

Tvorba histogramu

Počet intervalů

$k=1+3,3 \cdot \log(n)$ Sturgersovo pravidlo

$k = \text{celá část } 5 \cdot \log(n)$

$k = \text{odmocnina}(n)$

Šířka intervalů

$0,05R \leq h \leq 0,08R$

Rozdělení pravděpodobností

Transformace souboru dat s normálním rozdělením $N(\mu, \sigma^2)$ pomocí tzv. Z-transformace na soubor se standardním normálním rozdělením

$$Z = \frac{x - \mu}{\sigma}$$

Binomické rozdělení

střední hodnota: $E(x) = np$

rozptyl: $\sigma^2(X) = np(1 - p) = npq$

Testování statistických hypotéz

1) 1 výběrový soubor

χ^2 test

testovací kritérium

$$\chi^2 = \sum_{j=1}^k \frac{(n_{ej} - n_{oj})^2}{n_{oj}}$$

, n_e – experimentální četnosti; n_o – očekávané četnosti

kritická hodnota $\chi^2(1-\alpha; k-s-1)$, kde k je počet intervalů, s je počet parametrů rozdělení

Kolmogorov-Smirnovův test pro jeden výběr

testovací kritérium

$$D_1 = \max |F_{ej} - F_{oj}|$$

Kritická hodnota

Pro malá $n \leq 40$ z tabulek

Pro $n > 40$ dopočtu:

$$D_{2,0,05} = 1,36 \sqrt{n} \quad D_{2,0,01} = 1,53 \sqrt{n}$$

Jednovýběrový t-test o střední hodnotě

testovací kritérium

$$t = \frac{\bar{X} - \bar{\mu}}{s} \sqrt{n}$$

Kritická hodnota $t_{k(1-\alpha/2; n-1)}$ – oboustranná varianta testu

$t_{k(1-\alpha; n-1)}$ – jednostranná varianta testu

Testování odlehlých hodnot

Grubbsův test

Testovací kritérium pro dolní a horní odlehlou hodnotu

$$T_1 = \frac{\bar{x} - x_{(1)}}{s}$$

$$T_n = \frac{x_{(n)} - \bar{x}}{s}$$

kritická hodnota pro α a n je tabelována

Dean-Dixonův test

Testovací kritérium pro dolní a horní odlehlou hodnotu

$$Q_{\text{d}} = \frac{r_1 - r_2}{r_{n-1} - r_2}$$
$$Q_{\text{h}} = \frac{r_n - r_{n-1}}{r_n - r_2}$$

kritická hodnota pro α a n je tabelována

Testování statistických hypotéz

2) 2 výběrové soubory

Párový t-test na střední hodnotu

testovací kritérium $t = \frac{\bar{d}}{s_d/\sqrt{n}}$

kritická hodnota $t_k(1-\alpha/2; n-1)$ – oboustranná varianta testu

kritická hodnota $t_k(1-\alpha; n-1)$ – jednostranná varianta testu

F test

Testovací kritérium

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

Kritická hodnota $F_k(1-\alpha/2; n_1-1, n_2-1)$

Wilcoxonův test

testovací statistiky T^+ a nebo T^- (součet kladných a záporných rozdílů v pořadí párovaných hodnot)

Kritická hodnota pro příslušnou α a n (počet párů s nenulovým rozdílem) - tabelována

Kolmogorov-Smirnovův test shody pro dva výběry

Testovací kritérium

$$D_2 = \max_j |F_{1j} - F_{2j}|$$

kritická hodnota – dopočtu ze vzorce

$$D_{2,005} = 0,36 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$
$$D_{2,001} = 0,53 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Mann-Whitney test

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U' = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

kde R_1 je součet pořadí v souboru 1, R_2 je součet pořadí v souboru 2.
kritická hodnota $U_k(\alpha; n_1, n_2)$ - tabelována

Korelační a regresní analýza

kovariance

kovariance základního souboru

$$\text{COV}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

kovariance výběrového souboru

$$\text{COV}(X, Y) = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{n-1}$$

výpočet směrnice regresní přímky $y = ax + b$

$$a = \frac{\text{COV}(X, Y)}{S_x^2}$$

Pearsonův korelační koeficient

$$r_{xy} = \frac{\text{COV}(X, Y)}{S_x S_y}$$

Test na korelační koeficient

Testovací kritérium

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2}$$

kritická hodnota $t_{k(1-\alpha/2; n-2)}$

Spearmanův koeficient pořadové korelace

$$SR = \frac{1}{n} \sum_{i=1}^n d_i^2$$

Intervaly spolehlivosti

Interval spolehlivosti pro střední hodnotu

$$\bar{x} - t_{1-\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{1-\alpha/2} \frac{s}{\sqrt{n}}$$

Interval spolehlivosti pro rozptyl

$$\frac{(n-1) \cdot s^2}{\chi_{1-\alpha/2}^2} < \sigma^2 < \frac{(n-1) \cdot s^2}{\chi_{\alpha/2}^2}$$