Глоссарий

Генеральная совокупность (ГС) — множество всех исследуемых объектов.

Практически всегда ГС — это набор чисел: значения интересующей аналитика переменной, измеренной для всех объектов, которые он изучает. Например, количество позиций в чеках, пробитых всеми кассами торгового центра за всё время его работы, или суммарная стоимость всех позиций для каждого чека.

Параметр генеральной совокупности — термин с двумя значениями:

- Функция, которую можно рассчитать на множестве всех чисел из ГС. Например, среднее и стандартное отклонение всех чисел, входящих в ГС, — параметры ГС.
- Параметр теоретического распределения, которому подчиняется ГС.

Для любой конкретной ГС все её параметры — это числа.

Часто к параметрам ГС применяют термин «истинный»: например, среднее всех чисел из ГС называют истинным средним.

Параметры ГС принято обозначать греческими буквами: например, истинное среднее — μ, истинное стандартное отклонение — σ.

Выборка — доступный набор данных из генеральной совокупности.

Размер выборки (обычно обозначается n) — количество элементов в ней.

Случайная выборка — выборка, вероятность попасть в которую одинакова для всех объектов из генеральной совокупности. Страты — непересекающиеся группы объектов генеральной совокупности, различающиеся по значениям исследуемой переменной и, чаще всего, по размеру. Стратифицированная выборка — выборка, составленная путём объединения случайных выборок, взятых из всех страт генеральной совокупности пропорционально их размеру. Например, если генеральная совокупность разбита на страты А и Б и объектов в страте А в три раза больше, чем в страте Б, а выборку есть возможность взять размера 100, то из страты А берётся случайная выборка размера 75, а из страты Б — размера 25. Эти случайные выборки объединяются в итоговую выборку.

Репрезентативность выборки — способность выборки отражать свойства генеральной совокупности. Репрезентативность выборки зависит от способа её получения, но не от её размера.

Статистика — функция, которую можно рассчитать только на выборочных данных. Например, **выборочное среднее** — среднее всех чисел из выборки. **Статистика** — это случайная величина: на разных выборках она принимает разные значения. Раз это случайная величина, у неё есть распределение, математическое ожидание и дисперсия.

Проверка гипотез

Термин «выборочная» добавляется к называнию функции, чтобы показать, что это статистика. Например, выборочное стандартное отклонение — стандартное отклонение, рассчитанное на данных выборки.

Статистики принято обозначать латинскими буквами.

Выборочное распределение статистики — распределение её значений на всех возможных выборках фиксированного размера, которые можно взять из ГС, или теоретическое распределение, близко его описывающее.

Оценка — процесс нахождения по выборке числа, близкого к интересующему параметру ГС, или интервала, с высокой вероятностью содержащего параметр ГС. Результат этого процесса — число или интервал — тоже называют оценкой. Статистику, которую для этого используют, тоже называют оценкой.

Точечная оценка — оценка параметра ГС числом, то есть точкой на числовой оси. Для точечных оценок вместо латинских букв иногда используют греческие буквы, которыми обозначается оцениваемый параметр, с «крышечкой»: например, $\hat{\sigma}$ — это оценка σ.

Интервальная оценка — оценка параметра ГС интервалом: промежутком на числовой оси.

Смещение оценки — разница между параметром ГС и математическим ожиданием его оценки, то есть статистики, которую используют для получения близкого к этому параметру числа.

Несмещённая оценка — статистика, математическое ожидание которой равно оцениваемому ей параметру ГС. Например, доказано, что выборочное среднее — несмещённая оценка истинного среднего.

Поправка Бесселя — изменение формулы расчёта дисперсии (стандартного отклонения) набора данных: замена в знаменателе n на n-1. Выборочная дисперсия (стандартное отклонение), рассчитанная с поправкой Бесселя — несмещённая оценка истинной дисперсии (стандартного отклонения).

Центральная предельная теорема (ЦПТ) — доказанное утверждение о том, что:

- Выборочное распределение выборочного среднего будет стремиться к нормальному по мере роста размера выборки n.
- У этого нормального распределения будут следующие параметры: математическое ожидание — истинное среднее ГС (обычно обозначается \mu), стандартное отклонение — истинное стандартное отклонение, делённое на корень из размера выборки ($\sigma/\sqrt{n}$).

Стандартная ошибка среднего (SE) — стандартное отклонение этого нормального распределения: $SE = \sigma/\sqrt{n}$. По выборке SE найти нельзя: в формуле есть σ — параметр ГС. Иначе говоря, SE — это не статистика

Оценённая стандартная ошибка среднего (ESE) — стандартная ошибка среднего, оценённая по выборке: $ESE = s/\sqrt{n}$, где s — несмещённая оценка σ, рассчитанная с поправкой Бесселя. Эту формулу можно записать и так: $\hat{\sigma}/\sqrt{n}$. ESE — это статистика.

Формулирование статистических гипотез и логика их проверки

Статистическая гипотеза — утверждение о параметре генеральной совокупности.

Поскольку параметр ГС — это число, то часто используемые гипотезы имеют следующий вид:

- параметр ГС равен некоторому числу,
- параметр ГС не равен некоторому числу,
- параметр ГС больше некоторого числа,
- параметр ГС меньше некоторого числа.

Нулевая и альтернативная гипотезы (H_0 и H_1) — взаимоисключающие статистические гипотезы, формулируемые для конкретной исследовательской задачи. Для количественной проверки исследовательскую задачу нужно перевести на язык статистики — например, сформулировать в виде такого набора статистических гипотез.

По умолчанию считается, что нулевая гипотеза верна, если только данные не позволят отвергнуть её в пользу альтернативной гипотезы. Подтвердить или доказать гипотезу статистическими, да и в целом научными, методами нельзя. Только отвергнуть или не отвергнуть.

H_0 чаще всего формулируется в виде «параметр ГС равен некоторому числу», так как это позволяет — принимая H_0 верной — построить одно выборочное распределение для статистики, которой оценивается этот параметр. С одним распределением работать проще, чем со множеством распределений.

Для формулировки H_1 используется один из оставшихся вариантов:

- параметр ГС не равен числу из H_0 — **двусторонняя** H_1 ,
- параметр ГС больше числа из H_0 — **правосторонняя** H_1 ,
- параметр ГС меньше числа из H_0 — **левосторонняя** H_1 .

Помимо сформулированных гипотез, у аналитика в распоряжении есть выборка из ГС, то есть данные, на основании которых нужно или отвергнуть H_0 в пользу H_1 или нет.

Проверка гипотез

На выборке можно рассчитать статистику — оценку параметра, о котором сформулированы гипотезы. Если значение этой статистики на имеющейся в наличии выборке оказывается слишком далеко от предположенного в H_0 числа, то H_0 следует отвергнуть в пользу H_1. Если нет — то отвергать не следует.
Для принятия этого решения нужно выбрать уровень статистической значимости (его принято обозначать α). Он задаёт порог: насколько маловероятным должно быть это отклонение, чтобы уже не считать его случайным, а считать значимым.

> 👉 Значимость в статистике — это антоним случайности. Статистически значимое отличие наблюдения от предположения — не отличие, имеющее какую-то важность, а просто не считающееся случайным.

Наблюдаемое значение — посчитанная на выборке статистика — может отклониться от значения, предположенного в H_0, и случайным образом. Но если разница слишком велика, то есть вероятность получить такое или ещё более существенное отклонение меньше α, то H_0 отвергается.

Статистический тест (или **статистический критерий**) — алгоритм или правило, по которому H_0 отвергается или не отвергается в пользу H_1.

Типичный алгоритм проведения статистического теста:

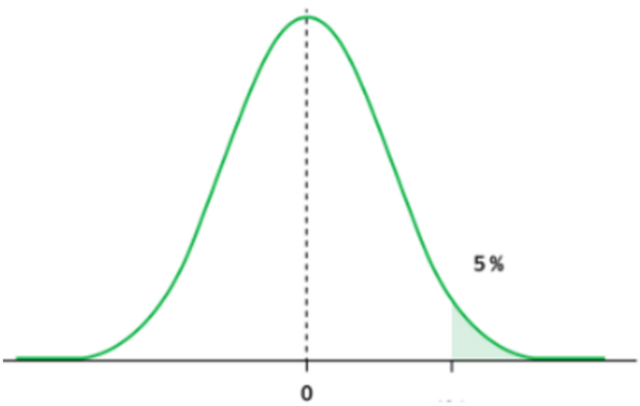
1. Собрать данные — значения исследуемой переменной, то есть получить выборку.
2. Сформулировать H_0 и H_1, исходя из исследовательского вопроса. В рамках спринта мы делали это только для такого параметра ГС, как истинное среднее, для него и будем дальше строить этот алгоритм.
3. Выбрать уровень статистической значимости. Часто используются значения 1% или 5%.
4. Принимая H_0 верной, построить выборочное распределение выборочного среднего, используя ЦПТ. Предположенное в H_0 значение служит его математическим ожиданием, а ESE — стандартным отклонением. На практике строится не нормальное распределение, а близкое к нему распределение Стьюдента — о нём ниже.
5. С помощью построенного распределения найти вероятность получить наблюдаемое значение — посчитанное на выборке среднее — или ещё более далёкое от него. Эта вероятность называется p-value.
6. Что имеется в виду под «ещё более далёким», зависит от выбора H_1: например, если H_1 правосторонняя, более далёкими от предполагаемого в H_0 значениями будут считаться те, что больше наблюдаемого, то есть справа от него.

> 👉 α — это пороговый уровень, и его общепринятые значения в $1\ \%$ или $5\ \%$ — не более чем конвенция, условная договорённость. Если вы получили p-value, близкий к выбранному уровню значимости, к результату теста нужно относиться с осторожностью. В этой ситуации можно провести дополнительное исследование, не полагаясь только на результаты этого теста. (См. «Двухвыборочный t-тест для независимых выборок» в разделе «Примеры кода».)

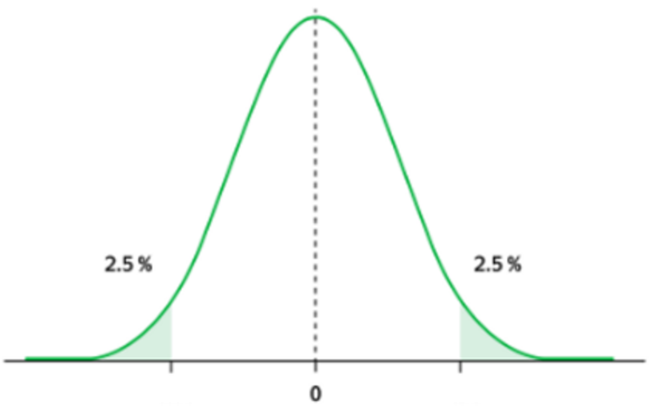
7. Сравнить полученное p-value и выбранный уровень статистической значимости. Если p-value $\leq \alpha$, можно сделать вывод, что наблюдаемое значение слишком далеко от предполагаемого в нулевой гипотезе, и отвергнуть её. Если же p-value превышает α, то нулевая гипотеза не отвергается. Это и есть результат статистического теста: отвержение или неотвержение H_0 в пользу H_1.

Дополнительное определение к пункту 5: промежутки на числовой оси, при попадании в которые выборочного среднего H_0 отвергается, называются критическими интервалами. Вместо расчёта p-value и сравнения его с α можно по α построить критические интервалы и по попаданию или не попаданию в них выборочного среднего сделать вывод об отвержении или неотвержении H_0 соответственно. Результат будет всегда одинаковым, p-value просто удобнее считать и интерпретировать, поэтому обычно считают его.

Вот как выглядят критические интервалы для правосторонней и двусторонней H_1 с уровнем значимости 5%, отмеченные на выборочном распределении выборочного среднего (H_0 здесь: «истинное среднее равно нулю»):



Односторонний тест



Двусторонний тест

Важно отметить: стандартное отклонение выборочного распределения — ESE — оценивается по выборке и зависит от того, насколько большой разброс имеет генеральная совокупность.

Результат проверки гипотезы зависит от дисперсии ГС прямым образом: одно и то же отклонение, посчитанной на выборке статистики от предположенного в H_0 значения, может быть значимым, если дисперсия ГС маленькая, или нет, если дисперсия ГС большая.

Например, если до рекламной кампании большинство клиентов обычно заказывало товаров на 2–3 тысячи рублей со средним в 2.5 тысячи, а после стали заказывать в среднем на 4 тысячи, то отличие, скорее всего, будет значимым: видимо, реклама сработала. Если же до рекламной кампании разброс был больше, и клиенты заказывали и на 50 рублей, и на 5 тысяч, нулевая гипотеза о том, что продажи не выросли и средний заказ остался равен 2.5 тысячам рублей, может оказаться не отвергнутой с тем же выборочным средним в 4 тысячи. При большом разбросе стоимости заказов выборочное среднее могло оказаться равно 4 тысячам случайно. Эта логика показана на данных в комментарии после первого примера в разделе «Примеры кода».

Распределение Стьюдента и t-тесты

Распределение Стьюдента

Если размер выборки меньше 30, выборочное распределение выборочного среднего имеет немного большую дисперсию, чем нормальное распределение. Такое распределение называется **распределением Стьюдента**. Его обычно и используют для проверки гипотез о среднем, хотя при $n > 30$ оно практически неотличимо от нормального распределения, — просто чтобы всегда использовать одно и то же.

Статистические тесты, проверяющие гипотезу о среднем и использующие распределение Стьюдента, называются **t-тестами**.

Параметр распределения Стьюдента, определяющий близость его формы к нормальному распределению, называется **количеством степеней свободы**. При проведении t-теста с одной выборкой или с двумя зависимыми выборками он равен $n - 1$, то есть размеру выборки (или обеих зависимых выборок), уменьшенному на единицу.

Если выборок две, они независимы, и их дисперсии полагаются равными (см. подраздел «Двухвыборочный t-тест для независимых выборок») — количество степеней свободы распределения Стьюдента равно сумме их размеров, каждый из которых уменьшен на единицу: $n_1 + n_2 - 2$. Если выборки независимы и их дисперсии полагаются неравными, количество степеней свободы рассчитывается сложнее. О зависимости и независимости выборок и том, почему для двух выборок строится всего одно распределение — следующий раздел.

Типы t-тестов и их проведение

Типы t-тестов

В этом спринте мы рассмотрели часто используемые t-тесты. Они предполагают верными либо нулевую гипотезу о равенстве истинного среднего ГС некоторому выбранному аналитиком числу, либо нулевую гипотезу о равенстве истинных средних двух генеральных совокупностей друг другу. В первом случае используется одна выборка — такой тест также называют **одновыборочным**. Во втором — две выборки, такой тест называют **двухвыборочным**.

Одновыборочный t-тест проводится по пошаговому алгоритму, приведённому выше.

Двухвыборочный t-тест можно проводить для независимых или для зависимых выборок.

Двухвыборочный t-тест для независимых выборок

Если нужно провести такой тест — то есть сравнить средние двух ГС, из каждой из которых независимо взяты две выборки, — алгоритм следующий. Нулевая гипотеза формулируется как «истинные средние этих двух ГС равны», и проводится одновыборочный t-тест о равенстве нулю *разности* истинных средних этих двух генеральных совокупностей.

В качестве статистики при этом берётся разница между выборочными средними двух выборок, взятых из этих двух ГС. То есть у двухвыборочного t-теста для независимых выборок «под капотом» — одновыборочный t-тест.

При этом возникает нюанс. Для проведения теста нужно построить выборочное распределение статистики «разница между выборочными средними». Нужно оценить стандартное отклонение этого распределения, и тут есть два варианта:

- Оценить стандартное отклонение, объединив обе выборки (так и происходит по умолчанию) — этот вариант подходит, если дисперсии обеих ГС можно считать равными.
- Оценить стандартное отклонение каждой ГС по отдельности по взятым из них выборкам, и уже стандартное отклонение для статистики посчитать по ним.

Этот выбор аналитик делает, исходя из своего представления об исследуемых генеральных совокупностях: равны ли их дисперсии или нет.

Показано, что t-тест устойчив даже к разнице дисперсий генеральных совокупностей в несколько раз, поэтому использовать расчёт стандартных отклонений по отдельности имеет смысл, только если предполагаемая аналитиком разница дисперсий действительно велика. Особенно с учётом того, что объединённая выборка имеет больший размер, чем каждая из двух изначальных выборок по отдельности, а значит, стандартное отклонение выборочного распределения статистики будет оценено по ней точнее.

Двухвыборочный t-тест для зависимых выборок

Другой случай двухвыборочного t-теста — ситуация, когда выборки зависимы. Это значит, что рассматривается одна и та же генеральная совокупность и некоторая переменная измерена два раза для одних и тех же объектов наблюдения — до и после какого-то изменения. Поэтому выборки обязательно должны быть одного размера и согласованы друг с другом: первое число в каждой выборке — это измерения до и после для первого объекта, второе — для второго и так далее.

Этот тест тоже сводится к одновыборочному: для каждого объекта рассчитывается разница между значениями до и после, то есть получается выборка таких разниц для всех объектов. По этой выборке проводится одновыборочный t-тест о равенстве среднего значения этой разницы нулю.

Вывод о формулировке альтернативной гипотезы

Таким образом, двухвыборочные t-тесты всегда сводятся к одновыборочному t-тесту, для которого строится одно выборочное распределение. Следовательно, все перечисленные t-тесты могут быть односторонними или двусторонними, то есть альтернативная гипотеза может быть сформулирована и как двусторонняя, и как левосторонняя, и как правосторонняя. Подробнее в таблице ниже.

Выбор нужного теста и код на python

Схема выбора нужного t-теста

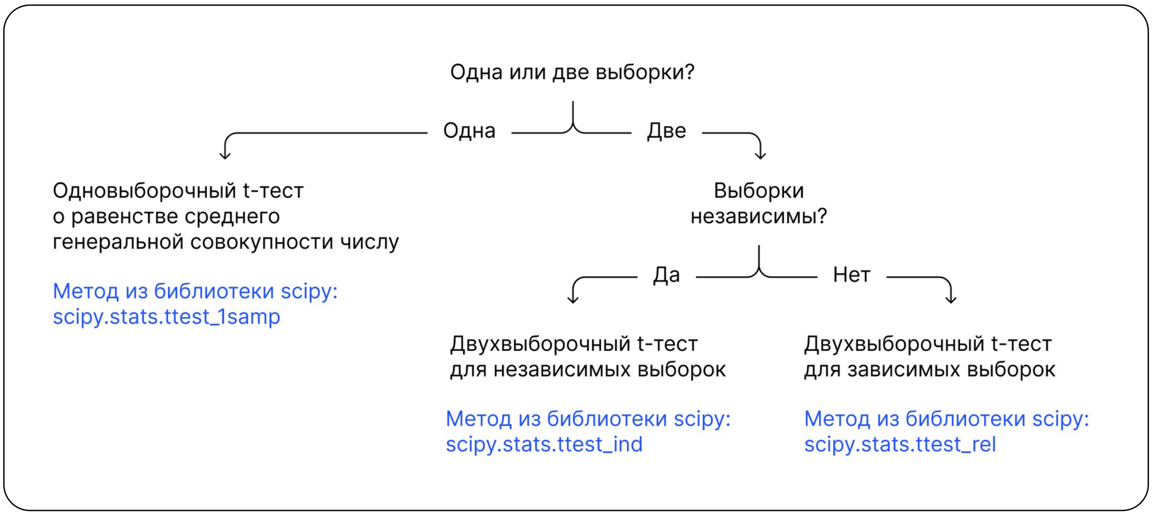


Таблица с методами Python для проведения t-тестов

В таблице указано, какие методы из библиотеки scipy используются для проведения рассмотренных вариантов t-тестов и какие аргументы им передаются.

В последней колонке описано, какая альтернативная гипотеза будет проверяться в зависимости от того, двусторонняя она, левосторонняя или правосторонняя.

Сколько выборок	**Какая гипотеза проверяется**	**Нужный метод и его аргументы**	**Выбор альтернативной гипотезы: аргумент `alternative`**
Одна выборка	О равенстве среднего генеральной совокупности числу, зафиксированному в нулевой гипотезе	`scipy.stats.ttest_1samp` Аргументы: набор данных выборки и предполагаемое среднее	Если не указывать или указать `alternative='two-sided'` — проверится двусторонняя гипотеза. Если указать `alternative='less'`, проверится левосторонняя альтернативная гипотеза о том, что истинное среднее меньше предполагаемого в H_0 значения. Если указать `alternative='greater'`, проверится правосторонняя гипотеза о том, что истинное среднее больше предполагаемого значения.
Две независимые выборки	О равенстве средних двух генеральных совокупностей	`scipy.stats.ttest_ind` Аргументы: наборы данных — две выборки. Необязательный аргумент: `equal_var=False` — указывайте, если полагаете, что дисперсии генеральных совокупностей не равны (если не указывать — принимает значение `True`).	Если не указывать или указать `alternative='two-sided'` — проверится двусторонняя гипотеза. Если указать `alternative='less'`, проверится левосторонняя альтернативная гипотеза о том, что истинное среднее генеральной совокупности, из которой взяли первую выборку, меньше, чем истинное среднее генеральной совокупности, из которой взяли вторую. Если указать `alternative='greater'`, проверится правосторонняя гипотеза.
Две зависимые (парные) выборки	О равенстве среднего значения генеральной совокупности до и после изменения	`scipy.stats.ttest_rel` Аргументы: наборы данных — парные выборки до и после изменения	Если не указывать или указать `alternative='two-sided'` — проверится двусторонняя гипотеза. Если указать `alternative='less'`, проверится левосторонняя альтернативная гипотеза о том, что истинное среднее генеральной совокупности до изменения меньше, чем истинное среднее генеральной совокупности после изменения. Если указать `alternative='greater'`, проверится правосторонняя гипотеза.

Примеры кода

Одновыборочный t-тест

```
from scipy import stats as st
import pandas as pd

# выборка
sample = pd.Series([33.073, 18.429, 23.118, 37.544, 28.771, 36.69,
                    29.697, 27.436, 22.671, 28.733, 19.844, 16.498,
                    29.955, 15.14, 19.748, 35.685, 30.516, 18.997,
                    19.351, 26.674, 32.112, 24.332, 31.48, 30.069,
                    25.106, 17.921, 30.905, 22.625, 19.615, 28.638])

print('Среднее выборки:', sample.mean())

# Предположенное в нулевой гипотезе число.
# С помощью теста проверяется гипотеза о том, что этому
# числу равно истинное среднее генеральной совокупности,
# из которой взята выборка
value = 28

# выбранный уровень статистической значимости
alpha = 0.05

results = st.ttest_1samp(
    sample, value,
    alternative='two-sided')
# если передать методу ttest_1samp значение аргумента alternative:
# 'two-sided' (или не указывать этот аргумент) — будет проверена
                          # двусторонняя альтернативная гипотеза
# 'less' — будет проверена левосторонняя альтернативная гипотеза
# 'greater' — будет проверена правосторонняя альтернативная гипотеза

# выведем полученное p-value
print('p-значение:', results.pvalue)

# вывод об отвержении или неотвержении нулевой гипотезы
if results.pvalue < alpha:
    print('Отвергаем нулевую гипотезу')
else:
    print('Не отвергаем нулевую гипотезу')
```

Проверка гипотез

Результат выполнения кода:

```
Среднее выборки: 26.045766666666667
p-значение: 0.09985381732739748
Не отвергаем нулевую гипотезу
```

Выборочное среднее равно 26.0457(6) и отличается от 28 — предположенного в нулевой гипотезе значения. Но это отличие недостаточно велико, чтобы сделать вывод о том, что истинное среднее ГС, из которой выборка взята, отличается от 28 — и при уровне статистической значимости 1%, и при 5%.

Если бы выборка оказалась с тем же или близким средним значением, но менее разбросанной вокруг него, то ESE было бы меньше, и нулевая гипотеза могла бы быть отвергнута. Например, если в качестве выборки в коде выше задать набор данных [26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,26,27], выборочное среднее оказывается почти таким же — 26.0(45), но p-value практически равно нулю: 5.84609739930291e-22, — в таком случае H_0, конечно, отвергается.

Таким образом, оценённый по выборке разброс ГС имеет ключевое значение при проверке гипотез. Одна и та же разница между предполагаемым в H_0 значением и выборочным средним может быть значимой, если разброс выборки маленький, и незначимой, если он большой.

Проверка гипотез

Двухвыборочный t-тест для независимых выборок

```
from scipy import stats as st
import numpy as np

# две выборки, взятые независимо друг от друга из двух ГС
sample1 = [7.1, 7.3, 9.8, 7.3, 6.4, 10.5, 8.7,
           17.5, 3.3, 15.5, 16.2, 0.4, 8.3]
sample2 = [12.1, 24.3, 10.4, 19.9, 19.7, 12.5, 17.6,
           5.0, 29.4, 10.5, 10.8, 13.4, 9.4, 8.7,
           2.5, 18.8, 4.8, 29.0, 7.7, 18.6, 16.7,
           14.2, 10.6]
print(f'среднее первой выборки: {np.mean(sample1)}')
print(f'среднее второй выборки: {np.mean(sample2)}')

# выбранный уровень статистической значимости
alpha = 0.05

results = st.ttest_ind(
    sample1,
    sample2,
    alternative='less')
# если передать методу ttest_ind значение аргумента alternative:
# 'two-sided' (или не указывать этот аргумент) — будет проверена
                         # двусторонняя альтернативная гипотеза
# 'less' — будет проверена левосторонняя альтернативная гипотеза
# 'greater' — будет проверена правосторонняя альтернативная гипотеза

# выведем полученное p-value
print('p-значение:', results.pvalue)

# вывод об отвержении или неотвержении нулевой гипотезы
if results.pvalue < alpha:
    print('Отвергаем нулевую гипотезу')
else:
    print('Не получилось отвергнуть нулевую гипотезу')
```

Результат выполнения кода:

```
среднее первой выборки: 9.1
среднее второй выборки: 14.200000000000001
p-значение: 0.014993918336210341
Отвергаем нулевую гипотезу
```

Проверка гипотез

- Этот код проверяет одностороннюю гипотезу: о том, что среднее ГС, из которой взята первая выборка, меньше, чем среднее ГС, из которой взята вторая выборка.
- Нулевая гипотеза отвергнута при уровне статистической значимости, равном 5%, но она уже не была бы отвергнута при уровне статистической значимости, равном 1%. Это значит, что не стоит делать однозначного вывода о том, что среднее ГС, из которой взята первая выборка, меньше, чем среднее ГС, из которой взята вторая выборка. Возможно, следует провести дополнительное исследование.
- Метод ttest_ind принимает на вход выборки разного размера.

Двухвыборочный t-тест для зависимых выборок

```
from scipy import stats as st

# две зависимые выборки: например, измерения одной
# переменной для одних и тех же объектов до и после изменения:
sample1 = [203, 219, 172, 220, 173, 225, 151, 222, 205, 213, 155, 168, 219,
           136, 158, 217, 183, 183, 183, 187, 196, 158, 185, 208, 202, 142,
           193, 145, 202, 192, 183, 190, 177, 182, 207, 177, 173, 150]
sample2 = [156, 184, 222, 194, 210, 158, 175, 187, 165, 198, 188, 181, 205,
           183, 189, 163, 202, 183, 142, 206, 124, 169, 222, 183, 186, 133,
           149, 208, 212, 142, 155, 183, 187, 178, 181, 162, 166, 130]
print(f'среднее первой выборки: {np.mean(sample1)}')
print(f'среднее второй выборки: {np.mean(sample2)}')

# выбранный уровень статистической значимости
alpha = 0.01

results = st.ttest_rel(
            sample1,
            sample2,
            alternative='greater')
# если передать методу ttest_rel значение аргумента alternative:
# 'two-sided' (или не указывать этот аргумент) — будет проверена
                          # двусторонняя альтернативная гипотеза
# 'less' — будет проверена левосторонняя альтернативная гипотеза
# 'greater' — будет проверена правосторонняя альтернативная гипотеза

# выведем полученное p-value
print('p-значение:', results.pvalue)

# вывод об отвержении или неотвержении нулевой гипотезы
if results.pvalue < alpha:
    print('Отвергаем нулевую гипотезу')
else:
    print('Не получилось отвергнуть нулевую гипотезу')
```

Проверка гипотез

Результат выполнения кода:

```
среднее первой выборки: 185.625
среднее второй выборки: 179.15
p-значение: 0.11459737791969755
Не получилось отвергнуть нулевую гипотезу
```

- Этот код проверяет одностороннюю гипотезу: о том, что до изменения (первая выборка) среднее значение переменной в генеральной совокупности было больше, чем после изменения (вторая выборка).
- Хотя среднее второй выборки и меньше, чем первой, нулевую гипотезу о равенстве средних до и после изменения отвергнуть не удалось.
- Метод ttest_rel принимает на вход только выборки одинакового размера.