Статистика 6

Фонин А.Ю.

May 7, 2021

Exercise 1

$n = 20$, $\bar{X} = 266000$, std $= 24000$

Проверяем гипотезы

$$H_0 : \quad \mu = 250000$$
$$H_1 : \quad \mu > 250000$$

при уровне значимости $\alpha = 0.01$

Используем ЦПТ

$$z = \sqrt{n}\frac{\bar{X} - \mu}{\sigma} \xrightarrow{D} N(0,1) \text{ при верной } H_0$$

Тогда статистика на нашей выборке равна

$$z(X_1 \dots X_n) \approx 3$$

Доверительная область равна$(-\infty, z_{1-\alpha}) = (-\infty; 2.32)$ Значение статистики попало в критическую область, тогда отвергаем нулевую гипотезу

Exercise 2

Гипотетическое распределение

Br	Y	R	O	G	Bl
30%	20%	20%	10%	10%$_0$	10%

Наша выборка

Br	Y	R	O	G	Bl
117	135	79	41	36	38

Используем критерий согласия χ^2 для проверки, что наша выборка пришла из гипотетического распределения

$$H_0 : p_i = p(X = x_i)$$
$$H_1 : \exists j : p_j \neq p(X = x_i)$$

Статистика χ^2 имеет вид:

$$\mathbb{X}^2 = \sum_{i=1}^{m} \frac{(\theta_i - E_i)^2}{E_i} \xrightarrow[n\to\infty]{D} \chi^2_{m-1}, \text{ где } E_i = np_i; \quad \theta_i = \#\{X_i = x\}$$

Критическая область при m=6 и нашем α будет

$$[11; +\infty)$$

И так как статистика по нашей выборке попадает в нее

$$\mathbb{X}^2(X_1, \ldots, X_n) = 29.51\$$$

Отвергаем нулевую гипотезу

Exercise 3

Наши данные: $\sigma_0^2 = 0.0001$, $n = 15$, $std = 0.014$, $\alpha = 0.1$

Проверяем следующую нулевую гипотезу

$$H_0 : \quad \sigma = \sigma_0$$
$$H_1 : \quad \sigma > \sigma_0$$

Воспользуемся статистикой

$$t = \frac{(n-1)\bar{S}_0^2}{\sigma_0^2} \sim \chi^2_{n-1} \quad \text{где} \quad S_0^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

Тогда статистика по нашей выборке равна

$$t(X_1 \ldots X_n) = \frac{n \cdot std^2}{\sigma_0^2} = \frac{15 \cdot 0.0002}{0.0001} = 30$$

$X_{n-1,1-\alpha}$- квантиль уровня $1 - \alpha$. Тогда Критическая зона $(21, +\infty)$

Ответ: отвергаем H_0

Exercise 4

p_1 - вероятность, что мужчина остановится и спросит

p_2 - вероятность, что женщина остановилится и спросит

Выдвигаем гипотезу при $\alpha = 0.05$

$$H_0 : \quad p_1 = p_2$$
$$H_1 : \quad p_1 > p_2$$

Тогда

$$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{1}{m} + \frac{1}{n}}\sqrt{\tilde{p}(1-\tilde{p})}} \xrightarrow{D} N(0,1), \text{ где } \tilde{p} = \frac{n\hat{p}_1 + \hat{p}_2 m}{n+m}$$

А

$$\hat{p}_1 = \frac{300}{811}; \quad \hat{p}_2 = \frac{255}{750}$$

$t(X_1 \dots X_n) = 1.23$ не лежит в критической области

Ответ: нет оснований считать, что девушка чаще спрашивает

Exercise 5

Данные:

$n_1 = 16 \quad \bar{X}_1 = 6,82 \quad \bar{S}_1^2 = 0,64$

$n_2 = 19 \quad \bar{X}_2 = 6,25 \quad \overline{S_2^2} = 0,75$

Выдвигаем гипотезу при $\alpha = 0.05$

$$H_0 : \quad \mu_1 = \mu_2$$
$$H_1 : \quad \mu_1 > \mu_2$$

У нас нет предположений о нормальном распределении, тогда используем ЦПТ

$$z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \longrightarrow N(0,1)$$

$$z(X_1, X_2) = 2.021$$

Критическая область $(Z_{1-\alpha}, +\infty)$, где $Z_{1-\alpha}$ квантиль нормального распределения

$2.021 \notin (1,64; +\infty)$

Нет оснований отвергнуть нулевую гипотезу.