечает на вопрос о том, с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух и более эмпирических распределениях.

Следовательно, критерий χ^2 Пирсона, используется:

1) для сопоставления *эмпирического* распределения признака с *теоретическим* – равномерным, нормальным или каким-то иным;

2) для сопоставления *двух, трех или более эмпирических* распределений одного и того же признака.

Принцип вычисления основывается на процедуре выравнивания вариационного ряда по кривой нормального распределения, то есть предполагает обязательное нахождение теоретических частот.

Итак, критерий χ^2 для проверки H_0 имеет вид:

$$\chi^2_{набл} = \sum_{(i)}\sum_{(j)} \frac{(n_{ij} - \tilde{n}_{ij})^2}{\tilde{n}_{ij}},$$

где ● nij – наблюдаемая (эмпирическая) частота, число объектов в ячейке на пересечении i-й строки таблицы сопряженности признаков и j-го столбца, так называемая фактическая клеточная частота;

● ñij − ожидаемая по H_0 (теоретическая) частота в этой же ячейке;

● $i = 1, 2, ..., r$ (то есть r − число строк таблицы сопряженности признаков);

● $j = 1, 2, ..., c$ (то есть c − число столбцов таблицы сопряженности признаков).

Как видим из формулы, первое, что нам необходимо сделать, – это найти ожидаемые (теоретические) частоты.

Ожидаемые (теоретические) частоты (ñij) вычисляют по формуле:

$$\tilde{n}_{ij} = \frac{r_i c_j}{N},$$

где ● r_i – маргинал i-й строки;

● c_j – маргинал j-го столбца;

● N – объем выборочной совокупности.

Дальнейшая последовательность действий довольно проста, мы проиллюстрировали ее на Примере 3.4.1, представленном чуть ниже.

Заметим, что для таблицы 2×2 число степеней свободы всегда равно ***1***.

При расчете критерия *Хи*-квадрат Пирсона должно соблюдаться следующее условие: достаточно большим должно быть число наблюдений ($n \geq 30$). Если полученное значение $\chi^2_{набл}$ меньше критических значений, представленных в таблице, то делается вывод о том, что эти расхождения между эмпирическими и теоретическими частотами распределения могут быть случайными, формулируется предположение о близости эмпирического распределения к нормальному и принимается гипотеза H_0. Фактически же это означает, что искомая связь между признаками если и есть, то, вероятно, она слишком слаба.

Итак, рассмотрим реальный пример применения критерия χ^2 Пирсона.

Пример 3.4.1. Предположим, мы провели социологическое исследование для известного «Модного дома». В данном случае изучение взаимосвязи между возрастом респондентов и их отношением к «классической» моде могло бы быть очень информативным и полезным для заказчика исследования. Таблица сопряженности признаков «Возраст» и «Отношение к классической моде» с абсолютными частотами двумерного распределения представлена ниже (см. *Табл. 3.1*).

Таблица 3.1

Таблица сопряженности по признакам «Возраст» и «Отношение к классической моде» (абсолютные частоты)

Возраст / ***Отношение***	*Молодежь*	*Люди среднего возраста*	*Пожилые люди*	**ИТОГО** *(маргинальные частоты по строкам –* r_i *для числа строк – r)*
Очень нравится классический стиль одежды	26	13	5	44
Скорее нравится, чем нет	20	11	8	39
Скорее не нравится, чем нравится	9	10	20	39
Совершенно не нравится «классика» в одежде	7	10	15	32
ИТОГО *(маргинальные частоты по столбцам –* c_j*; для числа столбцов – c)*	62	44	48	n=154

Попытаемся подтвердить/опровергнуть наличие/отсутствие взаимосвязи между признаками при помощи критерия χ^2 Пирсона. Если полученное наблюдаемое значение критерия $\chi^2_{набл}$ будет меньше критического значения $\chi^2_{кр}$, по заданному уровню значимости (α), в соответствии с числом степеней свободы (***df*** ***)***, то гипотеза *H₀* принимается и делается вывод о вероятном отсутствии связи между признаками. По сути это значило бы, что и молодые, и пожилые, и люди среднего возраста (с большой вероятностью) сделают одинаковые выборы, определяя свое отношение к классической моде. А если отличий нет, значит, признак возраста никак не влияет на это отношение. И наоборот, если $\chi^2_{набл}$ будет больше $\chi^2_{кр}$, гипотеза *H*$_0$ отвергается в пользу альтернативной *H*$_1$ и делается вывод о вероятном наличии связи.

Итак, найдем $\chi^2_{набл}$ и процедуру поиска опишем при помощи расчетной таблицы, представленной ниже.

Таблица 3.2

Расчетная таблица нахождения критерия $\chi^2_{набл}$ для двумерного распределения по признакам «Возраст» и «Отношение к моде»

№ п/п ячейки	n_{ij} *Набл.,эмп. частота*	$\tilde{n}_{ij}$ *Ожид. частота*	$(n_{ij}-\tilde{n}_{ij})$	$(n_{ij}-\tilde{n}_{ij})^2$	$\frac{(n_{ij}-\tilde{n}_{ij})^2}{\tilde{n}_{ij}}$
1	26	17,7	8,3	68,89	3,89
2	13	12,6	0,4	0,16	0,01
3	5	13,7	-8,6	73,96	5,41
4	20	15,7	4,3	18,49	1,18
5	11	11,1	-0,1	0,01	0,00
6	8	12,2	-4,2	17,64	1,45
7	9	15,7	-6,7	44,89	2,86
8	10	11,1	-1,1	1,21	0,11
9	20	12,2	7,8	60,84	4,99
10	7	12,9	-5,9	34,81	2,7
11	10	9,1	0,9	0,81	0,09
12	15	10	5	25	2,5
					$\sum = \mathbf{25,18}\ (\chi^2_{набл})$

Описание вычислительных действий:

1) выписываем слева направо наблюдаемые частоты;

2) для каждой из них по соответствующей формуле вычисляем ожидаемые (теоретические) частоты, то есть те, которые могли бы стоять в ячейках таблицы в «идеальном» случае независимости признаков:

«1»: $\tilde{n}_{11} = \frac{44\times62}{154} = 17{,}7;$ «4»: $\tilde{n}_{21} = \frac{39\times62}{154} = 15{,}7;$

«2»: $\tilde{n}_{12} = \frac{44\times44}{154} = 12{,}6;$ «5»: $\tilde{n}_{22} = \frac{39\times44}{154} = 11{,}1;$

«3»: $\tilde{n}_{13} = \frac{44\times48}{154} = 13{,}7;$ «6»: $\tilde{n}_{23} = \frac{39\times48}{154} = 12{,}2;$

«7»: $\tilde{n}_{31} = \frac{39 \times 62}{154} = 15{,}7;$ «10»: $\tilde{n}_{41} = \frac{32 \times 62}{154} = 12{,}9;$

«8»: $\tilde{n}_{32} = \frac{39 \times 44}{154} = 11{,}1;$ «11»: $\tilde{n}_{42} = \frac{32 \times 44}{154} = 9{,}1;$

«9»: $\tilde{n}_{33} = \frac{39 \times 48}{154} = 12{,}2;$ «12»: $\tilde{n}_{43} = \frac{32 \times 48}{154} = 10.$

Представленные выше результаты вычислений занесены в третий столбец расчетной таблицы;

3) далее находим разность между наблюдаемыми и ожидаемыми частотами;

4) полученные разности возводим в квадрат;

5) делим квадрат разностей на теоретические частоты;

6) все результаты деления складываем – эта сумма и является искомым наблюдаемым значением критерия χ^2 Пирсона.

В итоге мы получили $\chi^2_{набл}$ = **25,18.** Однако само по себе это число нам ни о чем сказать не может. Полученное значение критерия мы должны сравнить с критическим значением $\chi^2_{кр}$ по таблице (см. Приложение 2, *табл. Б*). В таблице критическое значение χ^2 находится по заданному уровню значимости α (0,05; 0,01 или 0,001) и числу степеней свободы *df*, которое находится по формуле:

df = (r - 1) (c -1),

где ● r – число строк таблицы сопряженности,

● *с* – число столбцов таблицы сопряженности.

Число степеней свободы для нашего случая будет следующим:

df = (4 - 1)(3 - 1) = 6.

По таблице критических значений χ^2 находим, что наблюдаемое значение критерия, соответствующее df = 6**,** превышает критические значения по всем уровням значимости:

$$\chi^2_{кр} = \begin{cases} 12{,}59 \text{ для } p \le 0{,}05 \\ 15{,}03 \text{ для } p \le 0{,}02 \\ 16{,}81 \text{ для } p \le 0{,}01 \end{cases}$$

$\chi^2_{набл}$ =25,18 > $\chi^2_{кр}$ =16,81.

Вывод: Гипотеза *H₀* отвергается, распределения по разным возрастным группам различаются в своем отношении к классической моде.

Т-критерий Стьюдента (критерий, основанный на нормальном распределении). *Т*-критерий Стьюдента − общее название для статистических тестов, в которых статистика критерия имеет распределение Стьюдента. Распределение Стьюдента, по сути, представляет собой сумму нескольких нормально распределенных случайных величин. Чем больше величин, тем больше вероятность, что их сумма будет иметь нормальное распределение. Таким образом, количество суммируемых величин определяет важнейший параметр формы данного распределения − число степеней свободы (***df***).

Наиболее часто t-критерий Стьютдента используется для:

1) установления сходства-различия средних арифметических значений в двух независимых выборках или для установления сходства-различия двух эмпирических распределений;

2) установления сходства-различия двух дисперсий в двух зависимых выборках.

Ограничения в применении t-критерия Стьюдента [10; 16]:

1) это параметрический критерий, поэтому необходимо, чтобы распределение признака, по крайней мере, не отличалось от нормального распределения;

2) ***разные формулы расчета***

1 *для независимых выборок:*

$$t_{\text{набл}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{(n_1-1)\delta_1^2 + (n_2-1)\delta_2^2}} \times \sqrt{\frac{n_1\, n_{2(n_1+n_2-2)}}{(n_1+n_2)}},$$

где ● $\bar{x}_1$ и $\bar{x}_2$ – средние арифметические, соответственно, в 1-й и 2-й сравниваемых выборках;

● n_1, n_2 – количество респондентов в 1-й и 2-й выборках;

● σ1, σ2 – средние квадратические отклонения в 1-й и 2-й выборках.

В данном случае *количество степеней свободы для нахождения критического значения критерия* (*$t_{\text{кр}}$*): ***df*** = n1+ n2 − 2.

2 *для зависимых выборок* (с учетом корреляция результатов, поскольку измерения проводятся на одних и тех же респондентах в различных условиях (x и y)':

$$t_{\text{набл}} = \frac{(\sum d_i)}{\sqrt{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}} \times \sqrt{\frac{n-1}{n}},$$

где ● $d_i = x_i - y_i$, то есть разность значений признака в разных условиях (x и y)' для каждого респондента;

● n – количество респондентов в выборке (объем выборки).

В данном случае *количество степеней свободы для нахождения критического значения критерия* (*$t_{кр}$*): ***df*** = n – 1. Проверяется статистическая гипотеза о соответствии распределения разностей t-распределению Стьюдента с нулевым средним значением.

Гипотезы:

1 *для независимых выборок*:

H_0: средние значения признака в обеих выборках не различаются, линейная связь между выборками отсутствует;

H_1: средние значения признака в обеих выборках статистически значимо различаются (гипотеза сдвига).

2 *для зависимых выборок (т.е. для выборок, состоящих из одних и тех же респондентов, опрошенных в разное время)*:

H_0: разности оценок респондентов в двух состояниях не отличаются от нуля;

H_1: разности оценок испытуемых в двух состояниях статистически значимо отличаются от нуля.

И все-таки наиболее часто t-критерии применяются для проверки равенства средних значений в двух выборках. Такое сравнение может быть интересным и даже необходимым, если исследователь хочет сравнить результаты нескольких аналогичных исследований, проведенных, например, в разные годы или в разных странах. Рассмотрим на примере [9, с. 64].

Пример 3.4.2 (*независимые выборки*). Предположим, имеется две независимые выборки школьников, интеллект которых развивали в течение некоторого времени по двум различным методикам. Требуется установить, какая из методик лучше (см. *Табл. 3.3*). Предварительно было выяснено, что начальный уровень интеллекта был одинаковым в обеих выборках. Однако после применения методик средний показатель интеллекта респондентов оказался выше. Задача состоит в том, чтобы разобраться, случайны ли эти различия или закономерны, то есть зависят ли они от самой методики? Такая задача сравнения двух методик может быть переформулирована на язык статистики как задача сравнения средних арифметических значений интеллекта в обеих выборках.

Таблица 3.3

Экспериментальные данные по двум выборкам

Числовые характеристики выборки	*1-я выборка*	*2-ая выборка*
n	30	32

$\bar{x}$	103	110
δ	10	12

Гипотезы:

H_0: средние значения уровня интеллекта в обеих выборках не различаются (а следовательно, эффективность методик одинаковая);

H_1: средние значения уровня интеллекта в обеих выборках статистически значимо различаются (а следовательно, эффективность методик разная).

Понятно, что в данном случае для получения эмпирического значения *t*-критерия используется формула для сравнения средних показателей двух независимых выборок, то есть формула «1».

Подставив числа, зафиксированные в таблице 3.3, в формулу $t_{набл}$ для независимых выборок, получим следующее числовое выражение:

$$t_{набл} = \frac{|103-110|}{\sqrt{(30-1)\times 10^2+(32-1)\times 12^2}} \times \sqrt{\frac{30\times 32(30+32-2)}{(30+32)}} \approx 2{,}486.$$

Для того чтобы сделать вывод о принятии/отвержении гипотезы H_0, необходимо сравнить наблюдаемое (эмпирическое) значение критерия – $t_{набл}$ с критическими табличными значениями $t_{кр}$, с помощью специальной таблицы критических значений t (см. Приложение 2, табл. *B*). По аналогии с критерием χ^2, для того, чтобы осуществить это сравнение, нужно найти число степеней свободы (*df*) . В нашем примере ***df*** = 30 + 32 – 2 = 60.

$$t_{кр} = \begin{cases} 2{,}0 \text{ для } p \le 0{,}05 \\ 2{,}66 \text{ для } p \le 0{,}01 \end{cases}$$

Полученное наблюдаемое значение t-критерия превышает критическое для α=*0,05*, но оказывается меньше критического для α = 0,01, то есть:
2,0 < 2,486 < 2,66.

Вывод: H_0 гипотеза отклоняется. А следовательно, можно сделать вывод о статистически значимом различии средних арифметических значений в двух выборках для $\rho \le 0{,}05$ и о преимуществах второй методики по сравнению с первой (т.к. средний показатель интеллекта по второй выборке выше).

Пример 3.4.3 *(для зависимых выборок)*. Допустим, проводится измерение уровня ситуативной напряженности в рабочем коллективе, состоящем из восьми человек до и после психологического тренинга, проведенного специалистом. Измерение осуществляется посредством метода анкетирования, с использованием специально разработанного, специфического инструментария. Основной вопрос, который интересует исследователя: приводит ли тренинг к изменению уровня напряженности?

Гипотезы:

H_0: разности оценок ситуативной тревожности респондентов до и после тренинга не отличаются от нуля. Следовательно, тренинг не дает результатов;

H_1: разности оценок ситуативной тревожности респондентов до и после тренинга статистически значимо отличаются от нуля, а следовательно, посредством данного тренинга возможно влиять на изменение уровня напряженности в коллективе.

Результаты исследования были занесены в специальную таблицу, в которую нами были добавлены дополнительные «расчётные» столбцы (см. *Табл. 3.4*).

Таблица 3.4

Показатели тревожности ДО и ПОСЛЕ тренинга

№ п/п респондента	*Показатель тревожности до тренинга* (x_i)	*Показатель тревожности после тренинга* (y_i)	«Расчетные» столбцы	
			$d_i = x_i - y_i$	$d_i^2 = (x_i - y_i)^2$
1	30	20	10	100
2	33	17	16	256
3	41	21	20	400
4	50	43	7	49
5	36	39	-3	9
6	45	11	34	1156
7	31	28	3	9
8	25	20	5	25
***N* = 8**			$\sum d_i = 92$	$\sum d_i^2 = 2004$

Подставив в формулу «2» для нахождения t-критерия найденные значения $\sum d_i$ и $\sum d_i^2$, получим:

$$t_{\text{набл}} = \frac{92}{\sqrt{2004 - \frac{92^2}{8}}} \times \sqrt{\frac{8-1}{8}} \approx 2,798$$

Как и в предыдущих примерах, для того чтобы осуществить сравнение $t_{\text{набл}}$ с $t_{\text{кр}}$, необходимо найти число степеней свободы (*df*). В данном случае ***df*** = 8−1= 7.

В соответствии с полученным числом степеней свободы, находим по таблице критические значения и получаем:

$$t_{\text{кр}} = \begin{cases} 2,365 \text{ для p} \le 0,05 \\ 3,499 \text{ для p} \le 0,01 \end{cases}$$

Отсюда: 2,365< 2,798 <3,499.

Вывод: Принимается H_1 гипотеза. Различия в уровнях тревожности до и после тренинга следует признать статистически значимыми (*p*<0,05), так как наблюдаемое (эмпирическое) значение превышает первое критическое (табличное), но меньше второго. Следовательно, посредством данного тренинга действительно можно добиться снижения уровня напряженности в коллективе в критических ситуациях.

Все разновидности критерия Стьюдента являются параметрическими и основаны на дополнительном предположении о нормальности выборки данных. Поэтому перед применением критерия Стьюдента рекомендуется выполнить предварительную проверку нормальности. Если гипотеза нормальности отвергается, можно проверить другие распределения, если и они не подходят, то следует воспользоваться непараметрическими статистическими тестами.

F-критерий Фишера (для сравнения дисперсий). *F - критерий Фишера* также является параметрическим критерием и используется для сравнения дисперсий двух вариационных рядов [10; 13, с. 388−389].

F-критерий Фишера используется:

1) в регрессионном анализе, позволяет оценивать значимость линейных регрессионных моделей. В частности, он используется для проверки целесообразности включения или исключения независимых переменных (признаков) в регрессионную модель;

2) в дисперсионном анализе, позволяет оценивать значимость факторов и их взаимодействия.

F-критерий Фишера основан на дополнительных предположениях о независимости и нормальности выборок данных. Поэтому перед его применением рекомендуется выполнить предварительную проверку нормальности распределения.

Наблюдаемое (эмпирическое) значение критерия вычисляется по формуле:

$$\varphi_{\text{набл}} = \frac{\delta_1^2}{\delta_2^2}$$

где ● σ_1^2 – бо́льшая дисперсия;

● σ_2^2 – меньшая дисперсия рассматриваемых вариационных рядов.

С помощью данного критерия проверяется гипотеза H_0: дисперсии сравниваемых совокупностей равны между собой, то есть $H_0=\{\delta_x^2=\delta_y^2\}$.

Критическое значение критерия Фишера так же как и во всех предыдущих случаях, касающихся иных критериев, находится по специальной таблице (см. Приложение 2, *табл. Ж*), исходя из уровня значимости α и степеней свободы:

$\boldsymbol{df_1}$ = n_1−1 (числителя);

$\boldsymbol{df_2}$ = n_1−1 (знаменателя),

где n_1 и n_2 – число респондентов, составивших 1-ю и 2-ю выборки, соответственно.

Если вычисленное значение критерия $\varphi_{\text{набл}}$ больше критического $\varphi_{\text{кр}}$ для определенного уровня значимости (α) и соответствующих чи сел степеней свободы для числителя и знаменателя (df_1 и df_2), то гипотеза H_0 отвергается и принимается альтернативная гипотеза H_1; дисперсии считаются различными.

Ниже проиллюстрировано применение критерия Фишера [13, с. 389].

Пример 3.4.4. Дисперсия такого показателя, как стрессоустойчивость для учителей составила 6,17 (n_1 = 32), а для менеджеров 4,41 (n_2 = 33). Определим, можно ли считать уровень дисперсий примерно одинаковым для данных выборок на уровне значимости α=0,05.

Для ответа на поставленный вопрос определим наблюдаемое значение критерия:

$\varphi_{\text{набл}} = \frac{6{,}17}{4{,}41} \approx 1{,}4$, при этом df_1 = 32−1=31, а df_2 = 33−1 = 32.

Исходя из проделанных вычислений, критическое значение критерия $\varphi_{\text{кр}}$ для уровня значимости α=0,05, с учетом степеней свободы равных 31 и 32, $\varphi_{\text{набл}}$=1,4 < $\varphi_{\text{кр}}$ =2. Следовательно, нулевая гипотеза о равенстве генеральных дисперсий на уровне значимости 0,05 принимается.

Вывод: Так как наблюдаемое (эмпирическое) значение критерия Фишера меньше критического (табличное) значения, то статистически значимых различий дисперсий в первой и второй группах нет и, следовательно, уровень стрессоустойчивости у учителей и менеджеров одинаков.

Вопросы и задания для самоконтроля

1. Дайте определения: нормальное распределение, нормальная площадь, нормальное отклонение, колоколообразное распределение.

2. Объясните, в каком смысле «нормальная» кривая является нормальной.

3. Что такое «сигма-единица», «сигма-точка»?

4. Каковы виды идеальных распределений?

5. Опишите, как распределения делятся в зависимости от их симметрии? Что могут означать подобные отклонения от «идеальных» распределений?

6. Установите долю случаев, лежащих между каждой парой сигма-точек на основной линии нормальной кривой. Представьте результат графически.

0,3 до 1,6	*1,1 до 1,2*	*–2,58 до +2,58*
0,3 до –1,6	*0,0 до 1,0*	*1,5 до 3,0*
0,1 до 0,2	*1,0 до 2,0*	*–2,3 до +2,3*

7. Между какими двумя сигма-точками на основной линии нормальной кривой лежат средние 50% случаев?

8. Объясните, почему доля случаев между 0 и 1,0 сигма не равна доле между 1,0 и 2,0?

9. Среднее нормальное распределение равно 75, а стандартное отклонение равно 3: А) какая доля значений лежит между 72 и 78? Б) какие значения находятся в верхнем дециле распределения? В) какая приблизительно доля лежит между 69 и 81?

10. Перечислите критерии нормальности распределения. Каков смысл этих критериев с точки зрения эмпирической социологии?

11. В чем разница между двусторонними и односторонними статистическими критериями?

12. Что понимается под мощностью статистического критерия?

13. «Нулевая гипотеза»: определение понятия, пример формулировки.

14. Чем статистическая гипотеза отличается, например, от той, которая формулируется социологом в программе эмпирического исследования?

15. Каковы типы критериев значимости? В чем между ними разница?

16. Определите сущность и опишите принцип нахождения критерия асимметрии (As).

17. В чем особенности критерия Пирсона Хи-квадрат (как критерия проверки гипотезы H_0*)?*

18. Что такое «число степеней свободы» и зачем обязательно определять это число при проверке статистического критерия?

19. В чем заключается сущность Т-критерия Стьюдента (как критерия, основанного на нормальном распределении)?

20. Чем отличаются параметрические и непараметрические критерии?

21. В каких случаях исследователя интересует F-критерий Фишера?