6. Проверка статистических гипотез: непрерывные случайные величины

6.1. Вводный пример

В течение ряда лет начальная школа регистрировала возраст чтения (типичный возраст, когда ребенок начинает учиться читать) детей в начале и конце каждого учебного года. Учителя обнаружили, что в течение 3-го года обучения (третьего года обучения детей в школе) уменьшение возраста чтения обычно распределяется со средним значением 1,14 года и стандартным отклонением 0,16 года. В этом году они собираются опробовать новую схему чтения: другие школы опробовали эту схему и обнаружили, что она привела к большему уменьшению возраста чтения. В конце года учителя будут использовать среднее уменьшение возраста чтения, $\bar{x}$, из 40 детей в 3-м классе, чтобы помочь им ответить на вопрос: "дает ли новая схема чтения лучшие результаты, чем старая в нашей школе?' Трудность ответа на этот вопрос заключается в том, что каждый ребенок прогрессирует с разной скоростью, так что разные значения$\bar{x}$ будут получены для разных групп детей. Легко проверить, будет ли $\bar{x}$ больше 1,14 лет. Нелегко узнать, отражает ли значение $\bar{x}$, превышающее 1,14 года, эффективность новой схемы или это просто случайное отклонение между детьми.

Далее описывается статистический метод для принятия решения. Этот процесс состоит из нескольких этапов.

6.2.Формулировка нулевой и альтернативной гипотезы

Есть две теории о том, как новая схема чтения работает в этой конкретной школе. Во-первых, использование новой схемы не дает улучшения. Эта теория называется **нулевой гипотезой**. Обозначается символом H_0. В этом примере, где среднее значение возраста чтения в прошлом составляет 1,14 года, нулевая гипотеза может быть выражена как: $H_0: \mu = 1.14$. Обратите внимание, что H_0 предлагает единственное значение для математического ожидания генеральной совокупности,μ, которое основано на прошлом опыте.

«Гипотеза» - это теория, которая считается верной, если не получены доказательства, указывающие на обратное. «Нуль» означает «ничто», а термин «нулевая гипотеза» означает «теория без изменений»; это «без изменений» по сравнению с тем, что можно ожидать из прошлого опыта.

Другая теория заключается в том, что новая схема чтения более эффективна, чем старая, то есть математическое ожидание генеральной совокупности увеличится. Это называется альтернативной гипотезой и обозначается символом H_1. Таким образом, альтернативная гипотеза в этом случае $H_1: \mu > 1.14$. Альтернативная гипотеза предлагает способ изменения μ, если новая схема чтения будет более эффективной, чем старая.

Процедура выбора между этими двумя противоположными теориями, называется **проверкой гипотезы** или иногда **проверкой значимости**. В этом примере проверка будет **односторонней**, потому что альтернативная гипотеза предполагает изменение математического ожидания только в одном направлении, в данном случае увеличение. Также возможна односторонняя проверка в другом направлении, в котором альтернативная гипотеза предполагает уменьшение математического ожидания. Проверки, в которых альтернативная гипотеза предполагает изменение математического ожидания в обоих направлениях, называются **двусторонними**.

Пример 6.2.1.

Для следующих ситуаций приведите нулевые и альтернативные гипотезы и покажите, будет ли проверка гипотез односторонней или двухсторонней.

(а) В прошлом спортсмен пробегал 100 метров в среднем за 10,3 секунды. Он следовал новой программе тренировок, которая, как он надеется, сократит время, необходимое ему для бега на 100 метров. Он собирается рассчитать время в следующих своих шести пробегах.

(б) Средняя масса мешочков с сахаром небольшой массы, сходящие с производственной линии, должна составлять 1,01 кг. Была исследована случайная выборка, чтобы проверить, были ли какое-либо изменение этой средней массы.

(в)Средний объем лимонада в бутылках должен быть не менее 2 литров. Была исследована случайная выборка, чтобы проверить, не упал ли средний объем ниже 2 литров.

Решение:

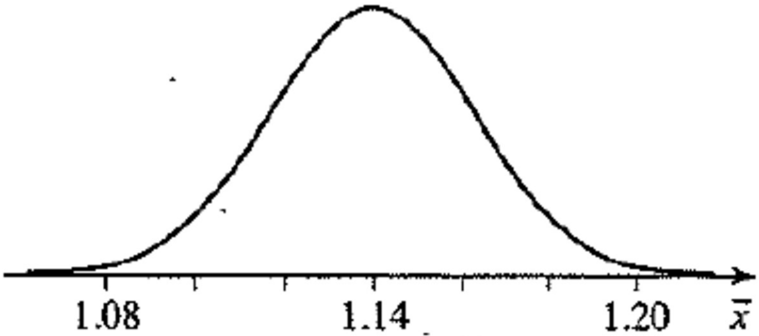
(а) Нулевая гипотеза предлагает единственное значение дляμ, $H_0: \mu = 10.3$, основанное на прошлых результатах спортсмена. Альтернативная гипотеза предлагает, как μ могло уменьшится, $H_1: \mu < 10.3$. Это односторонняя проверка.

(б) Нулевая гипотеза предлагает единственное значение для μ, $H_0: \mu = 1.01$, основанное на том, какой массы должны быть мешочки с сахаром. Альтернативная гипотеза предполагает, что μ могло измениться, $H_1: \mu \neq 1.01$ Это двусторонняя проверки.
(в) Единственное беспокойство в этой ситуации вызывает то, что средний объем составит менее 2 литров, поэтому подходит односторонняя проверка. Нулевая гипотеза $H_0: \mu = 2$ и альтернативная гипотеза $H_1: \mu < 2$

Для проверки гипотезы о математическом ожидании генеральной совокупности, $\boldsymbol{\mu}$, нулевая гипотеза предполагает значение $\boldsymbol{\mu_0}$, то есть
$\boldsymbol{H_0: \mu = \mu_0}$. Альтернативная гипотеза, H_1, предлагает способ, при которым $\boldsymbol{\mu}$ может отличаться от $\boldsymbol{\mu_0}$. Существуют три вида альтернативной гипотезы:
$\boldsymbol{H_1: \mu < \mu_0}$, односторонняя проверка на уменьшение;
$\boldsymbol{H_1: \mu > \mu_0}$, односторонняя проверка на увеличение;
$\boldsymbol{H_1: \mu \neq \mu_0}$, двухсторонняя проверка на изменение.

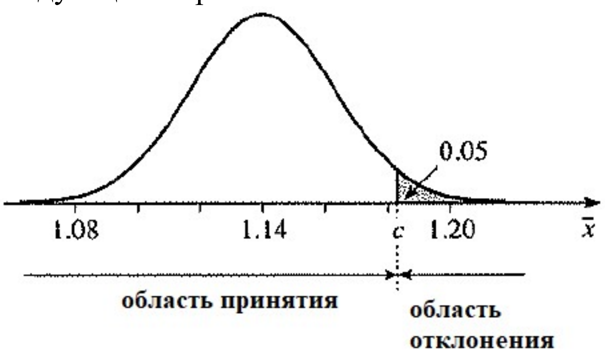
6.3. Критическое событие (критерии для нулевой гипотезы)

Как только вы определились со своими нулевыми и альтернативными гипотезами, следующим шагом будет разработка правила выбора между ними. Посмотрите еще раз на пример схемы чтения. Правило будет основано на выборочном среднем, $\bar{x}$. Учителей интересует новая схема только в том случае, если она улучшает среднее увеличение возраста чтения, и поэтому только значения $\bar{x}$, превышающие 1,14, могут привести к отказу от старой схемы в пользу новой. Вначале вы можете подумать, что любое значение $\bar{x}$, превышающее 1,14 года, покажет, что новая схема более эффективна. Немного поразмыслив, мы понимаем, что это слишком простое правило.
Можно получить выборочное среднее значение $\bar{x}$, которое больше 1,14, даже если новая схема чтения вообще неэффективна. Чтобы показать это, предположим, что нет никакой разницы между новой схемой чтения и старой. Тогда иμ, иσ будут одинаковыми по двум схемам, и случайная величина X будет распределена по нормальному закону распределения $N(1.14; 0.16^2)$.
Из центральной предельной теоремы известно, что если $X \sim N(\mu; \sigma^2)$, то $\bar{X}\left(\mu; \frac{\sigma^2}{n}\right)$. Значит для выборки из 40 учеников третьего класса, выборочное среднее, $\bar{X} \sim N\left(1.14; \frac{0.16^2}{40}\right)$, как показано на следующем рисунке:



Из этой диаграммы видно, что вероятность того, что среднее значение выборки будет больше, чем 1,14, равна $\frac{1}{2}$, даже если предположить, что среднее значение генеральной совокупности не изменилось. Насколько большим должно быть среднее значение выборки, прежде чем можно будет сделать вывод, что среднее значение генеральной совокупности будет больше чем 1.14? Большинство людей согласятся, что если получено среднее значение выборки 2.00, то маловероятно, что среднее значение по-прежнему будет 1.14, но как насчет среднего значения выборки 1.19? Одним из способов решения этой проблемы является разделение возможных результатов на две области: **область отклонения** (или **критическую**) и **область принятия**.
Область отклонения будет содержать значения в верхней части распределения выборочного среднего $\bar{X}$, показанного на вышеуказанном рисунке.
Если среднее значение выборки находится в области отклонения, вы отклоняете H_0 в пользу H_1: вы заключаете, что среднее значение генеральной совокупности. Если среднее значение выборки находится в области принятия, вы не отклоняете H_0: нет достаточных доказательств того, что новая схема чтения более эффективна.

Область отклонения выбирается таким образом, чтобы было «маловероятно», что среднее значение выборки попадает в область отклонения, когда гипотеза H_0 верна. Но что мы подразумеваем под «маловероятным», но среди статистиков обычно принято считать, что событие с вероятностью 0,05 или меньше, то есть 1 к 20 или менее, «маловероятно». На следующем рисунке показана область отклонения и область принятия для примера схемы чтения с детьми. Значение c, которое разделяет области отклонения и принятия, называется **критическим значением**. Это значение можно рассчитать следующим образом.



Так как $\Phi(z) = 0.95$, то значение случайной величины z, распределенной по стандартному нормальному закону распределения, соответствующего значению *c,* равно z=1.645. (из таблицы нормального распределения)

Значения z и $\bar{x}$ относятся как: $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$, где Z~N(0;1).

Подставим известные нам значения: $1.645 = \frac{c-1.14}{\sqrt{\frac{0.16^2}{40}}}$

$c = 1.645\sqrt{\frac{0.16^2}{40}} + 1.14 = 1.18$, с точностью до трех значащих цифр.

Таким образом, область отклонения $\bar{X} \geq 1.18$ года.

В конце учебного года наблюдаемое значение для выборочной средней составило $\bar{x} = 1.19$ года. Поскольку это значение находится в области отклонения, вы можете сделать вывод, что наблюдаемый результат вряд ли будет объяснен случайным изменением: это, скорее всего, связано с увеличением среднего значения генеральной совокупности. Это говорит о том, что новая схема чтения дает лучшие результаты, чем старая.

В этом примере решение принимается с учетом значения выборочного среднего. Значение выборочного среднего $\bar{X}$ называется **тестовой статистикой** для этой проверки гипотезы. Область отклонения была определена таким образом, чтобы вероятность попадания в нее тестовой статистики, если H_0 верна, составляла не более 0,05 или 5%.

Эта вероятность называется **уровнем значимости** теста. Она дает вероятность отклонения H_0, когда она на самом деле верна. В этом примере она дает вероятность сделать вывод, что новая схема чтения лучше, даже если это не так. Вам может показаться, что это слишком высокий риск ошибиться, и вместо этого решите использовать уровень значимости, скажем, 0,01 или 1%.

Пример 6.3.1.

Найдите область отклонения для теста с уровнем значимости 1% на примере схемы чтения для детей.

Решение:

Теперь Φ(z)=0.99, что дает z=2.326. Подставляя в выражение: $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$

Получим: $2.326 = \frac{c-1.14}{\sqrt{\frac{0.16^2}{40}}}$. Отсюда: c=1.20, с точностью до трех значащих цифр. Область отклонения:

$\bar{X} \geq 1.20$

Наблюдаемое значение 1.19 лет больше не находится в области отклонения, и поэтому H_0 не отклоняется на уровне значимости 1%.

Вы можете почувствовать неудовлетворенность от того, что результат проверки гипотезы должен зависеть от выбранного уровня значимости. Этот момент более подробно обсуждается в разделе 8.

При двухсторонней проверке область отклонения состоит из двух частей, потому что как высокие, так и низкие значения $\bar{X}$ маловероятны, если нулевая гипотеза верна. Следующий пример иллюстрирует эту ситуацию.

Пример 6.3.2.

В прошлом машина производила канат, который имеет разрывную нагрузку, которая распределена по нормальному закону распределения со средним значением 1000 Н и стандартным отклонением 21 Н. Был введен новый процесс. Чтобы проверить, изменилась ли средняя разрывная нагрузка, отбирают выборку из 50 кусков веревки, измеряют разрывную нагрузку для каждого куска и рассчитывают среднее значение.

(а) Сформулируйте подходящие нулевые и альтернативные гипотезы для проверки, изменилась ли разрывная нагрузка.

(б) Взяв среднее значение выборки $\bar{X}$ в качестве статистики теста, найдите область отклонения проверки гипотезы с уровнем значимости 5%.

(в) Примерное среднее значение для 50 кусков веревки было 1003 Н. Какой вывод вы сделаете?

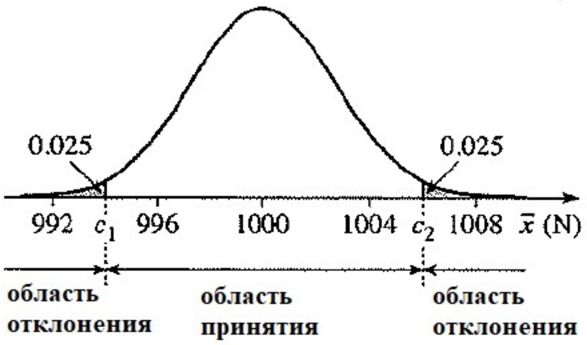
Решение:

(а) Нулевая гипотеза показывает значение, которое должна принимать средняя разрывная нагрузка, то есть $H_0: \mu = 1000$.

Альтернативная гипотеза утверждает, что μ мог измениться, то есть:

$H_1: \mu \neq 1000$

(б) Если H_0 верна, то выборочное среднее распределена по нормальному закону распределения как $\bar{X} \sim N\left(1000; \frac{21^2}{50}\right)$. На следующем рисунке показано распределение $\bar{X}$ с областями принятия и отклонения, с двумя критическими значениями обозначенными как c_1 и c_2.



Чтобы найти верхнее критическое значение c_2, используем $\Phi(z)=0.975$, поскольку вероятность 0,05 делится поровну между двумя «хвостами» распределения. Требуемое значение z равна 1.960.

Подставляя в выражение: $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$, получим: $1.960 = \frac{c_2 - 1000}{\sqrt{\frac{21^2}{50}}}$.

Отсюда: $c_2=1006$, с точностью до целого числа. По симметрии, $c_2=994$.

Таким образом область отклонения: $\bar{X} \leq 994$ и $\bar{X} \geq 1006$

(в) Наблюдаемое среднее значение выборки $\bar{x} = 1003$ не находится в области отклонения. Нет достаточных доказательств того, что среднее значение изменилось, и можно сделать вывод, что новый процесс является удовлетворительным.

Вот краткое изложение терминов и шагов проведения проверки гипотезы.

***Статистика теста* рассчитывается по выборке. Его значение используется, чтобы решить, следует ли отклонить нулевую гипотезу H_0.**

***Область отклонения* (или *критическая область*) - это такие значения тестовой статистики, для которой нулевая гипотеза H_0 отклоняется.**

***Область принятия* - это такие значения тестовой статистики, для которой нулевая гипотеза H_0 принимается (не отклоняется).**

Граничное значение (я) области отклонения называется (являются) критическим значением (ями).

***Уровень значимости* теста дает вероятность того, что статистика теста попадет в область отклонения, когда H_0 верна.**

Если отклоняется нулевая гипотеза H_0, то автоматически принимается альтернативная гипотеза H_1.

Шаги проведения проверки гипотезы:

Шаг 1. Сформулировать нулевую и альтернативные гипотезы.

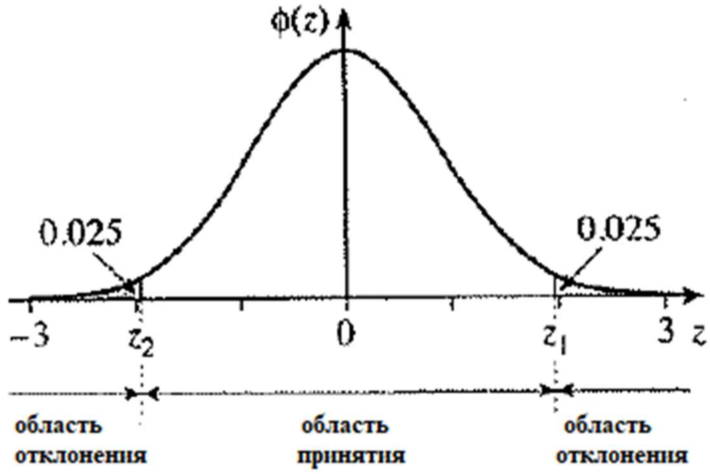
Шаг 2. Определиться с уровнем значимости.

Шаг 3. Определить критическое значение (я).

Шаг 4. Вычислите статистику теста (тестовую статистику).
Шаг 5. Определите, находится ли значение статистики теста в области отклонения или в области принятия.
Шаг 6. Сформулируйте заключение в зависимости от значения статистики теста.

6.4. Стандартизация статистики теста (тестовой статистики)

В предыдущем упражнении область отклонения для каждого пункта упражнения была разной, и вам нужно было найти ее, прежде чем получить результат проверки гипотезы. Возможно, вы заметили, что расчет может быть сокращен путем стандартизации значения $\bar{X}$ использованием:
$Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$, и принимая z в качестве тестовой статистики. Для этой тестовой статистики при заданном уровне значимости, вне зависимости односторонняя или двухсторонняя проверка, область отклонения для Z будет одинаковой. Например, следующий рисунок показывает область отклонения Z для двухсторонней проверке с уровнем значимости 5%.



Верхнее критическое значение получается из $\Phi(z_1) = 0.975$, то есть $z_1=1.960$ и, по симметрии, нижнее критическое значение $z_2=-1.960$. Таким образом, область отклонения $Z \geq 1.960$ и $Z \leq -1.960$, что можно записать более компактно, как $|Z| \geq 1.960$. Следующие примеры иллюстрируют это.

Пример 6.4.1.

Тест умственных способностей был составлен таким образом, что результат взрослого человека Великобритании распределен по нормальному закону распределения со средним значением 100 и стандартным отклонением 15. Врач желает проверить, отличаются ли результаты пациентов, страдающие конкретным заболеванием от среднего значения взрослых жителей Великобритании. Врач выбирает случайную выборку из 10 пациентов. Результаты тестирования этих пациентов:
119 131 95 107 125 90 123 89 103 103.
Выполните проверку гипотезы с уровнем значимости 5%, что результаты пациентов, страдающих этим конкретным заболеванием отличаются от средних результатов по стране при сдаче этого теста умственных способностей.

Решение:

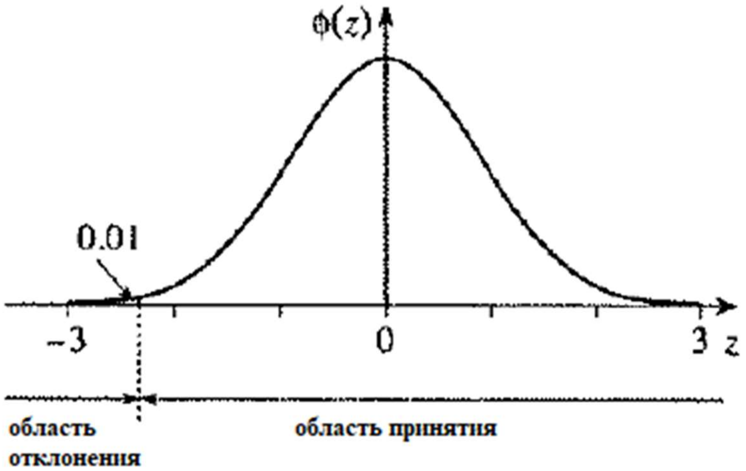
Нулевая и альтернативная гипотеза соответственно $H_0: \mu = 100$ и
$H_1: \mu \neq 100$. Это двухсторонняя проверка с уровнем значимости 5%. Как показано выше, область отклонения для тестовой статистики Z, с этим уровнем значимости: $|Z| \geq 1.960$. Если H_0 верна, то $\bar{X} \sim N\left(100; \frac{15^2}{10}\right)$.
Для этой выборки,
$\bar{x} = \frac{1}{10}(119 + 131 + 95 + 107 + 125 + 90 + 123 + 89 + 103 + 103) = 108.5$
При $\bar{x} = 108.5$, $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{(108.5-100)}{\sqrt{\frac{15^2}{10}}} = 1.792$, с точностью до трех цифр после запятой.
Наблюдаемое значение Z не находится в области отклонения, поэтому H_0 не отклоняется. На уровне значимости 5% недостаточно данных, чтобы предположить, что больные этим заболеванием отличаются от общей генеральной совокупности по результатам теста умственных способностей.

Пример 6.4.2.

Производитель заявляет, что срок службы его лампочек составляет 1500 часов, а стандартное отклонение - 30 часов. Владелец магазина подозревает, что лампочки не служат так долго, как утверждает производитель, потому что он получил ряд жалоб от клиентов. Он проверяет случайную выборку из шести лампочек и обнаруживает, что срок их службы составляет 1472, 1486, 1401, 1350, 1511, 1591 часов. Можно ли утверждать с уровнем значимости 1%, что срок службы лампочек короче, чем заявляет производитель?

Решение:

Нулевая и альтернативная гипотеза: $H_0: \mu = 1500$ и $H_1: \mu < 1500$ соответственно. Это односторонняя проверка на уменьшение с уровнем значимости 1%. На рисунке показана область отклонения для Z.



Критическое значение которого определяется из: Φ(z)=0.01. Напомним, что: $\Phi(z) = 1 - \Phi(z) = 1 - 0.01 = 0.99$. Тогда из таблицы нормального распределения: $-z = 2.326$ или $z = -2.326$. Таким образом, область отклонения: $Z \leq -2.326$. Если H_0 верна, то $\bar{X} \sim N\left(1500; \frac{30^2}{6}\right)$.

Для этой выборки:

$\bar{x} = \frac{1}{6}(1472 + 1486 + 1401 + 1350 + 1511 + 1591) = 1468.5$

При $\bar{x} = 1468.5$, $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{1468.5-150}{\sqrt{\frac{30^2}{6}}} = -2.572$, с точностью до трех цифр после запятой.

Наблюдаемое значение Z находится в области отклонения. На уровне значимости 1% имеются доказательства того, что лампочки производителя не служат так долго, как утверждает производитель. Можно обобщить этот метод следующим образом.

Тестовая статистика Z может использоваться для проверки гипотезы о среднем значении генеральной совокупности, $H_0: \mu = \mu_0$, для выборок, взятых из нормального распределения с известной дисперсией σ^2 . Для выборки размера n значение Z определяется как: $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

Область отклонения для Z зависит от H1 и используемого уровня значимости. Критические значения для некоторых обычно используемых областей отклонения приведены ниже:

Уровень значимости	*Двухсторонняя проверка* $H_1: \mu \neq \mu_0$	*Односторонняя проверка* $H_1: \mu > \mu_0$	*Односторонняя проверка* $H_1: \mu < \mu_0$
10%	± 1.645	*1.282*	-1.282
5%	± 1.960	*1.645*	-1.645
2%	± 2.326	*2.054*	-2.054
1%	± 2.576	*2.326*	-2.326

6.5. Выборки больших размеров

В рассмотренных выше примерах мы предполагали, что выборки были взяты из нормального распределения и что дисперсия этого распределения известна. На практике во многих случаях вы не можете быть уверены, что распределение генеральной совокупности является нормальным, и вы можете иметь или не иметь точную информацию о его дисперсии. Однако метод проверки гипотез, который был описан выше, все еще может применяться при условии, что размер выборки достаточно большой. Сначала рассмотрим распределение, из которого берется выборка. Хотя распределение может и не быть нормальным распределением, центральная предельная теорема говорит, что среднее значение выборки, $\bar{X}$ распределена по нормальному закону распределения при достаточно большом размере выборки. Далее рассмотрим дисперсию генеральной совокупности. Несмещенная оценка этого может быть рассчитана по выборке с использованием следующей формулы, показанной ранее: $s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)$. Ели объем выборки достаточной большой, то эта оценка достаточно точна, чтобы ее можно использовать вместо σ^2.

Тестовую статистику Z можно использовать для проверки гипотезы о среднем значении генеральной совокупности H_0: $\mu = \mu_0$ для <u>больших</u> выборок из <u>любой</u> совокупности.
Для выборки размера n значение Z определяется как: $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

Если дисперсия генеральной совокупности σ^2 неизвестна, ее можно заменить оценкой s^2.

Термин «большой» не очень точен. Эмпирическое правило будет означать, что «большой» означает размер выборки 30 и более.

Пример 6.5.1.

Новая хирургическая методика была разработана в попытке сократить время, которое пациенты должны проводить в больнице после определенной операции. В прошлом среднее время пребывания в больнице составляло 5,3 дня. Для первых 50 пациентов, на которых была опробована новая методика, среднее время, проведенное в больнице, составило 5,0 дней с оценочной дисперсией $0{,}4^2$ дней^2. Имеются ли доказательства на уровне значимости 2%, что новая методика сократила время, проведенное в больнице?

Решение:

Нулевая и альтернативная гипотеза: $H_0: \mu = 5.3$ и $H_1: \mu < 5.3$ соответственно. Распределение генеральной совокупности неизвестна, но так как размер выборки достаточно большой (>30), то можно использовать тестовую статистику Z. Это односторонняя проверка на уменьшение с уровнем значимости 2%. Из таблицы, приведенной выше, в этом случае область отклонения: $Z \leq -2.054$.

$z = \frac{\bar{x}-\mu}{\sqrt{\frac{s^2}{n}}} = \frac{5.0-5.3}{\sqrt{\frac{0.4^2}{50}}} = -5.303$, с точностью до трех цифр после запятой.

Рассчитанное значение z находится в области отклонения, поэтому H_0 отклоняется: то есть на уровне значимости 2% имеются доказательства того, что время в больнице было сокращено с помощью новой методики.

Пример 6.5.2.

Инспектору изделий с производственной линии требуется в среднем 21,75 секунды для проверки каждого изделия. После установки новой системы освещения время, t секунд, для проверки каждого из 50 случайно выбранных изделий из производственной линии суммируется следующим образом:
$\sum t = 1107$ и $\sum t^2 = 24592.35$

(a) Рассчитайте несмещенную оценку дисперсии времени проверки изделия в новой системе освещения.
(b) Проверьте с уровнем значимости 5%, есть ли доказательства того, что среднее время проверки изделий изменилось с 21,75 секунд.

Решение:

(a) Несмещенная оценка дисперсии генеральной совокупности:

$$s^2 = \frac{1}{n-1}\left(\sum t^2 - \frac{(\sum t)^2}{n}\right) = \frac{1}{49}\left(24592.35 - \frac{1107^2}{50}\right) = 1.701 \ldots$$

(б) Нулевая и альтернативная гипотеза: $H_0: \mu = 21.75$ и $H_1: \mu \neq 5.3$ соответственно. В данном случае неизвестны ни вид распределения, ни дисперсия генеральной совокупности, но так как размер выборки достаточно большой (>30), можно использовать Z в качестве тестовой статистики. Это двухсторонняя проверка на изменение с уровнем значимости 5% и для нее область отклонения (из таблицы выше) тестовой статистики Z: $|Z| \geq 1.960$.

Для данной выборки: $\bar{x} = \frac{1107}{50} = 22.14$.

$z = \frac{\bar{x}-\mu}{\sqrt{\frac{s^2}{n}}} = \frac{22.14-21.75}{\sqrt{\frac{1.701...}{50}}} = 2.114$, с точностью до трех знаков после запятой.

Это значение находится в области отклонения. Таким образом, с уровнем значимости 5% можно утверждать что среднее значение генеральной совокупности изменилось.

6.6. Альтернативный метод проведения проверки гипотезы.

Другой способ проведения проверки гипотезы состоит в том, чтобы рассчитать вероятность того, что статистика теста принимает наблюдаемое значение (или более экстремальное значение), и сравнить эту вероятность с уровнем значимости. Если вероятность меньше уровня значимости, нулевая гипотеза отклоняется. Результат считается «значимым» на данном уровне значимости. Следующий рисунок показывает, что этот метод всегда будет давать тот же результат, что и предыдущий метод.

Рисунок слева показывает область отклонения для Z для односторонней проверки на увеличение с уровнем значимости 5%. Если Z принимает значение в области отклонения, например 1.9, то из этого рисунка видно, что $P(Z \geq 1.9)$ меньше 0.05, что также приведет к отклонению H_0. Если Z принимает значение в области принятия, например 1.2, как показано на рисунке справа, то $P(Z \geq 1.2)$ больше 0.05, и H_0 не будет отклонено. Чтобы проиллюстрировать эту идею, снова посмотрите на Пример 6.5.1. В этом примере тестовая статистика приняла значение -5.30 (с точностью до 3 значащих цифр). Интерес представляют низкие значения Z, поэтому «более экстремальное значение», в данном случае значения меньше -5.30. Вероятность того, что Z примет это значение или более экстремальное значение, равна $P(Z \leq -5.30) = 1 - \Phi(5.30) = 0$.

Поскольку эта вероятность менее 2%, результат является значимым на уровне значимости 2%, и поэтому нулевая гипотеза отклоняется. Как и следовало ожидать, это то же самое заключение которое было сделано ранее. Однако этот способ записи результата дает больше информации: он показывает, что результат был значимым не только на уровне 2%, но и на гораздо более низком уровне значимости.

Предоставление вероятности (иногда называемой p-значением) вместо использования критических значений требует немного больше действий. Однако если использовать компьютеры (компьютерные программы), то легко определить вероятность, и большинство статистических программ дают результат проверки гипотезы в такой форме.

Следующий пример иллюстрирует этот подход в двухсторонней проверке.

Пример 6.6.1.

Станок предназначен для изготовления металлических стержней длиной 2 см со стандартным отклонением 0,02 см. Можно сказать, что длины стержней распределены по нормальному закону распределения. Машина перемещается в новое положение на заводе, и для того, чтобы проверить, изменилась ли настройка для средней длины, измеряются длины первых десяти стержней. Стандартное отклонение можно считать неизменным. Если эти длины в сантиметрах соответствуют приведенным ниже, проверьте на уровне значимости 5%, изменилась ли настройка или нет.

2.04 1.97 1.99 2.03 2.04 2.10 2.01 1.98 1.97 2.02

Решение:

Это двухсторонняя проверка с нулевой гипотезой, предполагающей, что среднее значение не изменяется.

Нулевая и альтернативная гипотеза: $H_0: \mu = 2$ и $H_1: \mu \neq 2$ соответственно.

Выборочное среднее $= \frac{1}{10}(2.04 + 1.97 + 1.99 + 2.03 + 2.04 + 2.10 +$
$+2.01 + 1.98 + 1.97 + 2.02) = 2.015$

Так как генеральная совокупность распределена по нормальному закону распределения то и выборочное среднее, $\bar{X}$, тоже распределена по нормальному закону распределения. Если H_0 верна, то $\bar{X} \sim N\left(2; \frac{0.02^2}{10}\right)$. Таким образом:

$$P(\bar{X} \geq 2.015) = P\left(Z \geq \frac{2.015 - 2}{\frac{0.02}{\sqrt{10}}}\right) = P(Z \geq 2.372) = 1 - \Phi(2.372) =$$

$$= 1 - 0.9912 = 0.0088 = 0.88\%$$

Поскольку это двухсторонняя проверка, эту вероятность следует сравнивать с половиной уровня значимости, то есть 2,5%, как показано на рисунке. Поскольку 0,88% составляет менее 2,5%, результат является значимым на уровне значимости 5%. Можно предположить, что средняя длина стержней, произведенных машиной, была изменена при перемещении станка.

Кроме того, вы можете рассчитать значение p для теста. Это вероятность того, что статистика теста отклоняется от среднего значения генеральной совокупности на наблюдаемое количество (или больше) в любом направлении. Таким образом, в этом случае значение p составляет $2 \cdot 0.88\% = 1.76\%$. Затем значение p сравнивают с уровнем значимости. Так как 1,76% <5%, результат значим на уровне 5%, как и раньше.

Любой метод проведения проверки гипотезы, использующий критические значения или использующий вероятности, является удовлетворительным, и обычно вам следует использовать метод, который для вас проще.

Задачи.

№1. Металлические стойки, используемые в здании, имеют среднюю длину 2,855 м. Длины стоек распределены по нормальному закону распределения со стандартным отклонением 0,0352 м. Партия из 15 стоек отправляется на строительную площадку и измеряется их длина. Средняя длина этой выборки составляет 2,841 м. Тест должен быть проведен на уровне значимости 5%, чтобы решить, является ли партия из указанной генеральной совокупности.

а) Сформулируйте вашу гипотезу и выразите область отклонения через Z.

б) Изложите заключение проверки гипотезы.

Ответ №1: а) $H_0: \mu = 2{,}855$, $H_1: \mu \neq 2{,}855$, $|Z| \geq 1{,}96$

б) $z = -1{,}540$, поэтому принимаем нулевую гипотезу $H_0: \mu = 2{,}855$, партия соответствует указанной генеральной совокупности.

№2. Случайная величина X распределена по нормальному закону распределения с неизвестным средним значением, но с известной дисперсией 12,4. Выборочное среднее значение случайной выборки из 10-ти наблюдений над случайной величиной X обозначено как $\bar{X}$. Областью принятия теста нулевой гипотезы $\mu = 25$ является $\bar{X} > 22{,}41$.

а) Сформулируйте альтернативную гипотезу.

б) Найдите уровень значимости теста.

в) Если бы гипотетическое значение μ было больше 25, был бы уровень значимости, соответствующий той же области принятия, больше или меньше, чем тот, который найден в пункте б)? Объясните свой ответ.

Ответ №2: а) $\mu < 25$ б) $z = -2{,}326$, 1%

в) $z < -2{,}326$, уровень значимости $< 1\%$, так как $z < -2{,}326$.

№3. Можно предположить, что масса содержимого фирменного греческого йогурта распределена по нормальному закону распределения со стандартным отклонением 2,58 грамма. Приобретается новая машина для наполнения коробочек йогурта, и случайная выборка из 20-ти коробочек йогурта, заполненных этой машиной, используется для проверки нулевой гипотезы о том, что средняя масса содержимого йогурта в коробочке составляет μ_0 граммов. Используя стандартное отклонение 2,58 грамма, находим область отклонения для теста, которая равна $\bar{X} \leq 208{,}81$ или $\bar{X} \geq 211{,}19$.

а) Найдите значение μ_0.

б) Найдите уровень значимости этого теста.

Ответ №3: а) $\mu_0 = 210$ б) 3,9%

№4. Ниша, страдающая диабетом, должна следить за уровнем глюкозы в крови, который меняется в течение дня. Результаты выборки из 75 показаний уровня глюкозы, x, взятых случайным образом в течение недели, суммированы следующим образом: $\sum x = 511{,}5$ и $\sum x^2 = 4027{,}89$.

а) Предполагая нормальное распределение, проверьте с уровнем значимости 5%, больше ли средний уровень глюкозы в крови Ниши, μ, чем 6,0.

б) Найдите область значений μ_0, для которых было бы принято, что $\mu > \mu_0$ с уровнем значимости 10%.

в) Укажите причину, будет ли заключение проверки пункта б) достоверным.

(i) если бы уже нельзя было предположить, что уровень глюкозы в крови имеет нормальное распределение.

(ii) если бы все 75 показаний были сделаны в выходные дни.

Ответ №4: а) $z = 2{,}630 > 1{,}96$, поэтому принимаем, что $\mu > 6{,}0$, то есть средний уровень глюкозы в крови Нишы больше чем 6,0

б) $\mu_0 < 6{,}420$

в) (i) допустимо; размер выборки достаточно велик для центральной предельной теоремы и замены σ на величину, оцененную по выборке.

(ii) снятие показаний только в выходные дни может быть вызвать смещение, другим стилем жизни в выходные, поэтому результаты не являются достоверными.

№5. Общий вес урожая, x кг, каждого из 64 бобовых растений измерены садоводом, и результаты суммированы следующим образом: $\sum x = 303{,}4$ и $\sum x^2 = 1615{,}96$. Найдите несмещенные оценки среднего значения и дисперсии генеральной совокупности (урожая всех бобовых растений этого вида). Садовод хочет проверить гипотезу о том, что средняя масса урожая на одно растение составляет 5 кг,

против альтернативной гипотезы о том, что средняя масса урожая составляет менее 5 кг. Проведите проверку с уровнем значимости 10%. Найдите наименьший уровень значимости, при котором проверка приведет к отклонению нулевой гипотезы.
Ответ №5: $\bar{x} = 4{,}741$, $s^2 = 2{,}820$, $z = -1{,}236 > -1{,}282$, поэтому принимаем нулевую гипотезу H_0, то есть средняя масса урожая на одно растение составляет 5 кг. Наименьший уровень значимости равен 10,8%

№6. Тренер по легкой атлетике использует легкоатлетический стадион, где он тренирует атлетов для бега на дистанцию по кругу. Отмечается время, затраченное новым спортсменом на прохождение одного круга в каждом из большого количества случаев. Среднее значение этих времен составляет 100 секунд, а стандартное отклонение-3 секунды. Распределение времени можно считать нормальным. Три месяца спустя, после того как спортсмен тренируется ежедневно, его время на прохождение одного круга в десяти случаях выглядит следующим образом.
96,8 101,2 98,2 99,6 98,0 95,6 98,0 100,0 95,2 97,4
Убедитесь, что эти значения показывают с уровнем значимости 5%, что у спортсмена улучшилось время прохождения одного круга.
Расчеты, основанные на большом количестве дальнейших испытаний, показывают, что на данном этапе среднее значение времени, затраченного спортсменом на прохождение одного круга, составляет 98 секунд, а стандартное отклонение-3 секунды. Каково было бы максимальное общее время прохождения 10-ти кругов в дальнейшем, которое обеспечило бы доказательства дальнейшего улучшения на уровне значимости 5%?
Ответ №6: $\bar{x} = 98$, $H_0\colon \mu = 100$, $H_1\colon \mu < 100$, $z = -2{,}108 < -1{,}645$, поэтому принимаем альтернативную гипотезу $\mu < 100$, Максимальное время прохождения 964,4 секунды показало бы дальнейшее улучшение результата спортсмена ($\mu < 98$).

№7. Станция скорой помощи обслуживает район, который включает в себя более 10000 домов. Было решено, что если среднее расстояние домов от станции скорой помощи будет больше 10 миль, то потребуется новая станция скорой помощи. Было измерено расстояние в x милях от станции до каждого из случайной выборки из 200 домов, результаты были суммированы в виде: $\sum x = 2092{,}0$ и $\sum x^2 = 24994{,}5$.
а) Вычислите, с точностью до 4-х значащих цифр, несмещенные оценки:
(i) среднего расстояния от всех домов (генеральной совокупности) до станции, μ миль.
(ii) дисперсии расстояния от всех домов (генеральной совокупности) до станции.
Объясните, что вы понимаете под «несмещенной оценкой».
б) Проведена проверка нулевой гипотезы $\mu = 10$ и альтернативной гипотезы $\mu > 10$ с уровнем значимости $\alpha\%$ на этой случайной выборке из 200 домов. Получили область отклонения нулевой гипотезы: $\bar{X} \geq 10{,}65$, где $\bar{X}$ обозначает выборочное среднее этой случайной выборки.
(i) Вычислите значение α.
(ii) Сформулируйте заключение теста с использованием выборочных данных.
в) Предположим, что нельзя предположить, что расстояния распределены нормально. Укажите, будут ли ответы в пункте а) и в пункте б) по-прежнему верны.
Ответ №7: а) (i) 10,46 (ii) 15,64
б) (i) приблизительно 1,0
(ii) так как $10{,}46 < 10{,}65$ то принимается нулевая гипотеза H_0. Среднее расстояние от всех домов до станции не больше 10-ти миль
в) Ответы части (a) требуют только случайности выборки, поэтому они все еще действительны. Часть (b) имеет достаточно большую выборку для центральной предельной теоремы, поэтому результаты тоже справедливы.