Задача 1

Тип гипотезы: гипотеза о среднем

Выборка напряжения батареек: 12,9; 11,6; 13,5; 13,9; 12,1; 11,9; 13,0 В

1. Формулировка гипотез

• Основная гипотеза (H_0): Среднее напряжение батареек равно 12 В $H_0: \mu = 12$

• Альтернативная гипотеза (H_1): Среднее напряжение батареек не равно 12 В $H_1: \mu \neq 12$

2. Наблюдаемое значение статистики критерия:

t-тест для одной выборки $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$, где

$\bar{x}$ – выборочное среднее, μ_0 – предполагаемое среднее (12 В)

s – выборочное стандартное отклонение, n – размер выборки

$$\bar{x} = \frac{12,9 + 11,6 + 13,5 + 13,9 + 12,1 + 11,9 + 13,0}{7} = 12,7$$

$$S = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(12,9-12,7)^2 + (11,6-12,7)^2 + \ldots + (13,0-12,7)^2}{6}} = 0,87$$

$$t = \frac{12,7 - 12}{0,87/\sqrt{7}} = 2,25$$

3. Статистика t имеет распределение Стьюдента с n−1=6 степенями свободы при H_0

4. Для уровня значимости $\alpha = 0,05$ критич. область двусторонняя

Для $\alpha = 0,05$ критич. значения: $\pm t_{0,025,6}$; $t_{0,025,6} = 2,447$

5. На уровне значимости 0,05 нет оснований отвергнуть H_0. Среднее напряжение батареек можно считать равным 12 В.

Задача 2

• Тип гипотезы: гипотеза о равенстве средних для парных наблюдений

H_0: Средние значения жира, измеренные двумя методами, одинаковы. $H_0: \mu_D = 0$

H_1: Средние значения жира, измеренные двумя методами, различаются $H_1: \mu_D \neq 0$

→ используем парный t-тест

$D_i = X_i - Y_i$ $\qquad t = \frac{\bar{D}}{S_D / \sqrt{n}}$ $\qquad \bar{D}$ – среднее разностей, S_D – станд. откл. разностей

$D_1 = 23{,}1 - 22{,}7 = 0{,}4$ $\qquad D_6 = 48{,}3 - 46{,}8 = 1{,}5$

$D_2 = 23{,}2 - 23{,}6 = -0{,}4$ $\qquad D_7 = 40{,}5 - 40{,}4 = 0{,}1$

$D_3 = 26{,}5 - 27{,}1 = -0{,}6$ $\qquad D_8 = 25{,}0 - 24{,}9 = 0{,}1$

$D_4 = 26{,}6 - 27{,}4 = -0{,}8$ $\qquad D_9 = 38{,}4 - 38{,}1 = 0{,}3$

$D_5 = 27{,}1 - 27{,}4 = -0{,}3$ $\qquad D_{10} = 23{,}5 - 23{,}8 = -0{,}3$

$$\bar{D} = \frac{0{,}4 - 0{,}4 - 0{,}6 - 0{,}8 - 0{,}3 + 1{,}5 + 0{,}1 + 0{,}1 + 03 - 0{,}3}{10} = 0{,}0$$

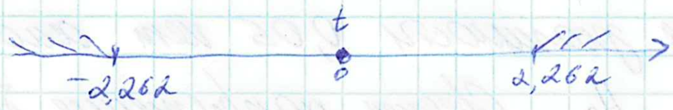
$$S_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{(0{,}4-0)^2 + (-0{,}4-0)^2 + \ldots + (-0{,}3-0)^2}{10-1}} = \sqrt{\frac{3{,}86}{9}} = 0{,}654$$

$$t = \frac{\bar{D}}{S_D / \sqrt{n}} = \frac{0{,}0}{0{,}654 / \sqrt{10}} = 0{,}0$$

Статистика t имеет распределение Стьюдента с $n-1=9$ степ. свободы при H_0

Для уровня значимости $\alpha = 0{,}05$ критическая область двусторонняя

Для $\alpha = 0{,}05$ крит. знач $\pm t_{0{,}025,9}$; $t_{0{,}025,9} = 2{,}262$



На уровне значимости 0,05 нет оснований отвергать H_0. Средние значения жира, измеренные двумя методами, можно считать равными

Задача 3

• Гипотеза о равенстве средних для независимых наблюдений

• H_0: Среднее давление у жителей городов и сельской местности одинаково $H_0: \mu_1 = \mu_2$

H_1: Среднее давление у жителей городов выше, чем у жителей сельской местности $H_1: \mu_1 > \mu_2$

Используем двухвыборочный t-тест для независимых выборок

Среднее значение для горожан ($\bar{X}_1$):

$$\bar{X}_1 = \frac{132+111+119+138+200+131+138+170+159+140}{10} = 143,8$$

Среднее значение для селян ($\bar{X}_2$):

$$\bar{X}_2 = \frac{115+190+127+155+148+121+116+121+197}{9} = 143,33$$

Стандартное отклонение для горожан (S_1):

$$S_1 = \sqrt{\frac{\Sigma(X_{1i}-\bar{X}_1)^2}{n_1-1}} = \sqrt{\frac{(132-143,8)^2+(111-143,8)^2+\ldots+(140-143,8)^2}{9}} = 30,01$$

Стандартное отклонение для селян (S_2):

$$S_2 = \sqrt{\frac{\Sigma(X_{2i}-\bar{X}_2)^2}{n_2-1}} = \sqrt{\frac{(115-143,33)^2+(190-143,33)^2+\ldots+(197-143,33)^2}{8}} = 32,27$$

t-статистика:

$$t = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} = \frac{143,8-143,33}{\sqrt{\frac{30,01^2}{10}+\frac{32,27^2}{9}}} = \frac{0,47}{\sqrt{90,06+115,87}} = \frac{0,47}{14,35} = 0,033$$

Статистика t имеет распределение Стьюдента с $n_1+n_2-2=17$ степенями свободы при H_0

Для уровня значимости $\alpha=0,05$ критич. область правосторонняя.

Для $\alpha=0,05$ критич. знач: $t_{0,05,17}$; $t_{0,05,17}=1,740$

На уровне значимости 0,05 у нас нет оснований отвергать H_0. Среднее давление у жителей городов и сельской местности можно считать одинаковым.

Задача 4

Гипотеза о доле (вероятности)

H_0: Доля мужчин равна 52%, $H_0: p = 0{,}52$

H_1: Доля мужчин не равна 52%, $H_1: p \neq 0{,}52$

Используем z-тест для доли

Выборочная доля мужчин $\hat{p} = \frac{2500}{5000} = 0{,}5$

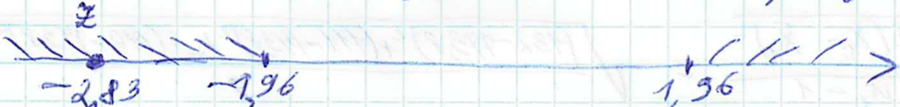
Z-статистика:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0{,}5 - 0{,}52}{\sqrt{\frac{0{,}52\cdot(1-0{,}52)}{5000}}} = \frac{-0{,}02}{0{,}00707} = -2{,}83$$

Z-статистика имеет нормальное распределение при H_0

Для уровня значимости $\alpha = 0{,}05$ критич. область двусторонняя

Для $\alpha = 0{,}05$ критич. значение: $\pm Z_{0.025}$; $Z_{0{,}025} = 1{,}96$



На уровне значимости 0,05 есть основания отвергнуть H_0. Доля мужчин не равна 52%

$\Rightarrow$ Смертность у мужчин и женщин разная.

Задача 5

Гипотеза о разности долей (вероятностей)

H_0: Доли заболевших одинаковы в обеих группах; $H_0: p_1 = p_2$

H_1: Доли заболевших отличаются; $H_1: p_1 \neq p_2$

Используем Z-тест для разности долей

Доли заболевших в группах:

$$\hat{p}_1 = \frac{3}{175} = 0{,}0171 \qquad \hat{p}_2 = \frac{32}{200} = 0{,}16$$

Общая доля заболевших $\hat{p} = \frac{3+32}{175+200} = \frac{35}{375} = 0{,}0933$

Z - статистика:

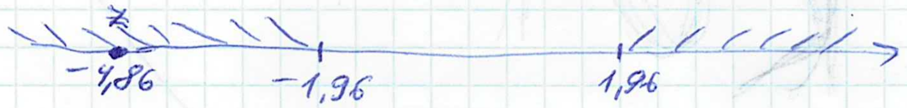
$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0{,}0171 - 0{,}16}{\sqrt{0{,}0933 \cdot (1 - 0{,}0933) \cdot \left(1/175 + \frac{1}{200}\right)}} =$$

$$= \frac{-0{,}1429}{\sqrt{0{,}0933 \cdot 0{,}9067 \cdot 0{,}0102}} = \frac{-0{,}1429}{0{,}0294} = -4{,}86$$

Z - статистика имеет нормальное распределение при H_0

Для уровня значимости $\alpha = 0{,}05$ критич. область двусторонняя

Для $\alpha = 0{,}05$ критич. значения: $\pm Z_{0.025}$; $Z_{0.025} = 1{,}96$



На уровне значимости 0,05 есть основания отвергнуть H_0. Доли заболевших отличаются, что указывает на эффективность лекарства от аллергии.