Лекция 11

ҚАЗАҚСТАННЫҢ АШЫҚ УНИВЕРСИТЕТІ

ОСНОВЫ СТАТИСТИКИ

Проверка гипотез. Использование Z-оценок и t-критерия Стьюдента

Здравствуйте!
В результате изучения материала, вы научитесь определять и разбирать ситуации, в которых уместны однотипные проверки гипотез, объяснять разницу между одно- и двусторонними тестами и указывать, когда каждый подходит. Определять и объяснять об ошибки типа I и типа II, использовать распределение Стьюдента, чтобы проверить значимость среднего значения для небольшой выборки.

О понятии гипотезу мы начал говорить в прошлой лекции. Вспомним что такое гипотеза. В большинстве случаев исследователь имеет дело с выборкой, которая всегда ошибочна. Выборка не может показать, например, чему в точности равна средняя или доля по генеральной совокупности. По выборке можно получить только оценку, т.е. приближенное значение этой характеристики (параметра). Чтобы по таким оценкам делать строгие выводы, необходимо вначале рассчитать, а затем сделать поправку на возможные отклонения оценки от истинного значения. Представим, что мы много-много раз (скажем, 1000) извлекаем выборки из некоторой генеральной совокупности и в каждой из них рассчитываем среднее арифметическое. Если выборки достаточны большие (более 30-ти наблюдений), то в силу действия центральной предельной теоремы выборочные средние будут распределены по нормальному закону с истинным средним в центре. Смоделировать такой эксперимент MS Excel.

Возьмем «генеральную совокупность», пусть даже с равномерным распределением от 0 до 1000. Извлечем из нее 1000 выборок по 30 наблюдений и отобразим распределение средних на гистограмме.

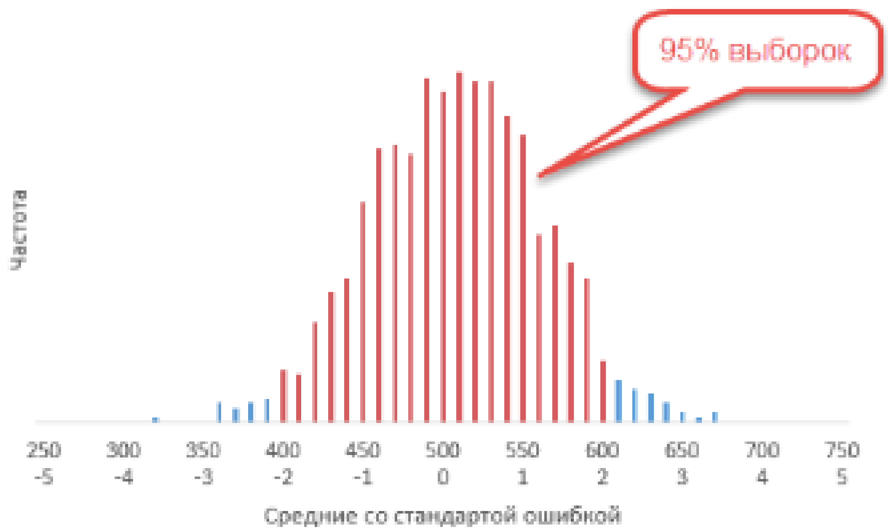

Рисунок 11.1 Распределение средних

В 95% среднее окажется в пределах ± 1,96 стандартных ошибок от истинной средней (матожидания). В остальных 5% средние отклонятся дальше. При однократном эксперименте мы имеем довольно мало шансов получить выборку со средней, выходящей за пределы ± 1,96 стандартной ошибки. И гораздо меньше шансов получить выборку со средней, выходящей за пределы ± 3 стандартной ошибки (3 случая из 1000). Это известные свойства нормального распределения.

Метод проверки гипотез

В реальности истинная средняя по генеральной совокупности неизвестна и ее значение можно только предполагать. Такое предположение называется статистической гипотезой, обозначается H. Если предположение противоречит наблюдаемым данным, то гипотезу отклоняют, как ложную; если не противоречит, то не отклоняют. Степень противоречия определяется вероятностью, которая в свою очередь зависит от того, как далеко фактическая выборочная средняя отклоняется от гипотезы. Если она (вероятность) достаточно маленькая, то противоречие считается доказанным (не забывая о возможной ошибке). Для расчета вероятности выбирают вероятностно-статистическую модель, которая описывает поведение оценки при многократном повторении эксперимента. В случае со средней арифметической в большой выборке подойдет стандартное нормальное распределение.

Рисунок 11.2 Распределение выборочных средних, в случае, если гипотеза верна.

Теперь нужно определить, какова вероятность извлечь из такой генеральной совокупности имеющуюся выборочную среднюю. Если она окажется в зоне близкой к центру, то это не противоречит гипотезе, ведь такое вполне может произойти в силу случайности. Но если она окажется далеко, например, выйдет за пределы ± 1,96 стандартные ошибки, то это будет означать что, либо произошло маловероятное событие, либо выдвинутая гипотеза ложна и ее следует отклонить.

Правила проверки гипотезы (статистического вывода) показаны на рисунке.

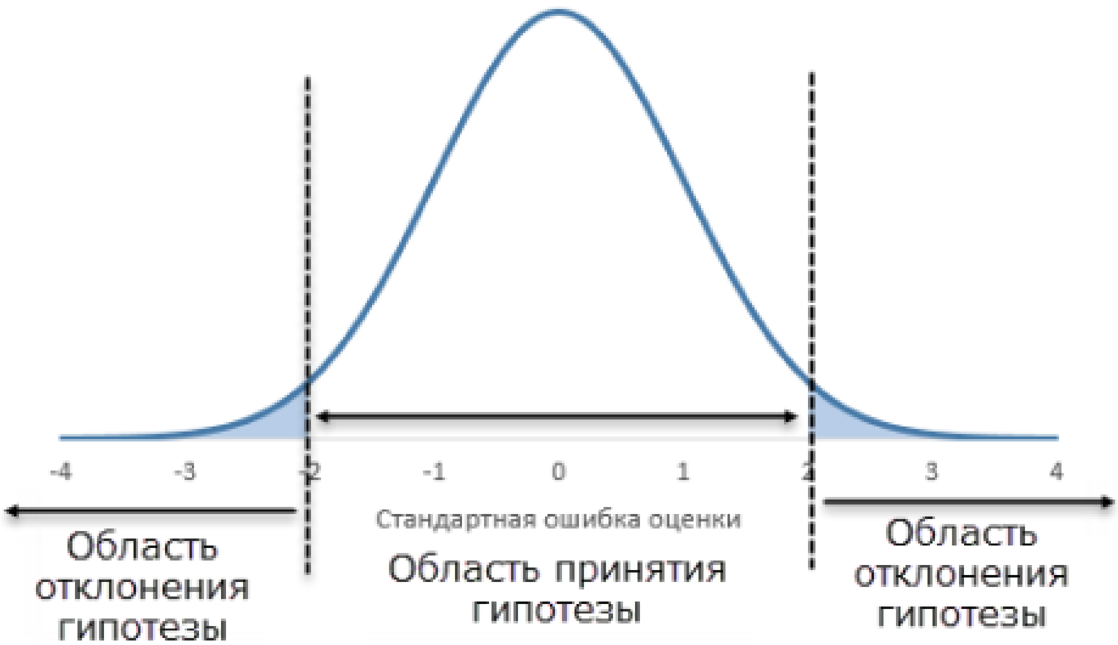

Рисунок 11.3 Правила проверки гипотезы.

Предельное значение, которое разделяет области принятия и отклонения гипотезы, называется критическим уровнем. Область отклонения гипотезы – критическая область. Вероятность, соответствующая критической области, – уровень значимости, обозначается греческой буквой α (альфа). Например, α = 0,05 означает, что уровень значимости равен 5%. Очевидно, что между критическим уровнем и уровнем значимости существует функциональная взаимосвязь.

Чтобы определить, в какую область попадает выборочная средняя, нужно рассчитать статистический критерий. Большие значения критерия, как правило, свидетельствуют в пользу того, что отличие не случайно и, соответственно, гипотеза не верна. Статистический критерий для нормальной модели – это обычная z-оценка, рассчитываемая по известной формуле.
Формула 11.1

$$z = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

где:
z – критерий;
$\bar{x}$ – наблюдаемое среднее арифметическое;
μ – гипотетическая средняя в генеральной совокупности;
s – среднеквадратическое отклонение выборочных данных;
n – объем выборки.
Если рассчитанный критерий оказывается по модулю больше, чем критическое значение, т.е. попадает в критическую область, значит, гипотеза отклоняется как ложная (точнее, маловероятная).

Рисунок 11.4 Случай, когда гипотеза отклоняется.

Если критерий не выходит за критическое значение, то гипотеза не отклоняется.

Рисунок 11.5 Критерий принятия гипотезы.

Уровень значимости задается исходя из практических соображений. Часто берут 0,05, для которого критический уровень равен 1,96 (в нормальной модели). Если $\alpha = 0{,}01$, то критический уровень – 2,58. Все это легко получить из таблиц стандартного нормального распределения. В зависимости от выбранной вероятностно-статистической модели вид распределения и способ расчета критерия производится по-разному. Но суть остается прежней: статистический критерий сравнивается с критическим значением, который задается исходя из желаемого уровня значимости.

P-value

Изложенная выше методика на сегодняшний день несколько устарела. Дело в том, что, сравнивая критерий с критическим уровнем, мы не видим «силу доказательства». Ведь критерий может попасть в область, соответствующую 5% уровню значимости, а может и в 1% значимости (т.е. отклониться еще дальше). В обоих случаях гипотеза отклоняется, но уверенность, с которой это делается, будет разной. Одно дело «скорее всего» (как при 5%-м уровне), а другое «наверняка» (как при 1% уровне). Поэтому проверку гипотезы делают по наблюдаемому уровню значимости, который обычно называют P-value (или P-значение).
P-level – это вероятность получить наблюдаемое или еще более ярко выраженное отклонение оценки от гипотезы, если она верна. Геометрически это площадь под кривой, которая начинается от статистического критерия.

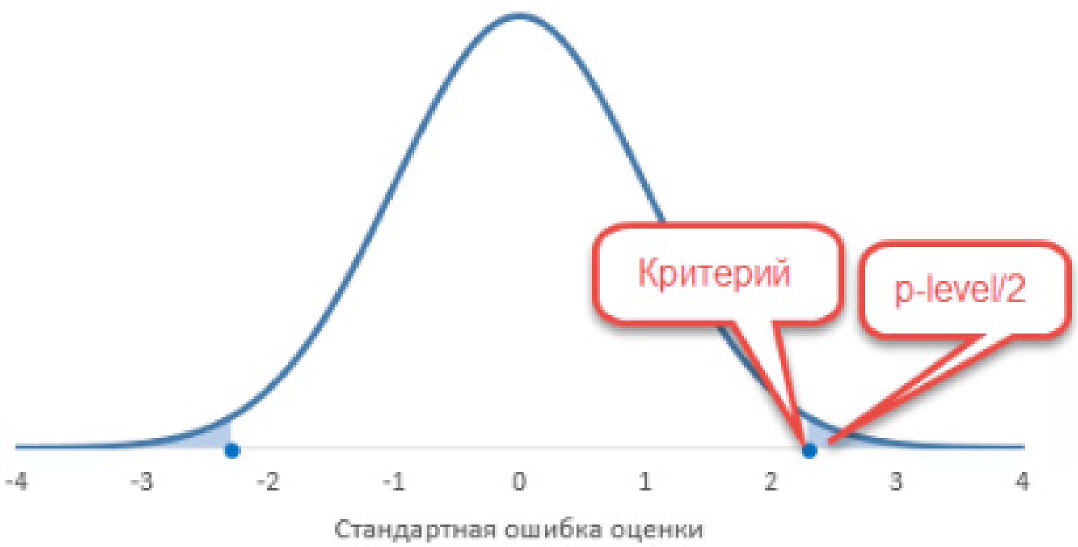

Рисунок 11.6 Определение P-уровня.

Общий P-value на данном рисунке складывается из двух частей, т.к. гипотеза рассматривает отклонение в любую из сторон.
Например, если статистический критерий равен 1,96, то вероятность получить по модулю такое или еще большее значение, равна 0,05. Это и есть P-level, который в данном случае совпал с уровнем значимости. Но если критерий равен 3, то вероятность получить такое или еще больше отклонение (по модулю) равна всего 0,0027. Т.к. мы считаем возможным отклонение в обе стороны, P-level складывается из двух частей.

Рисунок 11.7 Двустороннее отклонение.

Итак, правило проверки гипотезы по наблюдаемому уровню значимости следующее: если P-level меньше, чем заданный уровень значимости (например, 0,05), то гипотеза отклоняется. В противном случае не отклоняется (не отвергается). В примере выше P-level = 0,0027, что гораздо меньше, чем 0,05. Следовательно, гипотеза отвергается.

1 и 2 сторонний критерий.
Рассмотрим еще несколько важных понятий. Выше был показан двухсторонний критерий, это когда проверка на отклонение производится в обе стороны.
Иногда имеет смысл рассматривать отклонение только в одну сторону. Например, если заранее известно, что отклонение от гипотезы возможно только в сторону увеличения, то левый хвост не рассматривают. Такой критерий называется односторонним. Использование одностороннего критерия вместо двухстороннего при заданном уровень значимости (α) приводит увеличению мощности критерия (его способности обнаружить эффект), что очень даже хорошо. Но про мощность поговорим в другой раз.
Вот, как на диаграмме выглядит односторонний критерий.

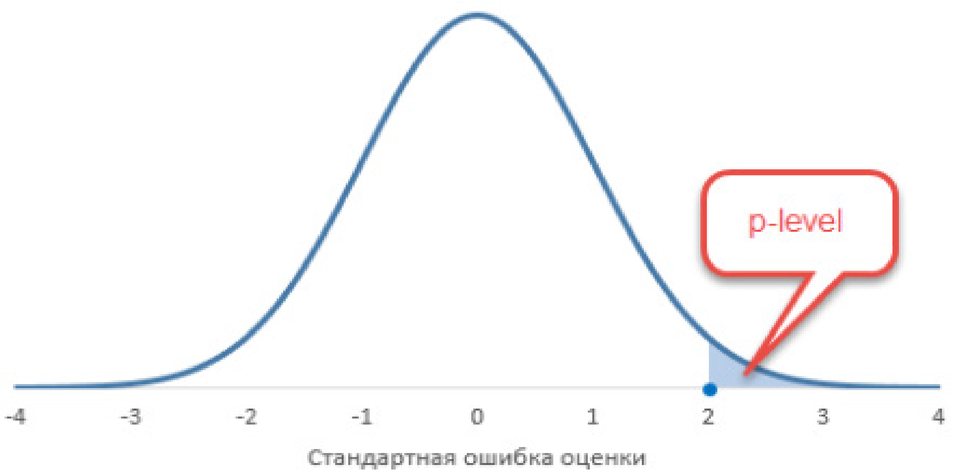

Рисунок 11.8 Односторонний критерий.

Однако одностороннюю гипотезу нужно формировать заранее. Нельзя для повышения убедительности выводов после проведения анализа менять двухсторонний критерий на односторонний. Это будет подгонка фактов под теорию, что увеличивает вероятность совершить ошибку.

Альтернативная гипотеза

Проверяемая гипотеза называется основной или нулевой. Она подразумевает некоторый status quo, когда между проверяемыми данными нет отличий. Гипотеза остается в силе, если оценка отклонятся не слишком далеко и находится в зоне возможных случайных колебаний.
Кроме основной (нулевой) гипотезы рассматривают альтернативную или конкурирующую. Формально, альтернативная гипотеза – это любое предположение о параметрах распределения, не совместимое с нулевой гипотезой. Однако на практике разнообразие проверяемых и альтернативных гипотез довольно ограничено. Например, основная гипотеза (нулевая) заключается в том, что средняя равна некоторому значению, а альтернативная – не равна этому значению.

Нулевая гипотеза обозначается H0, альтернативная H1. Краткая запись условия задачи при использовании двухстороннего критерия имеет следующий вид:
H0: $\mu = a$
H1: $\mu \neq a$

Если рассматривается односторонний критерий, то запись может иметь такой вид:
H0: $\mu = a$
H1: $\mu > a$

При отклонении нулевой гипотезы, автоматически принимается альтернативная.

Следует отметить, что предметом доказательства, как правило, является именно конкурирующая гипотеза. То есть проверяя равенство средних в двух выборках, исследователя интересует их различие, которое должно подтвердить влияние некоторого воздействия на предмет исследования (новое лекарство, новых способ обработки материала и др.). Если есть влияние, то будет и различие, если нет, то средние будут отличаться не очень сильно, в пределах случайных колебаний оценок.

Статистический вывод

Заострим внимание на корректности статистических выводов. Вместо выражения «гипотеза не отклоняется» часто говорят «гипотеза принимается». В целом, это выражение также приемлемо, если его понимать правильно, т.е. если считать, что принимается именно гипотеза (одно из возможных объяснений), а не конкретное утверждение. Но понимают его часто неправильно, подразумевая, что в случае не отклонения гипотезы принимается сама идея гипотезы. Например, если гипотеза о равенстве вероятностей в двух выборках не отклоняется, то делают заключение, что, мол, вероятности действительно равны. Такое заключение ошибочно.

На самом деле принятие гипотезы означает, что она не противоречит данным и может рассматриваться до тех пор, пока не будет доказано обратное. Принятие гипотезы не может доказать ее правильность, для этого есть лишь один способ: исследовать все анализируемое

явление в целом, собрав генеральную совокупность. По выборке можно только опровергнуть маловероятные или невозможные предположения, противоречащие фактическим данным, сузив тем самым круг для поиска истины.
Проще говоря, выдвинув ту или иную гипотезу, исследователь задает вопрос: может ли такое быть, чтобы при имеющихся данных имело место вот это событие (нулевая гипотеза об отсутствии различий или взаимосвязей)? Ответа здесь только два: 1) да, может; 2) нет, не может. Нулевую гипотезу можно только опровергнуть, но не доказать.

Далее мы рассмотрим пятиступенчатую модель для организации всех проверок гипотезы.
Шаг 1. Создание предположений и выполнение требований теста.
Шаг 2. Изложение нулевой гипотезы.
Шаг 3. Выбор распределения выборки и определение критической области.
Шаг 4. Вычисление тестовой статистики.
Шаг 5. Принятие решения и интерпретация результатов теста.

Мы рассмотрим каждый шаг в отдельности.
Шаг 1. Создание предположений и выполнение требований теста. Любое применение статистики основано на ряде предположений. Для проверки гипотез с выборочным средним должны быть выполнены три предположения:
• Образец был выбран в соответствии с правилами.
• Проверяемая переменная является интервалом в уровне измерения.
• Распределение выборки всех возможных выборочных средств нормальное по форме. Это позволит нам использовать стандартное нормальное распределение, чтобы найти области под распределением выборки. Мы можем быть уверены, что это предположение выполняется с помощью больших выборок (см. Центральную предельную теорему).
Обычно мы будем формулировать эти предположения в сокращенной форме в качестве математической модели для теста. Например, Модель: случайная выборка. Уровень измерения интервал-отношение – нормальное распределение выборки.

Шаг 2. Изложение нулевой гипотезы. Нулевая гипотеза - это всегда утверждение «без разницы», но ее точная форма будет варьироваться в зависимости от проводимого теста. В случае единичной выборки нулевая гипотеза утверждает, что выборка происходит из популяции с определенной характеристикой.

Шаг 3. Выбор распределения выборки и определение критической области. Распределение выборки – это вероятностный критерий, по которому измеряется результат конкретной выборки. Предполагая, что нулевая гипотеза верна (и только этим предположением), мы можем привязать значения к среднему и стандартному отклонению распределения выборки и измерить вероятность любого конкретного результата выборки.

Шаг 4. Вычисление тестовой статистики. Чтобы оценить вероятность исхода выборки, мы должны преобразовать значение выборки в Z-оценку. Нахождение этой Z-оценки называется вычислением тестовой статистики, и результирующее значение будет обозначаться как Z (получено), чтобы отличать его от значения, которое обозначает начало критической области.

Шаг 5. Принятие решения и интерпретация результатов теста. Если тестовая статистика попадает в критическую область, наше решение будет отклонять нулевую гипотезу. Если тестовая статистика не попадает в критическую область, мы не можем отвергнуть нулевую гипотезу.
Эта пятиступенчатая модель послужит основой для принятия решений, проверке гипотез.

Точный характер и способ выражения наших решений будут различными для разных ситуаций. Тем не менее, пятиступенчатая модель будет использована во всех случаях.

Распределение t-критерия Стьюдента для проверки гипотезы о средней и расчета доверительного интервала в MS Excel.
Проверка статистической гипотезы позволяет сделать строгий вывод о характеристиках генеральной совокупности на основе выборочных данных. Гипотезы бывают разные. Одна из них – это гипотеза о средней (математическом ожидании). Суть ее в том, чтобы на основе только имеющейся выборки сделать корректное заключение о том, где может или не может находится генеральная средняя (точную правду мы никогда не узнаем, но можем сузить круг поиска). Общий подход в проверке гипотез мы рассмотрели ранее, поэтому сразу к делу. Предположим для начала, что выборка извлечена из нормальной совокупности случайных величин X с генеральной средней μ и дисперсией σ2 (в реальности так не бывает, но мы рассмотрим в качестве теоретического примера). Средняя арифметическая из этой выборки, очевидно, сама является случайной величиной. Если извлечь много таких выборок и посчитать по ним средние, то они также будут иметь нормальное распределение с математическим ожиданием μ и дисперсией.
Формула 11.2:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

Тогда случайная величина:

Формула 11.3

$$\frac{\bar{X} - \mu}{\sigma_{\bar{x}}}$$

будет иметь стандартное нормальное распределение со всеми вытекающими отсюда последствиями. Например, с вероятностью 95% ее значение не выйдет за пределы ±1,96.
Однако такой подход будет корректным, если известна генеральная дисперсия. В реальности, как правило, она не известна. Вместо нее берут оценку – **несмещенную выборочную дисперсию:**
Формула 11.4
где:

$$s_{\bar{x}}^2 = \frac{s^2}{n}$$

$$s^2 = \frac{\sum(X_i - \bar{X})^2}{n-1}$$

Формула 11.5

Возникает вопрос: будет ли генеральная средняя с вероятностью 95% находиться в пределах ±1,96s$\bar{x}$. Другими словами, являются ли распределения случайных величин

$$\frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \quad \text{и} \quad \frac{\bar{X} - \mu}{s_{\bar{X}}} \quad \text{эквивалентными.}$$

Сгенерируем 20 тысяч нормальных выборок из 6-ти наблюдений со средней ($\bar{X}$) 50 и среднеквадратичным отклонением (σ) 10. Затем нормируем выборочные средние, используя генеральную дисперсию:

$$\frac{\bar{X}_i - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X}_i - 50}{10/\sqrt{6}}$$

Получившиеся 20 тысяч средних сгруппируем в интервалы длинной 0,1 и подсчитаем частоты. Изобразим на диаграмме фактическое (Norm) и теоретическое (ENorm) распределение частот выборочных средних.

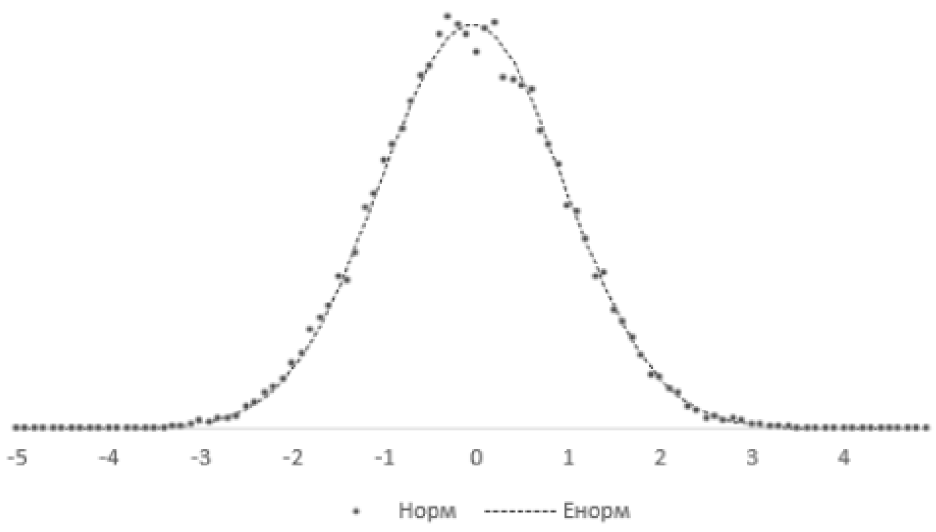

Рисунок 11.9 Распределение частот выборочных средних.

Точки (наблюдаемые частоты) практически совпадают с линией (теоретическими частотами). Оно и понятно, ведь данные взяты из одной и то же генеральной совокупности, а отличия – это лишь ошибки выборки. Проведем новый эксперимент. Нормируем средние, используя выборочную дисперсию.

$$\frac{\bar{X}_i - \mu}{s_{\bar{X}_i}} = \frac{\bar{X}_i - 50}{s_{\bar{X}_i}}$$

Снова подсчитаем частоты и нанесем их на диаграмму в виде точек, оставив для сравнения линию стандартного нормального распределения. Обозначим эмпирическое частоты средних, скажем, через букву t.

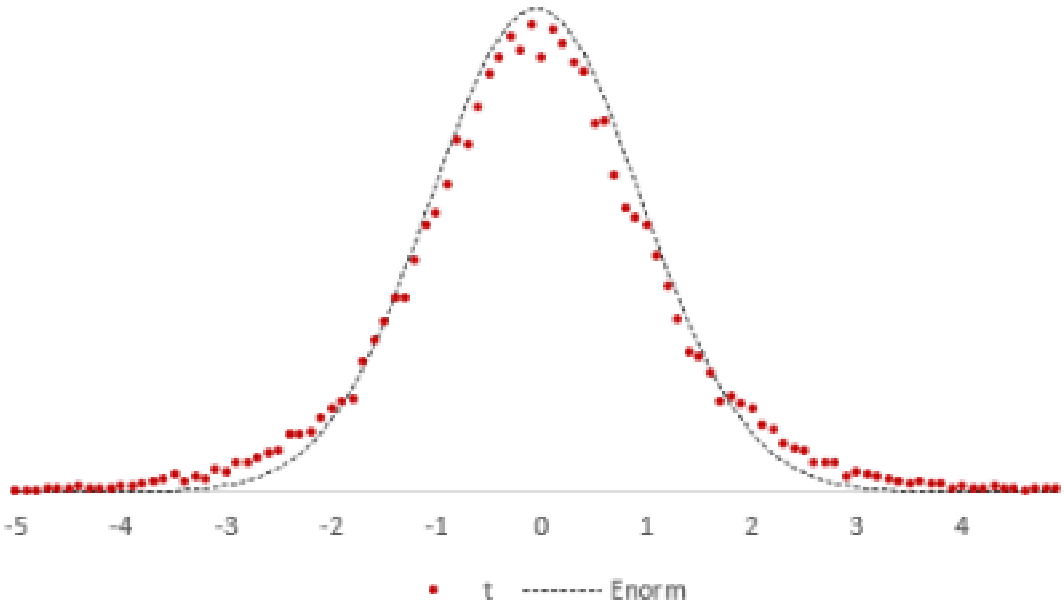

Рисунок 11.10 Линия стандартного нормального распределения.

Видно, что распределения на этот раз не очень-то и совпадают. Близки, да, но не одинаковы. Хвосты стали более «тяжелыми». Почему так получается? Объяснение заключается в том, что случайная величина зависит не только от ошибки выборки (числителя), но и от стандартной ошибки средней (знаменателя), которая также является случайной величиной.

Применим всю вышеизложенную теорию для примера.
Для случайной выборки из 152 уголовных дел, рассмотренных в местном суде, средний срок тюремного заключения составил 27,3 месяца.

Значительно ли это отличается от среднего тюремного срока для уголовников по штату (y = 28,7)? Мы будем использовать пятиступенчатую модель для организации процесса принятия решений.

Шаг 1. Сделайте допущения и проведите тестирование
Требования.
Модель: случайная выборка.
Уровень измерения интервал-отношение Распределение выборки в норме.
Это большая выборка (A> 100), а длина предложения является переменной отношения между значениями, поэтому мы можем сделать вывод, что предположения модели выполнены.

Шаг 2. Изложение нулевой гипотезы (H0). Нулевая гипотеза гласит, что среднее предложение на местном уровне (для всех уголовных дел) равно среднему по региону. В обозначениях: H0: μ=28,7. Вопрос исследования не определяет направление, он только спрашивает, отличаются ли местные предложения от (не выше или ниже) среднего значения по штату. Это предполагает двусторонний тест: H1: μ≠28,7.

Шаг 3. Выбор распределения выборки и определение критического региона. Это большая выборка, поэтому мы можем использовать Приложение А, чтобы установить обогащенную область и указать критические оценки как Z-оценки (в отличие от t-оценок). Распределение выборки = распределение Z, альфа= 0,05, Z (критическое значение) = ±1,96

Шаг 4. Вычисление тестовой статистики. Необходимая информация для проведения проверки нулевой гипотезы

$$\bar{X} = 27.3, \quad s = 3.7, \quad N = 152, \; \mu = 28.7$$

Тестовая статистика Z (получена) равна

$$Z = \frac{\bar{X}-\mu}{s/\sqrt{N-1}} = \frac{27.3-28.7}{3.7/\sqrt{152-1}} = \frac{-1.40}{3.7/\sqrt{151}} = \frac{-1.40}{3.7/12.29} = \frac{-1.40}{0.30} = -4.67$$

Обратите внимание, что, хотя размер выборки велик, мы использовали «N - 1», а не «N» под знаком квадратного корня. Мы будем делать это всегда, когда σ неизвестно, даже когда влияние на значение Z будет очень небольшим.

Шаг 5. Принятие решения и интерпретация результатов теста. При Z (критическом) + 1,96 и полученном значении Z -4,67 нулевая гипотеза отвергается. Разница между тюремным заключением осужденных в местном суде и уголовных преступников по всему штату статистически значима. Разница настолько велика, что мы можем заключить, что это произошло не случайно. Решение отклонить нулевую гипотезу с вероятностью 0,05 ошибочно.

Использование t-распределения в тесте гипотезы.
Чтобы более подробно продемонстрировать использование распределения, рассмотрим примерную проблему. Обратите внимание, что с точки зрения пятиступенчатой модели изменения происходят в основном на этапах 3 и 4. На этапе 3 распределение выборки будет распределение t, и степени свободы (df) должны быть вычислены до определения критической области или оценки t (критической). На шаге 4 будет использоваться немного другая формула для вычисления статистики теста, t (получено). По сравнению с формулой для Z (получено), s заменит σ, а N - 1 заменит N. В частности,

$$t = \frac{\bar{X}-\mu}{s/\sqrt{N-1}} .$$

Исследователь задается вопросом, отличаются ли приезжие студенты от общего числа студентов с точки зрения академической успеваемости. Она собрала случайную выборку из 30 учеников пригородной зоны и узнала от регистратора, что средний балл по всем учащимся составляет 2,50 (μ= 2,50), но стандартное отклонение для населения (σ) никогда не вычислялось. Пример данных приведен здесь. Является ли выборка из популяции, которая имеет среднее значение 2,50?

$$\mu=2{,}50,\ \overline{X}=2{,}78,\ s=1.23,\ N=30,\ \sigma=?$$

Шаг 1. Создание предположений и выполнение требований теста.
Модель: случайная выборка
Уровень измерения интервал-отношение.
Распределение выборки в норме.

Шаг 2. Изложение нулевой гипотезы. Из гипотезы исследования видно, что исследователь не предсказал направление разницы. Это будет двусторонний тест.

$$H_0:\ \mu = 2.50$$

$$H_1:\ \mu \neq 2.50$$

Шаг 3. Выбор распределения выборки и определение критической области. Поскольку σ неизвестно, а размер выборки невелик, распределение t будет использовано для нахождения критической области. Альфа будет установлена на уровне 0,01.
Распределение выборки = распределение t

$\alpha=0{,}01$, df=(N-1)=29, t крит=±2,756

Шаг 4. Вычисление тестовой статистики.

$$t = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{N-1}}} = \frac{2.78-2.50}{1.23/\sqrt{29}} = \frac{0.28}{1.23/5.39} = \frac{0.28}{0.23} = 1.22$$

Шаг 5. Принятие решения и интерпретация результатов теста. Статистика теста не попадает в критическую область. Поэтому исследователь не может отклонить H0. Разница между средним по выборке (2,78) и средним по популяции (2,50) не является статистически значимой. Разница не превышает ожидаемого результата, если бы действовал только случайный шанс. Статистика теста и критические области показаны на рисунке:

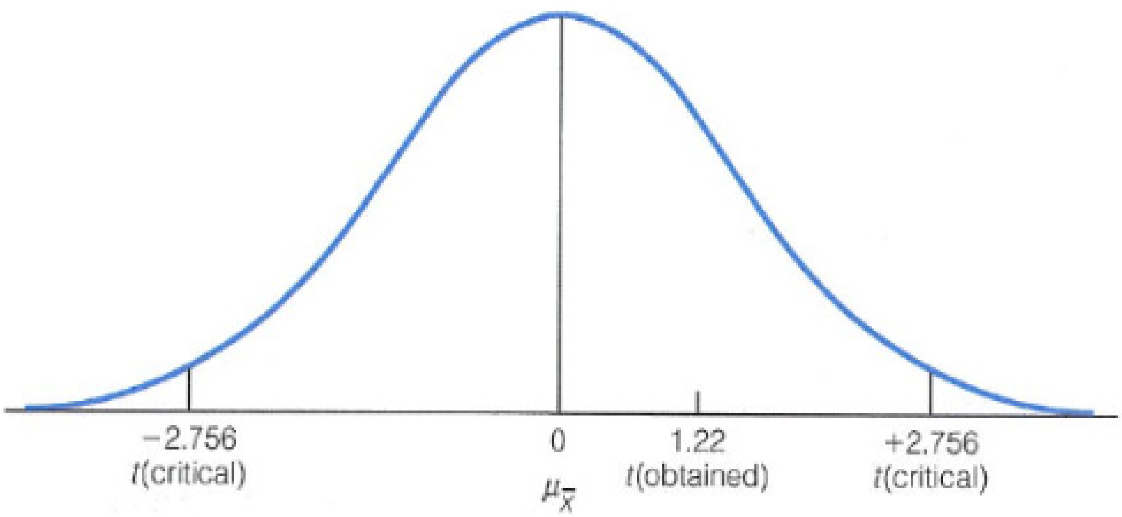

Рисунок 11.11 Статистика теста и критические области.

Подводя итоги, мы должны сделать выбор в отношении теоретического распределения, которое мы будем использовать для установления критической области. Выбор прост. Если стандартное отклонение совокупности (σ) известно или размер выборки велик, будет использоваться распределение Z. Если σ неизвестен, а выборка мала, будет использовано распределение t.

Проверка гипотезы для пропорций одного образца (большие образцы)
Во многих случаях мы работаем с переменными, которые не являются интервальными отношениями в уровне измерения. Одной из альтернатив в этой ситуации было бы использование выборочной пропорции (Ps), а не выборочного среднего, в качестве статистического теста. Как мы увидим, общие процедуры для тестирования пропорций одного образца такие же, как и для средств тестирования. Центральный вопрос по-прежнему остается: «Имеет ли популяция, из которой была отобрана выборка, определенную характеристику?». Мы по-прежнему проводим тест, основываясь на предположении, что нулевая гипотеза верна, и мы все еще оцениваем вероятность полученного результата выборки по выборочное распределение всех возможных результатов выборки. Наше решение в конце теста также, то же самое. Если полученная статистика теста попадает в критическую область (и, следовательно, маловероятно, учитывая предположение, что H0 верно), мы отклоняем H0.

Конечно, есть также некоторые важные различия в тестах значимости для пропорций выборки. Эти различия лучше всего связаны с точки зрения пятиступенчатой модели:
1. На шаге 1 мы предполагаем, что переменная измеряется на номинальном уровне.
2. На шаге 2 символы, используемые для формулировки нулевой гипотезы, отличаются, хотя это все еще утверждение «без разницы».
3. На шаге 3 мы будем использовать только распределение Z, чтобы найти области под распределением выборки и найти критическую область. Это будет уместно, если размер выборки велик. Мы не будем рассматривать небольшие выборочные проверки гипотезы о пропорциях в этом тексте.
4. На шаге 4 вычисления статистики теста форма формулы остается прежней. Таким образом, тестовая статистика Z (полученная) равна статистике выборки минус среднее значение распределения выборки, деленное на стандартное отклонение распределения выборки. Однако символы меняются, потому что мы основываем тесты на пропорциях выборки. Формула может быть сформулирована как:

Формула 11.6

$$Z = \frac{P_s - P_U}{\sqrt{P_U(1-P_U)/N}}$$

5. Шаг 5 принятия решения точно такой же, как и раньше. Если тестовая статистика Z (получена) попадает в критическую область, отклоните H0.
Пример должен прояснить эти процедуры.

Случайная выборка из 122 домохозяйств в районе с низким доходом показала, что 53 (или доля 0,43) домохозяйств возглавляли женщины. В целом по городу доля домохозяйств, возглавляемых женщинами, составляет 0,39. Являются ли домохозяйства соседями с низкими доходами?
Шаг 1. Создание предположений и выполнение требований теста.
Модель: случайная выборка
Уровень измерения номинальный. Распределение выборки нормальное по форме.
Шаг 2. Изложение нулевой гипотезы. Вопрос исследования спрашивает, отличается ли доля

выборки от доли населения. Будет использован двусторонний тест, поскольку направление для разницы не было предсказано.
H0: PU=0.39 (H1 : PU≠0.39)
Шаг 3. Выбор распределения выборки и установление критического значения
Альфа=0,01, двустронний Zкрит=±1,65
Шаг 4. Вычисление тестовой статистики.

$$Z = \frac{P_s - P_U}{\sqrt{P_U(1-P_U)/N}} = \frac{0{,}43-0{,}39}{\sqrt{0{,}39 \cdot 0{,}61/122}} = \frac{0{,}04}{\sqrt{0{,}24/122}} = \frac{0{,}04}{0{,}05} = 0,80$$

Шаг 5. Принятие решения и интерпретация результатов теста. Тестовая статистика,

Z (получено), не попадает в критическую область. Поэтому мы не можем отказаться от H0. Не существует статистически значимой разницы между сообществом с низкими доходами и городом в целом с точки зрения доли домохозяйств, возглавляемых женщинами. На рисунке показано распределение выборки, критическая область и Z (расчетного).

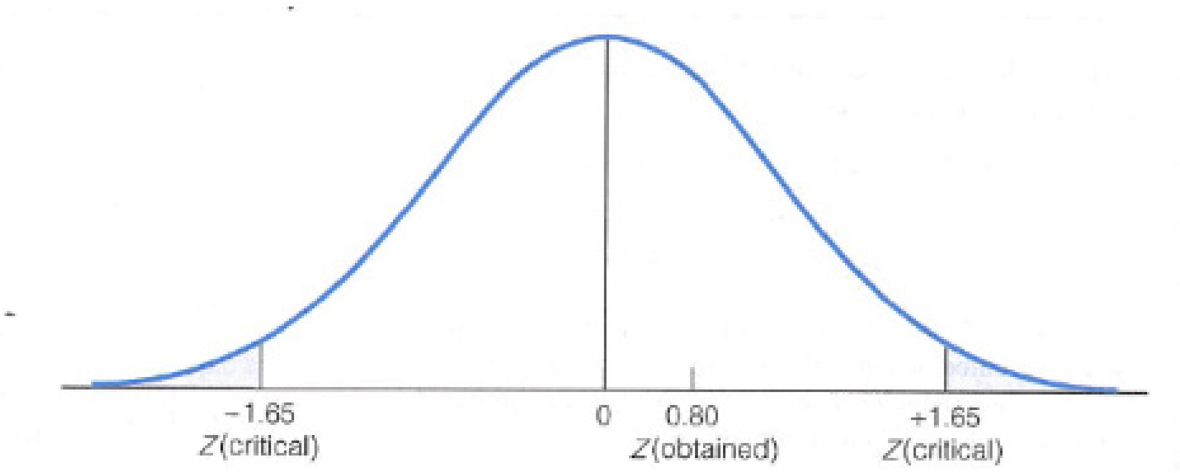

Рисунок 11.12 Распределение выборки, критическая область и Z

Выводы:

1. Мы увидели, как можно проверить нулевую гипотезу «без разницы» для средних значений и пропорций для одной выборки. В обоих случаях центральный вопрос заключается в том, имеет ли популяция, представленная выборкой, определенный характер.
2. Все тесты значимости включают в себя поиск вероятности наблюдаемого результата выборки, учитывая, что нулевая гипотеза верна. Если результат имеет низкую вероятность, мы отвергаем нулевую гипотезу. В обычной исследовательской ситуации мы хотим отказаться от нулевой гипотезы и тем самым поддержать гипотезу исследования.
3. Пятиступенчатая модель станет нашей основой для принятия решений в главах, посвященных проверке гипотез. Однако то, что мы делаем на каждом этапе, будет зависеть от конкретного проводимого теста.
4. Если мы можем предсказать направление различий в определении гипотезы исследования. требуется односторонний тест. Если направление не может быть предсказано, подходит двусторонний тест.

5. Есть два вида ошибок при проверке гипотез.
Ошибка типа I, или альфа, отвергает истинную нулевую гипотезу; Ошибка типа II, или бета, не позволяет отвергнуть ложную нулевую гипотезу. Вероятности совершения этих двух типов ошибок обратно пропорциональны и не могут быть одновременно минимизированы в одном и том же тесте. При выборе альфа-уровня мы стараемся сбалансировать достоверность этих двух ошибок.