

Конспект по теме «Статистические тесты»

Стандартная ошибка среднего

Распределение выборочных средних — это распределение, показывающее, какие значения принимает среднее значение случайной выборки из генеральной совокупности при многократном повторении эксперимента.

Стандартная ошибка среднего (Standard Error, SE) — это стандартное отклонение выборочных средних, которое зависит от количества наблюдений в выборке и стандартного отклонения генеральной совокупности.

$$SE = \frac{\sigma}{\sqrt{n}}.$$

Стандартная ошибка среднего (SE) указывает, насколько большая ожидается погрешность при измерении среднего значения по выборке. Чем больше дисперсия в самих данных и чем меньше выборка, тем больше погрешность.

Статистические гипотезы

Цель проверки **статистических гипотез** — на основе небольшой выборки сделать обоснованные предположения о генеральной совокупности.

Нулевая гипотеза, обозначаемая как H_0 , это исходное предположение о данных.

Альтернативная гипотеза, обозначаемая как H_1 , представляет собой предположение, которое противоположно нулевой гипотезе.

Критические значения формируют область допустимых значений для статистики, при условии, что нулевая гипотеза верна.

Ошибка первого рода — это ситуация, когда исследователь отвергает нулевую гипотезу, хотя она на самом деле верна.

Уровень значимости α — это вероятность совершения ошибки первого рода, то есть отклонить нулевую гипотезу, когда она верна.

Z-тест для среднего значения — это статистический тест, который используется, чтобы определить, отличается ли среднее значение выборки от предполагаемого среднего значения генеральной совокупности.

В процессе Z-теста вычисляется z-статистика, которая представляет собой отношение разности между средним значением выборки и средним значением генеральной совокупности к стандартной ошибке среднего.

$$z = \frac{\bar{x} - \mu}{SE},$$
$$SE = \frac{\sigma}{\sqrt{n}}.$$

Алгоритм проведения Z-теста:

1. Сформулировать H_0 и H_1 .
2. Выбрать α , найти соответствующие критические значения.
3. Вычислить выборочное среднее.
4. Вычислить z-статистику.
5. Сравнить z-статистику с критическими значениями:
6. Принять решение:
 - Если z-статистика вываливается, отклонить нулевую гипотезу.
 - Если z-статистика не вываливается, принимаем решение, что отклонить нулевую гипотезу нельзя.

P-value

P-value — вероятность получить определённое или ещё более экстремальное значение статистического критерия при условии, что нулевая гипотеза верна.

Ошибка первого рода, α — это ситуация, когда исследователь отвергает нулевую гипотезу, хотя она на самом деле верна.

Ошибка второго рода, β — это ситуация, когда исследователь не отвергает нулевую гипотезу, хотя верна альтернативная.

Мощность статистического теста, $1 - \beta$ — это вероятность правильно отклонить нулевую гипотезу, когда она действительно неверна.

T-тест и распределение Стьюдента

Оценка стандартной ошибки (Estimated Standard Error, ESE) — это стандартное отклонение выборочных средних, вычисленное на основе стандартного отклонения выборки.

$$ESE = \frac{s_X}{\sqrt{n}}.$$

В тесте для среднего, когда неизвестно стандартное отклонение генеральной совокупности, используется **t-статистика**. Это мера, которая помогает определить статистическую значимость различия между средним значением выборки и предполагаемым средним значением генеральной совокупности:

$$t = \frac{\bar{x} - \mu}{ESE}.$$

Распределение Стьюдента, или t-распределение, это вероятностное распределение, которое описывает поведение t-статистики. Оно имеет «тяжёлые» хвосты, это позволяет учесть большую неопределённость в оценках при малом объёме данных.

Степени свободы — это концепция в статистике, которая относится к количеству независимых значений в наборе данных, которые могут свободно варьироваться при расчёте некоторой статистической меры.

В контексте t-распределения Стьюдента, степени свободы определяют форму распределения и рассчитываются по формуле:

$$df = n - 1.$$

Алгоритм проведения Т-теста:

1. Формулируем H_0 , H_1
2. Выбираем α
3. Считаем выборочное среднее
4. Считаем выборочное стандартное отклонение
5. Считаем t-статистику
6. Считаем количество степеней свободы по формуле:
 $df = n - 1$,
где n — количество наблюдений в выборке
7. Считаем p-value
8. Сравниваем p-value с уровнем значимости:
 - Если p-value меньше уровня значимости, принимаем решение, что надо отклонить нулевую гипотезу.
 - Если p-value больше уровня значимости, то оснований отклонять нулевую гипотезу нет.

Сравнение групп

Одновыборочные тесты — тесты, которые позволяют исследовать одну выборку. **Двухвыборочные тесты** помогают сравнить две выборки.

Двухвыборочный Т-тест используют, когда наблюдения независимы и распределены нормально.

Для двухвыборочного Т-теста в случае, если размеры выборок равны:

- t -статистику рассчитывают по формуле:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{ESE},$$

$$\text{где } ESE = \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}},$$

- степени свободы рассчитывают по формуле:

$$df = \frac{(n - 1) \cdot (s_1^2 + s_2^2)^2}{s_1^4 + s_2^4},$$

\bar{x}_1 и \bar{x}_2 — выборочные средние для первой и второй выборки соответственно,

$\mu_1 - \mu_2$ — разность математических ожиданий двух выборок,

s_1 и s_2 — значения стандартных отклонений, рассчитанных на основе выборок,

где n — размер каждой выборки.

Бакетный тест используют, когда распределение наблюдений значительно отличается от нормального. Обычно данные в каждой выборке разбивают на 100 бакетов, далее вычисляют среднее значение по каждому бакету. К полученным данным применяют двухвыборочный Т-тест.

Z-тест для пропорций помогает сравнить доли определённых наблюдений в двух выборках. Для него z -статистику рассчитывают так:

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{ESE},$$

$$\text{где } ESE = \sqrt{\frac{\bar{p}_1 \cdot (1 - \bar{p}_1)}{n_1} + \frac{\bar{p}_2 \cdot (1 - \bar{p}_2)}{n_2}},$$

\bar{p}_1 — расчётная доля в первой выборке,

\bar{p}_2 — расчётная доля во второй выборке,

n_1 — количество наблюдений в первой выборке,

n_2 — количество наблюдений во второй выборке.