

Popisná štatistika

Modus

$$\hat{x} = x_i + \frac{h}{2} \frac{n_{i+1} - n_{i-1}}{2n_i - n_{i+1} - n_{i-1}}$$

Medián

$$\tilde{x} = a + h \frac{\frac{n}{2} - (n_1 + n_2 + \dots + n_{i-1})}{n_i}$$

Momenty

$$m_3 = m_3^* - 3m_2^*m_1^* + 2m_1^{*3}$$

$$m_4 = m_4^* - 4m_3^*m_1^* + 6m_2^*m_1^{*2} - 3m_1^{*4}$$

Náhodný výber z normálneho rozdelenia

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Y = \frac{n-1}{\sigma^2} S_X^2 \sim \chi^2(n-1)$$

$$U = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$$

$$T = \frac{\bar{X} - \mu}{S_X} \sqrt{n} \sim t(n-1)$$

$$Z = \bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

$$T = \frac{\bar{X} - \bar{Y}(\mu_1 - \mu_2)}{S^*} \sim t(n+m-2), \quad S^* = \sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}} \sqrt{\frac{n+m}{nm}}$$

$$F = \frac{S_X^2}{\sigma_1^2} : \frac{S_Y^2}{\sigma_2^2} \sim F(n-1, m-1)$$

Intervaly spoľahlivosti pre parametre normálneho rozdelenia

1.

a) Interval spoľahlivosti pre strednú hodnotu μ , rozptyl σ^2 známy

$$\left\langle \bar{X} - k_\alpha \frac{\sigma}{\sqrt{n}}, \quad \bar{X} + k_\alpha \frac{\sigma}{\sqrt{n}} \right\rangle$$

b) Interval spoľahlivosti pre strednú hodnotu μ , rozptyl σ^2 neznámy

$$\left\langle \bar{X} - t_\alpha(n-1) \frac{S_X}{\sqrt{n}}, \quad \bar{X} + t_\alpha(n-1) \frac{S_X}{\sqrt{n}} \right\rangle$$

$$\left\langle \bar{X} - t_{2\alpha}(n-1) \frac{S_X}{\sqrt{n}}, \quad \infty \right\rangle, \quad \left\langle -\infty, \quad \bar{X} + t_{2\alpha}(n-1) \frac{S_X}{\sqrt{n}} \right\rangle$$

c) Interval spoľahlivosti pre rozptyl σ^2 , stredná hodnota μ známa

$$\left\langle \frac{n S_0^2}{\chi_{\frac{\alpha}{2}}^2(n)}, \quad \frac{n S_0^2}{\chi_{1-\frac{\alpha}{2}}^2(n)} \right\rangle$$

$$\left\langle \frac{n S_0^2}{\chi_\alpha^2(n)}, \quad \infty \right\rangle, \quad \left\langle -\infty, \quad \frac{n S_0^2}{\chi_{1-\alpha}^2(n)} \right\rangle$$

d) Interval spoľahlivosti pre rozptyl σ^2 , stredná hodnota μ neznáma

$$\left\langle \frac{(n-1)S_X^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S_X^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right\rangle$$

$$\left\langle \frac{(n-1)S_X^2}{\chi_{\alpha}^2(n-1)}, \infty \right\rangle, \left(-\infty, \frac{(n-1)S_X^2}{\chi_{1-\alpha}^2(n-1)} \right)$$

2.

a) Interval spoľahlivosti pre rozdiel stredných hodnôt $(\mu_1 - \mu_2)$, rozptyly σ_1^2 a σ_2^2 známe

$$\left\langle \bar{X} - \bar{Y} - k_{\alpha} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \bar{X} - \bar{Y} + k_{\alpha} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right\rangle$$

b) Interval spoľahlivosti pre rozdiel stredných hodnôt $(\mu_1 - \mu_2)$, rozptyly $\sigma_1^2 \neq \sigma_2^2$ neznáme

$$\left\langle \bar{X} - \bar{Y} - t_{\alpha}(n+m-2) \sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}, \bar{X} - \bar{Y} + t_{\alpha}(n+m-2) \sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}} \right\rangle$$

$$\left\langle \bar{X} - \bar{Y} - t_{2\alpha}(n+m-2) \sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}, \infty \right\rangle$$

$$\left(-\infty, \bar{X} - \bar{Y} + t_{2\alpha}(n+m-2) \sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}} \right)$$

c) Interval spoľahlivosti pre rozdiel stredných hodnôt $(\mu_1 - \mu_2)$, rozptyly $\sigma_1^2 = \sigma_2^2$ neznáme

$$\langle \bar{X} - \bar{Y} - t_{\alpha}(n+m-2)S^*, \bar{X} - \bar{Y} + t_{\alpha}(n+m-2)S^* \rangle$$

$$(-\infty, \bar{X} - \bar{Y} + t_{2\alpha}(n+m-2)S^*), \langle \bar{X} - \bar{Y} - t_{2\alpha}(n+m-2)S^*, \infty \rangle$$

d) Interval spoľahlivosti pre podiel rozptylov $\frac{\sigma_1^2}{\sigma_2^2}$

$$\left\langle \frac{S_X^2}{S_Y^2 F_{\frac{\alpha}{2}}(n-1, m-1)}, \frac{S_X^2}{S_Y^2 F_{1-\frac{\alpha}{2}}(n-1, m-1)} \right\rangle$$

$$\left\langle \frac{S_X^2}{S_Y^2 F_{\alpha}(n-1, m-1)}, \infty \right\rangle, \left(-\infty, \frac{S_X^2}{S_Y^2 F_{1-\alpha}(n-1, m-1)} \right)$$

Testovanie štatistických hypotéz

Testy hypotéz o parametroch normálneho rozdelenia

1. Testy hypotéz o parametroch μ a σ^2 normálneho rozdelenia

a) Test hypotézy o μ , rozptyl σ^2 známy

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu \neq \mu_0$$

Testovacia štatistika

$$U = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$$

H_0 zamietame ak $|U| \geq k_\alpha$.

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu > \mu_0 \quad \text{resp.} \quad H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu < \mu_0$$

H_0 zamietame ak $U > k_{2\alpha}$ resp. $U < -k_{2\alpha}$

b) Test hypotézy o μ , rozptyl σ^2 neznámy.

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu \neq \mu_0$$

Testovacia štatistika

$$T = \frac{\bar{X} - \mu_0}{S_X} \sqrt{n}$$

H_0 zamietame ak $|T| \geq t_\alpha(n-1)$.

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu > \mu_0 \quad \text{resp.} \quad H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu < \mu_0$$

H_0 zamietame ak $T > t_{2\alpha}(n-1)$ resp. $T < -t_{2\alpha}(n-1)$.

c) Test hypotézy o σ^2 , stredná hodnota μ neznáma

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{proti} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Testovacia štatistika

$$\chi^2 = \frac{(n-1)S_X^2}{\sigma_0^2}$$

H_0 zamietame ak $\chi^2 \leq \chi_{1-\frac{\alpha}{2}}^2(n-1)$ alebo $\chi^2 \geq \chi_{\frac{\alpha}{2}}^2(n-1)$.

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{proti} \quad H_1 : \sigma^2 > \sigma_0^2 \quad \text{resp.} \quad H_0 : \sigma^2 = \sigma_0^2 \quad \text{proti} \quad H_1 : \sigma^2 < \sigma_0^2$$

H_0 zamietame ak $\chi^2 \geq \chi_\alpha^2(n-1)$ resp. $\chi^2 \leq \chi_{1-\alpha}^2(n-1)$.

2. Párový test

$$H_0 : \mu_1 - \mu_2 = \Delta \quad \text{proti} \quad H_1 : \mu_1 - \mu_2 \neq \Delta$$

Testovacia štatistika

$$T = \frac{\bar{X} - \Delta}{S_X} \sqrt{n}$$

H_0 zamietame ak $|T| \geq t_\alpha(n-1)$.

3. Test rovnosti parametrov dvoch normálnych rozdelení

a) Test rovnosti rozptylov, stredné hodnoty neznáme. F-test

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{proti} \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

Testovacia štatistika

$$F = \frac{S_X^2}{S_Y^2}$$

H_0 zamietame ak $F \leq F_{1-\frac{\alpha}{2}}(n-1, m-1)$ alebo $F \geq F_{\frac{\alpha}{2}}(n-1, m-1)$.

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{proti} \quad H_1 : \sigma_1^2 > \sigma_2^2 \quad \text{resp.} \quad H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{proti} \quad H_1 : \sigma_1^2 < \sigma_2^2$$

H_0 zamietame ak $F \geq F_{\alpha}(n-1, m-1)$ resp. $F \leq F_{1-\alpha}(n-1, m-1)$.

b) Test rovnosti stredných hodnôt, rozptyly rovnaké. Dvojjvýberový Studentov test

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 \neq \mu_2$$

Testovacia štatistika

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}} \sqrt{\frac{nm}{n+m}}$$

H_0 zamietame ak $|T| \geq t_{\alpha}(n+m-2)$.

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 > \mu_2 \quad \text{resp.} \quad H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 < \mu_2$$

H_0 zamietame ak $T > t_{2\alpha}(n+m-2)$ resp. $T < -t_{2\alpha}(n+m-2)$.

c) Test rovnosti stredných hodnôt, rozptyly rôzne. Cochranov- Coxov t-test

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 \neq \mu_2$$

Testovacia štatistika je

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

H_0 zamietame ak $|T| \geq t_{\alpha}^*$, kde $t_{\alpha}^* = \frac{\frac{S_X^2}{n} t_{\alpha}(n-1) + \frac{S_Y^2}{m} t_{\alpha}(m-1)}{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}$.

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 > \mu_2 \quad \text{resp.} \quad H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 < \mu_2$$

H_0 zamietame ak $T > t_{2\alpha}^*$ resp. $T < -t_{2\alpha}^*$.

Regresná a korelačná analýza

Regresná priamka

Systém normálnych rovníc pre priamku

$$\begin{aligned}\beta_0 n + \beta_1 \sum_{i=1}^n x_i &= \sum_{i=1}^n Y_i \\ \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i Y_i\end{aligned}$$

$$b_0 = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n x_i \sum_{i=1}^n Y_i x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}, \quad b_1 = \frac{n \sum_{i=1}^n Y_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}$$

$$b_0 = \bar{Y} - b_1 \bar{x}.$$

$$s_e^2 = \frac{\sum_{i=1}^n Y_i^2 - b_0 \sum_{i=1}^n Y_i - b_1 \sum_{i=1}^n x_i Y_i}{n - 2}.$$

Rozptyly odhadov

$$s^2(b_0) = \frac{s_e^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}, \quad s^2(b_1) = \frac{s_e^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Testy hypotéz o parametroch

$$H_0: \beta_1 = 0 \quad \text{proti} \quad H_1: \beta_1 \neq 0$$

Testovacia štatistika

$$T_1 = \frac{(b_1 - 0)}{s(b_1)}$$

H_0 zamietame na hladine významnosti α , ak $|T_1| \geq t_\alpha(n-2)$.

$$H_0: \beta_0 = 0 \quad \text{proti} \quad H_1: \beta_0 \neq 0$$

Testovacia štatistika

$$T_0 = \frac{(b_0 - 0)}{s(b_0)}$$

H_0 zamietame na hladine významnosti α , ak $|T_0| \geq t_\alpha(n-2)$.

Intervaly spoľahlivosti

100.(1 - α)% interval spoľahlivosti pre parameter β_1

$$\langle b_1 - t_\alpha(n-2) s(b_1); \quad b_1 + t_\alpha(n-2) s(b_1) \rangle$$

100.(1 - α)% interval spoľahlivosti pre parameter β_0

$$\langle b_0 - t_\alpha(n-2) s(b_0); \quad b_0 + t_\alpha(n-2) s(b_0) \rangle$$

Test adekvátnosti regresného modelu, analýza rozptylu pre regresnú priamku

$$S_C = \sum_{i=1}^n Y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n Y_i \right)^2$$
$$S_M = b_0 \sum_{i=1}^n Y_i + b_1 \sum_{i=1}^n x_i Y_i - \frac{1}{n} \left(\sum_{i=1}^n Y_i \right)^2$$
$$S_R = \sum_{i=1}^n Y_i^2 - b_0 \sum_{i=1}^n Y_i - b_1 \sum_{i=1}^n x_i Y_i$$

Test adekvátnosti regresného modelu

H_0 : model nie je adekvátny proti H_1 : model je adekvátny

Testovacie kritérium

$$F = \frac{S_M(n-2)}{S_R}$$

H_0 zamietame na hladine významnosti α , ak $F \geq F_\alpha(1, n-2)$.

Združené regresné priamky

$$\hat{Y} - \bar{Y} = r_{XY} \frac{s_y}{s_x} (x - \bar{x})$$
$$\hat{X} - \bar{x} = r_{XY} \frac{s_x}{s_y} (y - \bar{y})$$

Koeficient korelácie

$$r_{XY} = \frac{\sum_{i=1}^n x_i Y_i - n\bar{x}\bar{Y}}{\sqrt{\left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) \left(\sum_{i=1}^n Y_i^2 - n\bar{Y}^2 \right)}}$$

Index korelácie

$$I_{YX} = \sqrt{\frac{\sum_{i=1}^n \hat{Y}_i^2 - n\bar{Y}^2}{\sum_{i=1}^n Y_i^2 - n\bar{Y}^2}}$$