

DESKRIPTIVNÍ STATISTIKA

PRŮMĚR

$$M = \frac{\sum x}{n}$$

MEDIÁN INTERVALOVÝCH ČETNOSTÍ

$$Md = L_p + W \left(\frac{n}{2} - f_p \right) / f_m$$

POZICE PERCENTILU

$$k = n \times \frac{p}{100}$$

VARIAČNÍ ROZPĚTÍ

$$R = X_{max} - X_{min}$$

(INTER)KVARTILOVÉ ROZPĚTÍ

$$IQR = Q_3 - Q_1$$

ROZPTYL (VÝBĚROVÝ)

$$s^2 = \frac{\sum (x - M)^2}{n - 1}$$

$$\text{Populační } s^2 = \frac{\sum (x - M)^2}{n}$$

SMĚRODATNÁ ODCHYLKA

$$s = \sqrt{s^2}$$

TRANSFORMACE SKÓRŮ

$$z_i = \frac{x_i - M}{s}$$

$$T_i = 50 + 10z_i$$

$$IQ_i = 100 + 15z_i$$

PRAVDĚPODOBNOST

$$P(A) = \frac{m}{n}$$

„NEBO“

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

„A“

$P(A \cap B) = P(A) \times P(B) \Rightarrow$ pro nezávislé jevy

$P(A \cap B) = P(A) + P(B) - P(A \cup B) \Rightarrow$ pro závislé jevy

ŠANCE

$$O(A) = \frac{P(A)}{P(A')} = \frac{P(A)}{1 - P(A)}$$

POMĚR ŠANCÍ

$$OR_{12} = \frac{O_1}{O_2}$$

PODMÍNĚNÁ

PRAVDĚPODOBNOST

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

CELÝ BAYESŮV TEORÉM

$$\frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(A') \cdot P(B|A')}$$

KORELACE

KOVARIANCE

$$c_{xy} = \frac{\sum_{i=1}^n (x_i - M_x)(y_i - M_y)}{n - 1}$$

PEARSONŮV KORELAČNÍ KOEFICIENT

$$r_{xy} = \frac{c_{xy}}{s_x s_y}$$

$$r_{xy} = \frac{\sum_{i=1}^n z_{xi} z_{yi}}{n - 1}$$

-alternativa

$$= \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

PARCIÁLNÍ KORELACE

$$r_{BC.A} = \frac{r_{BC} - (r_{BA} \cdot r_{CA})}{\sqrt{1 - r_{BA}^2} \cdot \sqrt{1 - r_{CA}^2}}$$

SEMIPARCIÁLNÍ KORELACE

$$r_{B(C.A)} = \frac{r_{BC} - r_{BA} \cdot r_{CA}}{\sqrt{1 - r_{CA}^2}}$$

SPEARMANŮV KORELAČNÍ KOEFICIENT

$$r_s = 1 - \frac{6 \cdot \sum (i_x - i_y)^2}{n \cdot (n^2 - 1)^2}$$

KENDALŮV KOEFICIENT POŘADOVÉ KORELACE

$$\tau = \frac{2(K - D)}{n(n - 1)} = \frac{K - D}{K + D}$$

VNITŘNÍ KONZISTENCE

$$r_{tt} = \frac{k \cdot r_m}{1 + (k - 1) \cdot r_m}$$

TEST SIGNIFIKANCE PEARSONOVA KORELAČNÍHO KOEFICIENTU

H0: $\rho=0$: NORM.S.DIST (Z/sz; 1)

H0: $\rho=c$: NORM.S.DIST (Dz/sz; 1)

Dz = FISCHER (r) – FISCHER (c)

LINEÁRNÍ REGRESE

$$Y = Y' + e = f_x + e$$

$$Y' = a + bx$$

$$a = m_y - b m_x$$

$$b = r_{xy} \frac{s_y}{s_x}$$

$$Z_y = r \cdot z_x$$

ÚSPĚŠNOST PREDIKCE

$$s_{reg}^2 = s_{Y'}^2 = \frac{\sum (M_y - Y')^2}{n - 1}$$

$$s_{res}^2 = s_e^2 = \frac{\sum (Y - Y')^2}{n - 1}$$

$$R^2 = \frac{s_{reg}^2}{s_Y^2} = \frac{s_{Y'}^2}{s_Y^2}$$

$$s_{Y'}^2 = s_Y^2 \cdot R^2$$

$$s_Y^2 = s_{reg}^2 + s_{res}^2 = s_{Y'}^2 + s_e^2$$

$$s_{res}^2 = s_e^2 = s_Y^2 (1 - R^2)$$

$$d = 2r / \sqrt{1 - r^2}$$

INDUKCE

$$v = df$$

PRO PRŮMĚR

SMĚRODATNÁ CHYBA

$$s_M = \frac{s}{\sqrt{N}}; \sigma_M = \frac{\sigma}{\sqrt{N}};$$

INTERVAL SPOLEHLIVOSTI

$$CI = M \pm (z_{crit} \cdot \sigma_m)$$

$$CI = M \pm (t_{crit} \cdot s_m)$$

t: T.INV (1 - $\alpha/2$; df)

z: NORM.S.INV (1 - $\alpha/2$; 0; 1)

PRO KORELACI

SMĚRODATNÁ CHYBA

$$s_z = \frac{1}{\sqrt{n - 3}}$$

FISHEROVO Z

$$Z = FISHER(r) = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

INTERVAL SPOLEHLIVOSTI

$$CI = FISHERINV(Z \pm (z_{crit} \cdot s_z))$$

z: NORM.S.INV (1 - $\alpha/2$; 0; 1)

TESTY

T-TEST

PRO JEDEN VÝBĚR

SE	$\sigma_m = \frac{\sigma}{\sqrt{N}}$	$s_m = \frac{s}{\sqrt{N}}$
df	-----	$df = n - 1$
t	$z = \frac{m - \mu}{\sigma_m}$	$t = \frac{m - \mu}{s_m}$
funkce	$2*(1-NORM.S.DIST(z))$	$2*(1-T.DIST(t;df;1))$

PRO DVA NEZÁVISLÉ VÝBĚRY

SE	$s_d = s_{pooled} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
df	$df = n_1 + n_2 - 2$
t	$t = \frac{m_1 - m_2}{s_d} = \frac{d}{s_d}$
funkce	$2*(1-T.DIST(t;df;1))$
ES	$Cohenovo\ d = \frac{d}{s_{pooled}}$

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Když $n_1=n_2$ lze použít $s_{pooled} = \sqrt{\frac{s_1^2 + s_2^2}{2}}$

PRO DVA ZÁVISLÉ VÝBĚRY

Sd	$s_d = \frac{s_{pooled}}{\sqrt{N}}$
df	$df = n - 1$
t	$t = \frac{m_1 - m_2}{s_d} = \frac{d}{s_d}$
funkce	$2*(1-T.DIST(t;df;2))$
ES	$Cohenovo\ d = \frac{d}{s_{pooled}}$

$$s_{pooled} = \sqrt{s_1^2 + s_2^2 - 2r s_1 s_2}$$

$$CI = (d \pm t \cdot s_d)$$

PŘEVODY COHENOVA D

$$d' = \frac{m_1 - m_2}{s_{control}}$$

$$r = \sqrt{\frac{d^2}{d^2 + 4}}; \quad d = \frac{2r}{\sqrt{1 - r^2}}$$

CHI-KVADRÁT

TEST DOBRÉ SHODY

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^k \frac{(F_o - F_e)^2}{F_e}$$

$$df = k - 1$$

TEST HOMOGENITY

$$m_{ij} = \frac{n_i \cdot n_j}{n}$$

$$\chi^2 = \sum_{i=1}^s \sum_{j=1}^r \frac{(n_{ij} - m_{ij})^2}{m_{ij}} = \sum_{i=1}^k \frac{(F_o - F_e)^2}{F_e}$$

$$df = (r - 1)(s - 1)$$

STANDARDIZOVANÁ REZIDUA

$$R_{ij} = \frac{n_{ij} - m_{ij}}{\sqrt{m_{ij}}} = \frac{f_{oij} - f_{eij}}{\sqrt{f_{eij}}}$$

Síla vztahu v kontingenční tabulce

- tabulka 2x2 Phi $\varphi = \sqrt{\frac{\chi^2}{n}}$

- tabulka 3x3 a více Pearson C = $\sqrt{\frac{\chi^2}{\chi^2 + n}}$
(ve čtvercových tabulkách)

- tabulka r x s Cramerovo V = $\sqrt{\frac{\chi^2}{n(k-1)}}$

Výběrové rozložení relativní četnosti p

$$SD: \sqrt{p(1-p)/N}$$

$$CI: (p - z_{1-\alpha/2} \times SD) \quad CI: p \pm 2\sigma_p.$$

Pravděpodobnost výskytu alespoň 1 chyby I. typu u k nezávislých srovnání $p = 1 - (1 - \alpha)^K$

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

$$P_+ = \frac{SeP(D_+)}{SeP(D_+) + (1 - Sp)(1 - P(D_+)}}$$

permutace n prvku: n!

kombinace r prvku z n-prvkove množiny: $n! / (r!(n-r)!)$

$Md = X_{(N+1)/2}$ $X = \text{liché}$ $= X_{(N/2 + ((N+1)/2)/2)}$ $X = \text{sudé}$

ZNAMÉNKOVÝ TEST

$$z = (2Z_+ - n) / \sqrt{n}$$

$$P = 2*(1-NORM.S.DIST(z))$$

MEDIÁNOVÝ TEST

$$z = \frac{(ad - bc)\sqrt{n}}{\sqrt{(a+b)(b+d)(a+c)(c+d)}}$$

	Sk A	Sk B	Σ
<Md	a	b	a+b
>Md	c	d	c+d
Σ	a+c	b+d	n