

СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

Статистическая гипотеза – любое предположение о виде закона распределения генеральной совокупности или ее числовых характеристиках, которое можно проверить, опираясь на данные выборки

В процессе проверки формулируются следующие статистические гипотезы

- H_0 гипотеза об отсутствии различий (*нулевая гипотеза*);
- H_1 гипотеза о наличии различий (*альтернативная гипотеза*).

Статистический критерий – правило, по которому принимается или отклоняется проверяемая гипотеза

Критерии проверки статистических гипотез делятся на две группы:

1) параметрические критерии, основанные на конкретных законах распределения;

Применяются главным образом для анализа нормально распределенных количественных признаков

2) непараметрические критерии, которые анализируют данные, не подчиняющиеся предположениям о законах распределения.

Применяются в остальных случаях:

а) для анализа количественных признаков независимо от вида их распределения;

б) для анализа качественных признаков.

- **Критическая область** – это множество возможных значений статистического критерия, при котором **нулевая гипотеза отклоняется**.
- Вероятность попадания значения критерия в эту область равна **уровню значимости**.
- **Уровнем значимости α** называется вероятность ошибки 1-го рода, заключающейся в отклонении нулевой гипотезы, когда она верна.
- В медико-биологических исследованиях в качестве критического порога значимости традиционно выбирается уровень 0,05, что допускает наличие ошибки первого рода 5 раз в 100 сравнениях.
- **Критические значение статистического критерия** определяются по специальным статистическим таблицам.

Таким образом, для проверки статистических гипотез используют следующие способы:

I способ

1) Сравнение **расчетного** или эмпирического (рассчитанного по выборке) значения статистического критерия с **критическим значением** (из таблиц).

- Если ***Расчетное значение < Критического значения***, то различие статистически незначимо, т.е. принимаем гипотезу H_0 об отсутствии различий.
- Если ***Расчетное значение > Критического значения***, то наблюдаемое различие значимо, т. е. принимаем гипотезу H_1 .

II способ:

2) Сравнение полученного (достигнутого) значения p с критическим уровнем значимости α (равным $0,05$).

- Если $p > \alpha$, то принимают гипотезу H_0 об отсутствии различий (различия являются статистически незначимыми).
- Если $p < \alpha$, то принимают гипотезу H_1 (различия являются статистически значимыми).

Параметрические и непараметрические критерии

- **Параметрические критерии** – это критерии, применение которых требует обязательного знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров. Например, известно, что выборки извлечены из генеральных совокупностей с *нормальным законом* распределения и *одинаковыми дисперсиями*.
- **Непараметрические критерии** не требуют знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров. Непараметрические методы не основываются на оценке параметров (таких как среднее или стандартное отклонение) при описании выборочного распределения величины.

Параметрический t-критерий Стьюдента (t-критерий Стьюдента; t-test)

Согласно теории статистики должны быть выполнены два условия применимости критерия Стьюдента:

- Результаты наблюдений случайным образом извлекаются из нормально распределенных генеральных совокупностей.
- Эти нормальные генеральные совокупности имеют одинаковые (однородные) дисперсии или известное отношение дисперсий.

Методика проверки с помощью t-критерия

1. Проверить, что данные являются количественными.
2. Проверить выборки на соответствие нормальному закону распределению.
3. Проверить гипотезу о равенстве дисперсий (например, по критерию Ливиня). Нулевая гипотеза о равенстве дисперсий не отклоняется, если достигнутый уровень значимости $p > 0,05$
4. Формулируется нулевая гипотеза H_0 о том, что средние значения признака в сравниваемых группах не различаются; формулируется альтернативная гипотеза H_1 о том, что средние значения различны ($a_1 \neq a_2$).

5. Вычисляется расчетное значение t-критерия

6. Формулируем выводы

- 1) Сравнивается полученное значение $t_{\text{расч}}$ расчетное с $t_{\text{крит}}$ табличным критическим.
- Если $t_{\text{расч}} < t_{\text{крит}}$, то различие между средними значениями двух связанных выборок статистически незначимо, т.е. принимаем гипотезу H_0 .
- Если $t_{\text{расч}} > t_{\text{крит}}$, то наблюдаемое различие между средними значениями двух связанных выборок значимо на уровне значимости (принимаем гипотезу H_1).
- 2) Если $p < 0,05$, то нулевая гипотеза отклоняется, а принимается альтернативная гипотеза о том, что «средние значения признака в группах различны».