Проверка гипотезы о среднем значении *a* нормально распределенной совокупности

- Пусть нулевая гипотеза H_0 состоит в том, что $a = a_0$.
- В качестве альтернативной гипотезы рассмотрим следующие случаи:
  - $H_1 = \{a \neq a_0\}$;
  - $H_1 = \{a > a_0\}$;
  - $H_1 = \{a < a_0\}$.

Статистика критерия

Статистика критерия

$$t = \frac{(\bar{x} - a_0) \cdot \sqrt{n}}{s}$$

имеет распределение Стьюдента с $k = n - 1$ степенями свободы, где n – объем выборки.

Если вычисленное значение статистики попадает в критическую область, то гипотеза H_0 отвергается.

Критическая область

I. В случае $H_1 = \{a \neq a_0\}$ критическая область нулевой гипотезы H_0 имеет вид $|t| > t_{\text{кр}}$, где $t_{\text{кр}}$ - критическая точка распределения Стьюдента для заданного уровня значимости α и степеней свободы $k = n - 1$.

II. В случае $H_1 = \{a > a_0\}$;критическая область нулевой гипотезы H_0 имеет вид $\mathrm{t} > t_{\text{кр}}$.

III. В случае $H_1 = \{a < a_0\}$;критическая область нулевой гипотезы H_0 имеет вид $\mathrm{t} < -t_{\text{кр}}$.

Замечание

Если дисперсия σ^2 генеральной совокупности известна, то статистика критерия

$$t = \frac{(\bar{x} - a_0) \cdot \sqrt{n}}{\sigma}$$

имеет стандартное нормальное распределение и для нахождения критических значений использую таблицу функции Лапласа:

- ✓ для двусторонней критической области

$$\Phi(t_{\text{кр}}) = \frac{1-\alpha}{2};$$

- ✓ для односторонней критической области $\Phi(t_{\text{кр}}) = \frac{1-2\alpha}{2}$.

Пример 1

Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 5{,}2$ извлечена выборка объема $n = 100$ и по ней найдена выборочная средняя $\bar{x} = 27{,}56$.

Требуется при уровне значимости 0,05 проверить нулевую гипотезу $H_0 = \{a = 26\}$ при альтернативной гипотезе $H_1 = \{a \neq 26\}$.

Решение примера 1

Найдем наблюдаемое значение критерия

$$t_{\text{набл}} = \frac{(\bar{x} - a_0) \cdot \sqrt{n}}{\sigma} = \frac{(27{,}56 - 26) \cdot \sqrt{100}}{5{,}2} = 3.$$

Найдем критическую точку двусторонней критической области при альтернативной гипотезе $H_1 = \{a \neq 26\}$ и уровне значимости 0,05 из соотношения $\Phi(t_{\text{кр}}) = \frac{1-\alpha}{2}$

Решение примера 1

или с помощью встроенной функции Excel

$$t_{\text{кр}} = \text{НОРМ.СТ.ОБР}\left(1 - \frac{\alpha}{2}\right),$$

$$t_{\text{кр}} \approx 1{,}96,$$

и, следовательно $|t_{\text{набл}}| = 3 > t_{\text{кр}}$.

В этом случае гипотеза H_0 отклоняется.

Проверка гипотезы о равенстве средних двух независимых выборок нормально распределенных совокупностей

- Пусть нулевая гипотеза H_0 состоит в том, что $a_1 = a_2$.
- В качестве альтернативной гипотезы рассмотрим следующие случаи:
  - $H_1 = \{a_1 \neq a_2\}$;
  - $H_1 = \{a_1 > a_2\}$;
  - $H_1 = \{a_1 < a_2\}$.

Статистика критерия

Статистика критерия

$$t = \frac{(\overline{x_1} - \overline{x_2})}{s \cdot \sqrt{\frac{1}{n_1} - \frac{1}{n_2}}},$$

где

$$s^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2},$$

имеет распределение Стьюдента с $k = n_1 + n_2 - 2$ степенями свободы, где n_1 и n_2 - объемы выборок.

Если вычисленное значение статистики попадает в критическую область, то гипотеза H_0 отвергается.

Замечание

В случае известных значений дисперсий генеральных совокупностей статистика критерия имеет вид

$$t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

и подчиняется закону стандартного нормального распределения.

Пример 2

Из двух партий изделий, изготовленных на двух одинаково настроенных станках извлечены малые выборки объемов $n = 10$ и $m = 12$. Получены следующие результаты:

контролируемый размер изделий первого станка x_i	3,4	3,5	3,7	3,9
частота (число изделий) n_i	2	3	4	1

контролируемый размер изделий второго станка y_i	3,2	3,4	3,6
частота (число изделий) m_i	2	2	8

Пример 2

При уровне значимости 0,02 проверить гипотезу $H_0 = \{a_1 = a_2\}$ при альтернативной гипотезе $H_1 = \{a_1 \neq a_2\}$. Предполагается, что случайные величины распределены нормально.

Решение примера 2

По данным выборкам найдем средние значения

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{4} x_i \cdot n_i = 3{,}6 \text{ и}$$

$$\bar{y} = \frac{1}{m} \cdot \sum_{i=1}^{3} y_i \cdot m_i = 3{,}5$$

и исправленные выборочные дисперсии

$$s_1^2 = \frac{n}{n-1} \cdot \left(\frac{1}{n} \cdot \sum_{i=1}^{4} x_i^2 \cdot n_i - \bar{x}^2\right) \approx 0{,}027 \text{ и}$$

$$s_2^2 = \frac{m}{m-1} \cdot \left(\frac{1}{m} \cdot \sum_{i=1}^{3} y_i^2 \cdot m_i - \bar{y}^2\right) \approx 0{,}025,$$

$$s^2 = \frac{(n-1) \cdot s_1^2 + (m-1) \cdot s_2^2}{n+m-2} \approx 0{,}026.$$

Решение примера 2

Тогда статистика критерия

$$t_{\text{набл}} = \frac{(\bar{x} - \bar{y})}{s \cdot \sqrt{\frac{1}{n} - \frac{1}{m}}} \approx 4{,}8.$$

Найдем $t_{\text{кр}}$ с помощью таблицы распределения Стьюдента с уровнем значимости $\alpha = 0{,}02$ и $k = n + m - 2 = 10 + 12 - 2 = 20$ степенями свободы или с помощью функции, встроенной в Excel

$$t_{\text{кр}} = \text{СТЬЮДЕНТ.ОБР.2Х}(\alpha; k) \approx 2{,}53.$$

Решение примера 2

Так как

$$t_{\text{набл}} \approx 4{,}8 > t_{\text{кр}} \approx 2{,}53,$$

гипотезу $H_0 = \{a_1 = a_2\}$ отвергаем.

Проверка гипотезы о значении дисперсии нормально распределенной совокупности

- Пусть нулевая гипотеза H_0 состоит в том, что $\sigma^2 = \sigma_0^2$.
- В качестве альтернативной гипотезы рассмотрим следующие случаи:
  - $H_1 = \{\sigma^2 \neq \sigma_0^2\}$;
  - $H_1 = \{\sigma^2 > \sigma_0^2\}$;
  - $H_1 = \{\sigma^2 < \sigma_0^2\}$.

Статистика критерия

Статистика критерия

$$t = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

имеет распределение Пирсона (χ^2) с $k = n - 1$ степенями свободы, где n – объем выборки.

Если вычисленное значение статистики попадает в критическую область, то гипотеза H_0 отвергается.

Критическая область

I. В случае $H_1 = \{\sigma^2 \neq \sigma_0^2\}$ находим критические точки t_1 и t_2 порядков $1 - \frac{\alpha}{2}$ и $\frac{\alpha}{2}$, соответственно. Если $t_1 < t_{\text{набл}} < t_2$, то отклонять H_0 нет оснований.

II. В случае $H_1 = \{\sigma^2 > \sigma_0^2\}$ гипотеза H_0 отклоняется, если $t_{\text{набл}} > t_{\text{кр}}$, где $t_{\text{кр}}$ - критическая точка порядка α.

III. В случае $H_1 = \{\sigma^2 < \sigma_0^2\}$ гипотеза H_0 отклоняется, если $t_{\text{набл}} < t_{\text{кр}}$, где $t_{\text{кр}}$ - критическая точка порядка $1 - \alpha$.

Пример.

Из нормально распределенной совокупности извлечена выборка объема $n = 21$, для которой вычислено значение $s^2 = 10{,}3$.

Требуется проверить нулевую гипотезу $H_0 = \{\sigma^2 = 9\}$ при альтернативной гипотезе $H_1 = \{\sigma^2 > 9\}$ с уровнем значимости $\alpha = 0{,}05$.

Решение примера

Вычислим значение статистики

$$t_{\text{набл}} = \frac{20 \cdot 10{,}3}{9} \approx 22{,}89$$

Критерий правосторонний, поэтому находим критическую точку $t_{\text{кр}}$ распределения χ^2 порядка $\alpha = 0{,}05$ с $k = 21 - 1 = 20$ степенями свободы.

$$t_{\text{кр}} = 31{,}4$$

Так как $t_{\text{набл}} < t_{\text{кр}}$, то нет оснований отклонять H_0.

Сравнение двух выборочных дисперсий нормально распределенных совокупностей

- Пусть нулевая гипотеза H_0 состоит в том, что $\sigma_1^2 = \sigma_2^2$.
- В качестве альтернативной гипотезы рассмотрим следующие случаи:
  - $H_1 = \{\sigma_1^2 \neq \sigma_2^2\}$;
  - $H_1 = \{\sigma_1^2 > \sigma_2^2\}$.

Статистика критерия

Статистика критерия

$$t = \frac{s_1^2}{s_2^2}$$

имеет распределение Фишера со степенями свободы $k_1 = n_1 - 1$ и $k_2 = n_2 - 1$, где n_1 и n_2 – объемы выборок.

Если наблюдаемое значение статистики $t_{\text{набл}} > t_{\text{кр}}$ распределения Фишера порядка α, со степенями свободы k_1 и k_2, то гипотеза H_0 отвергается.

Замечание. $s_1^2 > s_2^2$.

Пример

По двум независимым выборкам объемов $n_1 = 11$ и $n_2 = 14$ найдены выборочные дисперсии $s_1^2 = 0{,}76$ и $s_2^2 = 0{,}38$.

Требуется проверить нулевую гипотезу $H_0 = \{\sigma_1^2 = \sigma_2^2\}$ при альтернативной гипотезе $H_1 = \{\sigma_1^2 > \sigma_2^2\}$ с уровнем значимости $\alpha = 0{,}05$.

Решение примера

Критическое значение статистики находим по таблицам распределения Фишера для $\alpha = 0{,}05$ и степеней свободы $k_1 = 10$ и $k_2 = 13$:

$$t_{\text{кр}} = 2{,}67.$$

При этом наблюдаемое значение статистики

$$t_{\text{набл}} = \frac{s_1^2}{s_2^2} = \frac{0{,}76}{0{,}38} = 2.$$

Так как $|t_{\text{набл}}| < t_{\text{кр}}$, то дисперсии данных выборок различаются незначительно с вероятностью $\beta = 1 - \alpha$.