VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE
FAKULTA INFORMATIKY A STATISTIKY
Katedra statistiky a pravděpodobnosti

STATISTIKA PRO INFORMATIKY

VZORCE PRO 4ST204

verze 1.02
poslední aktualizace: 26. 1. 2021

Popisná statistika

$$p_i = \frac{n_i}{n} \qquad \sum_{i=1}^{k} n_i = n \qquad \sum_{i=1}^{k} p_i = 1 \qquad i = 1, 2, ..., k$$

$$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$$

$$\tilde{x}_P, 0 < P < 1 \qquad \tilde{x}_P = x_{(z_P)}, \qquad nP < z_P < nP + 1,$$

$$\tilde{x}_P = \frac{x_{(z_P)} + x_{(z_P+1)}}{2}, \qquad nP = z_P$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i} \qquad \bar{x} = \sum_{i=1}^{k} x_i p_i$$

$$\bar{x}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \qquad \bar{x}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{x_i}} \qquad \bar{x}_H = \frac{1}{\sum_{i=1}^{k} \frac{p_i}{x_i}}$$

$$\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$$

$$R = x_{\max} - x_{\min}$$

$$s_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \qquad s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n} \right)^2$$

$$s_x^2 = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i}{\sum_{i=1}^{k} n_i} \qquad s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^{k} x_i^2 n_i}{\sum_{i=1}^{k} n_i} - \left(\frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i} \right)^2$$

$$s_x^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i \qquad s_x^2 = \overline{x^2} - \bar{x}^2 = \sum_{i=1}^{k} x_i^2 p_i - (\sum_{i=1}^{k} x_i p_i)^2$$

$$s_x^2 = \overline{s^2} + s_{\bar{x}}^2 = \frac{\sum_{i=1}^{k} s_{ix}^2 n_i}{\sum_{i=1}^{k} n_i} + \frac{\sum_{i=1}^{k} (\bar{x}_i - \bar{x})^2 n_i}{\sum_{i=1}^{k} n_i} \qquad \bar{x} = \frac{\sum_{i=1}^{k} \bar{x}_i n_i}{\sum_{i=1}^{k} n_i}$$

$$s_x^2 = \sum_{i=1}^{k} s_{ix}^2 p_i + \sum_{i=1}^{k} (\bar{x}_i - \bar{x})^2 p_i \qquad \bar{x} = \sum_{i=1}^{k} \bar{x}_i p_i$$

$$s_x = \sqrt{s_x^2} \qquad s^2 = \frac{n}{n-1} s_x^2 \qquad s = \sqrt{s^2} \qquad v_x = \frac{s_x}{\bar{x}}$$

Indexní analýza

Bazické a řetězové indexy (v čase uspořádaná řada hodnot $y_1, y_2, ..., y_T$)

$$I_{t/1} = \frac{y_t}{y_1} = I_{2/1}\, I_{3/2} ... I_{t/t-1} \qquad I_{t/t-1} = \frac{y_t}{y_{t-1}} = \frac{I_{t/1}}{I_{t-1/1}}$$

Individuální indexy a diference jednoduché

$$p = \frac{Q}{q} \qquad IQ = Iq \cdot Ip$$

$$Ip = \frac{p_1}{p_0} \quad \Delta p = p_1 - p_0 \qquad Iq = \frac{q_1}{q_0} \quad \Delta q = q_1 - q_0 \qquad IQ = \frac{Q_1}{Q_0} \quad \Delta Q = Q_1 - Q_0$$

Individuální indexy a diference složené

$$I_{\Sigma q} = \frac{\sum q_1}{\sum q_0} = \frac{\sum Iq \,.\, q_0}{\sum q_0} = \frac{\sum q_1}{\sum \frac{q_1}{Iq}} \qquad \Delta_{\Sigma q} = \sum q_1 - \sum q_0$$

$$I_{\Sigma Q} = \frac{\sum Q_1}{\sum Q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum IQ \,.\, Q_0}{\sum Q_0} = \frac{\sum Q_1}{\sum \frac{Q_1}{IQ}} \qquad \Delta_{\Sigma Q} = \sum Q_1 - \sum Q_0$$

$$\overline{Ip} = \frac{\overline{p}_1}{\overline{p}_0} = \frac{\frac{\sum Q_1}{\sum q_1}}{\frac{\sum Q_0}{\sum q_0}} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{\sum Q_1}{\sum \frac{Q_1}{p_1}}}{\frac{\sum Q_0}{\sum \frac{Q_0}{p_0}}} \qquad \Delta \overline{p} = \overline{p}_1 - \overline{p}_0 = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0}$$

Souhrnné indexy

$$Ip^{(L)} = \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{\sum Ip \,.\, p_0 q_0}{\sum p_0 q_0} = \frac{\sum Ip \,.\, Q_0}{\sum Q_0} \qquad Ip^{(P)} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{Ip}} = \frac{\sum Q_1}{\sum \frac{Q_1}{Ip}}$$

$$Ip^{(F)} = \sqrt{Ip^{(L)} .\ Ip^{(P)}}$$

$$Iq^{(L)} = \frac{\sum p_0 q_1}{\sum p_0 q_0} = \frac{\sum Iq \,.\, p_0 q_0}{\sum p_0 q_0} = \frac{\sum Iq \,.\, Q_0}{\sum Q_0} \qquad Iq^{(P)} = \frac{\sum p_1 q_1}{\sum p_1 q_0} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{Iq}} = \frac{\sum Q_1}{\sum \frac{Q_1}{Iq}}$$

$$Iq^{(F)} = \sqrt{Iq^{(L)} \cdot\ Iq^{(P)}}$$

Pravděpodobnost

Počet pravděpodobnosti

$P(A) = \frac{m}{n}$ $\quad P\left(\bar{A}\right) = 1 - P(A)$

$P(A \cap B) \leq \min(P(A), P(B))$ $\qquad P(A \cap B) = P(A)\, P(B)$

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

Náhodné veličiny

$P(x) = P(X = x)$ $\quad F(x) = P(X \leq x) = \sum_{x_j \leq x} P(x_j)$ $\quad P(x_1 < X \leq x_2) = \sum_{x_1 < x \leq x_2} P(x) = F(x_2) - F(x_1)$

$E(X) = \sum_x x\, P(x)$ $\quad D(X) = E(X^2) - \left[E(X)\right]^2 = \sum_x x^2 P(x) - \left[\sum_x x P(x)\right]^2$

$f(x) = F'(x)$ $\quad \int_{-\infty}^{\infty} f(x)dx = 1$ $\quad F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$

$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1)$

x_P, $0 < P < 1$ $\quad F\left(x_P\right) = P$ $\quad x_P = F^{-1}(P)$

$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$ $\quad D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} x f(x)dx\right]^2$ $\quad \sigma = \sqrt{D(X)}$

Pravděpodobnostní rozdělení

Alternativní rozdělení $A(\pi)$

$P\left(x\right) = \pi^x (1-\pi)^{1-x}$ $\quad x = 0, 1, \quad 0 < \pi < 1$

$E(X) = \pi$ $\quad D\left(X\right) = \pi(1-\pi)$

Binomické rozdělení $Bi(n,\pi)$

$P(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$ $\quad x = 0, 1, 2, \ldots, n, \quad n \in N, 0 < \pi < 1$

$E\left(X\right) = n\pi$ $\quad D\left(X\right) = n\pi\left(1-\pi\right)$

Poissonovo rozdělení $Po(\lambda)$

$P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ $\quad x = 0, 1, \ldots, \lambda > 0,$

$E(X) = \lambda$ $\quad D(X) = \lambda$

Hypergeometrické rozdělení $Hg(M, N, n)$; v R: *hyper*(, $m = M$, $n = N\text{-}M$, $k = n$)

$$P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\ x = max\left(0, M-N+n\right),\ \ldots,\ min(M, n),\ 1 \le n < N,\ 1 \le M < N$$

$$E(X) = n\frac{M}{N} \qquad D(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$$

Normované normální rozdělení $N(0,1)$

$$U = \frac{X-\mu}{\sigma} \qquad -\infty < u < \infty \qquad E(U) = 0 \qquad D(U) = 1$$

$$\Phi(u) = 1-\Phi(-u) \qquad u_P = -u_{1-P}$$

Normální rozdělení $N(\mu,\sigma^2)$

$$-\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma^2 > 0 \qquad E(X) = \mu \qquad D(X) = \sigma^2$$

$$u = \frac{x-\mu}{\sigma} \qquad F(x) = \Phi(u) = \Phi\left(\frac{x-\mu}{\sigma}\right) \qquad x_P = \mu + \sigma\, u_P$$

$$P(x_1 < X \le x_2) = P\left(\frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{x_2-\mu}{\sigma}\right) = P(u_1 < U \le u_2) = \Phi(u_2) - \Phi(u_1)$$

Chí-kvadrát rozdělení $\chi^2(\nu)$ $\quad x > 0,\ \nu \in N$

Rozdělení t (*Studentovo*) $t(\nu)$ $\quad -\infty < x < \infty,\ \nu \in N$ $\quad t_P(\nu) = -t_{1-P}(\nu)$

F rozdělení (*Fisherovo – Snedecorovo*) $F(\nu_1,\nu_2)$ $\quad x > 0,\ \nu_1,\nu_2 \in N$

$$F_P(\nu_1,\nu_2) = \frac{1}{F_{1-P}(\nu_2,\nu_1)}$$

Matematická statistika

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Bodové a intervalové odhady parametrů (teoretické intervaly spolehlivosti)

střední hodnota $\hat{\mu} = \bar{X}$ *odhad* $N\mu = N\bar{X}$ *odhad* $E(X) = \bar{X}$

normální rozdělení

a) σ^2 známý

$$P\left(\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(\bar{X} - u_{1-\alpha}\frac{\sigma}{\sqrt{n}} < \mu\right) = 1-\alpha, \qquad P\left(\mu < \bar{X} + u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

b) σ^2 neznámý

$$P\left(\bar{X} - t_{1-\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{1-\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha \qquad T \sim t(n-1)$$

$$P\left(\bar{X} - t_{1-\alpha}\frac{S}{\sqrt{n}} < \mu\right) = 1-\alpha, \qquad P\left(\mu < \bar{X} + t_{1-\alpha}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

obecné rozdělení, σ^2 neznámý, velký výběr ($n > 30$)

$$P\left(\bar{X} - u_{1-\alpha/2}\frac{S}{\sqrt{n}} < E(X) < \bar{X} + u_{1-\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(\bar{X} - u_{1-\alpha}\frac{S}{\sqrt{n}} < E(X)\right) = 1-\alpha, \qquad P\left(E(X) < \bar{X} + u_{1-\alpha}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

rozptyl σ^2 (normální rozdělení) $\hat{\sigma}^2 = S^2$

Parametr π alternativního rozdělení (odhad relativní četnosti základního souboru)

$\hat{\pi} = P \qquad N\pi = NP$

$$P\left(P - u_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}} < \pi < P + u_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}}\right) = 1-\alpha$$

$$P\left(P - u_{1-\alpha}\sqrt{\frac{P(1-P)}{n}} < \pi\right) = 1-\alpha \qquad P\left(\pi < P + u_{1-\alpha}\sqrt{\frac{P(1-P)}{n}}\right) = 1-\alpha$$

Testování statistických hypotéz

Střední hodnota normálního rozdělení

H_0	H_1	Testové kritérium	Kritický obor
$\mu = \mu_0$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	σ^2 známý $U = \dfrac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$ $\quad U \sim N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u\rvert \geq u_{1-\alpha/2}\}$
		σ^2 neznámý $T = \dfrac{\bar{X} - \mu_0}{S}\sqrt{n}$ $\quad T \sim t(n-1)$	$W_\alpha = \{t;\ t \geq t_{1-\alpha}\}$ $W_\alpha = \{t;\ t \leq -t_{1-\alpha}\}$ $W_\alpha = \{t;\ \lvert t\rvert \geq t_{1-\alpha/2}\}$

Střední hodnota, obecné rozdělení, velký výběr

H_0	H_1	Testové kritérium	Kritický obor
$E(X) = \mu_0$	$E(X) > \mu_0$ $E(X) < \mu_0$ $E(X) \neq \mu_0$	σ^2 neznámý ($n > 30$) $U = \dfrac{\bar{X} - \mu_0}{S}\sqrt{n}$ $\quad U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u\rvert \geq u_{1-\alpha/2}\}$

Parametr π alternativního rozdělení (velké výběry)

H_0	H_1	Testové kritérium	Kritický obor
$\pi = \pi_0$	$\pi > \pi_0$ $\pi < \pi_0$ $\pi \neq \pi_0$	$U = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$ $\quad U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u\rvert \geq u_{1-\alpha/2}\}$

Rovnost středních hodnot dvou rozdělení

velké nezávislé výběry

H_0	H_1	Testové kritérium	Kritický obor
$E(X_1) = E(X_2)$ $E(X_1) - E(X_2) = 0$	$E(X_1) > E(X_2)$ $E(X_1) < E(X_2)$ $E(X_1) \neq E(X_2)$	σ_1^2 a σ_2^2 neznámé $U = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ $\quad U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u\rvert \geq u_{1-\alpha/2}\}$

závislé výběry z normálního rozdělení (párový *t*-test)

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$T = \dfrac{\sqrt{n}\bar{D}}{S_D}$ $\quad T \sim t(n-1)$ $D_i = X_{1i} - X_{2i},\ i = 1, 2, \ldots, n$	$W_\alpha = \{t;\ t \geq t_{1-\alpha}\}$ $W_\alpha = \{t;\ t \leq -t_{1-\alpha}\}$ $W_\alpha = \{t;\ \lvert t\rvert \geq t_{1-\alpha/2}\}$

Chí-kvadrát test dobré shody

H_0 a H_1	Testové kritérium	Kritický obor
H_0: $\pi_j = \pi_{0,j}$ $\quad j = 1, \ldots, k$ H_1: non H_0	$G = \sum_{j=1}^{k} \dfrac{(n_j - n\pi_{0,j})^2}{n\pi_{0,j}}$ $\quad G \approx \chi^2(k-1)$	$W_\alpha = \{g;\ g \geq \chi^2_{1-\alpha}\}$ $n\pi_{0,j} \geq 5$

Analýza závislostí

Kontingenční tabulka (r x s)

$$\sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} = \sum_{i=1}^{r} n_{i+} = \sum_{j=1}^{s} n_{+j} = n \qquad n'_{ij} = \frac{n_{i+}n_{+j}}{n} \quad n'_{ij} \geq 5$$

H_0	H_1	Testové kritérium	Kritický obor
znaky jsou nezávislé	non H_0	$G = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-n'_{ij})^2}{n'_{ij}} \quad G \approx \chi^2((r-1)(s-1))$	$W_\alpha = \{g;\ g \geq \chi^2_{1-\alpha}\}$

$$C = \sqrt{\frac{G}{G+n}} \qquad V = \sqrt{\frac{G}{n(m-1)}},\ m = \min(r, s)$$

Tabulka 2 x 2 $\qquad G = \dfrac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1+}n_{+1}n_{2+}n_{+2}}$

Analýza rozptylu

$$S_y = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(y_{ji} - \overline{y}\right)^2 = S_{y.m} + S_{y.v} \qquad S_{y.m} = \sum_{j=1}^{k}\left(\overline{y}_j - \overline{y}\right)^2 n_j \qquad S_{y.v} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(y_{ji} - \overline{y}_j\right)^2$$

$$P^2 = \frac{S_{y.m}}{S_y}$$

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2 = ... = \mu_k$	non H_0	$F = \dfrac{\frac{S_{y.m}}{k-1}}{\frac{S_{y.v}}{n-k}} \qquad F \sim F(k-1, n-k)$	$W_\alpha = \{F;\ F \geq F_{1-\alpha}\}$

Regrese a korelace

regresní přímka $y = \beta_0 + \beta_1 x + \varepsilon$,

$$Y = b_0 + b_1 x,\ \ \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \qquad \text{minimum}_{\beta_0,\beta_1} \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

$$s_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{n-1} = \frac{n}{n-1}(\overline{xy} - \overline{x}.\overline{y})$$

$$b_1 = \hat{\beta}_1 = \frac{n\sum y_i x_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\overline{xy} - \overline{x}.\overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{s_{xy}}{s_x^2}$$

$$b_0 = \hat{\beta}_0 = \frac{\sum y_i \sum x_i^2 - \sum y_i x_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2} = \overline{y} - \hat{\beta}_1\overline{x}$$

Jiné regresní funkce $\quad Y = b_0 + b_1 x + b_2 x^2, \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$

$$Y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k\ , \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_k x_k$$

$$S_y = \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 \qquad S_T = \sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2 \qquad S_R = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 = \sum_{i=1}^{n}\hat{\varepsilon}_i^2$$

$$S_y = S_T + S_R \qquad s_R^2 = \frac{S_R}{n-p} \qquad s_R = \sqrt{\frac{S_R}{n-p}} = \sqrt{s_R^2}$$

$$R^2 = I^2 = \frac{S_T}{S_y} \qquad R_{ADJ}^2 = I_{ADJ}^2 = 1-(1-R^2)\frac{n-1}{n-p-1}$$

Test hypotézy o regresním parametru

H_0	H_1	Testové kritérium	Kritický obor
$\beta_j = 0$	$\beta_j \neq 0$	$T = \dfrac{\hat{\beta}_j}{s_{\hat{\beta}_j}}$ $T \sim t(n-p)$	$W_\alpha = \{t;\ \lvert t\rvert \geq t_{1-\alpha/2}\}$

Test o modelu $p = k + 1$

H_0	H_1	Testové kritérium	Kritický obor
$\beta_0 = c$ $\beta_1 = 0$ ... $\beta_k = 0$	non H_0	$F = \dfrac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}}$ $F \sim F(p-1, n-p)$	$W_\alpha = \{F;\ F \geq F_{1-\alpha}\}$

korelační koeficient

$$r_{yx} = r_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\sqrt{\left(\overline{x^2} - \bar{x}^2\right)\left(\overline{y^2} - \bar{y}^2\right)}} = \frac{s_{xy}}{s_x s_y}$$

H_0	H_1	Testové kritérium	Kritický obor
$\rho_{XY} = 0$	$\rho_{XY} \neq 0$	$T = \dfrac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}}$ $T \sim t(n-2)$	$W_\alpha = \{t;\ \lvert t\rvert \geq t_{1-\alpha/2}\}$

Časové řady

$$\bar{y} = \frac{\sum_{t=1}^{T} y_t}{T} \qquad \bar{y} = \frac{\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + \ldots + \frac{y_{T-1} + y_T}{2}}{T-1} = \frac{\frac{1}{2} y_1 + \sum_{t=2}^{T-1} y_t + \frac{1}{2} y_T}{T-1}$$

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} d_1 + \frac{y_2 + y_3}{2} d_2 + \ldots + \frac{y_{T-1} + y_T}{2} d_{T-1}}{d_1 + d_2 + \ldots + d_{T-1}}$$

$$\Delta y_t = y_t - y_{t-1} \qquad \overline{\Delta y} = \frac{\sum_{t=2}^{T} \Delta y_t}{T-1} = \frac{y_T - y_1}{T-1}$$

$$k_t = \frac{y_t}{y_{t-1}} \qquad \bar{k} = \sqrt[T-1]{k_2 \cdot k_3 \cdot \ldots \cdot k_T} = \sqrt[T-1]{\frac{y_T}{y_1}} \qquad \delta_t = k_t - 1 \qquad \bar{\delta} = \bar{k} - 1$$

Dekompozice časové řady

$y_t = T_t + C_t + S_t + \varepsilon_t \qquad y_t = T_t \cdot C_t \cdot S_t \cdot \varepsilon_t$

Modelování trendu

Trendové funkce

$T_t = \beta_0 + \beta_1 t \qquad \hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t \qquad T_t = \beta_0 + \beta_1 t + \beta_2 t^2 \qquad \hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$

$$MSE = \frac{\sum_{t=1}^{T} (y_t - \hat{y}_t)^2}{T}$$

Klouzavé průměry

$$m = 2p + 1 \qquad \hat{T}_t = \bar{y}_t = \frac{y_{t-p} + \ldots + y_{t-1} + y_t + y_{t+1} + \ldots + y_{t+p}}{m}$$

$$m = 2p \qquad \hat{T}_t = \bar{y}_t = \frac{1}{2m}\left(y_{t-p} + 2y_{t-p+1} + \ldots + 2y_{t-1} + 2y_t + 2y_{t+1} + \ldots + 2y_{t+p-1} + y_{t+p}\right)$$

Modelování sezónnosti

Regresní metoda s umělými proměnnými (lineární trend, sezónnost délky 4)

$y_t = T_t + S_t + \varepsilon_t = \beta_0 + \beta_1 t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + \varepsilon_t$

$y_t^* = \beta_0^* + \beta_1^* t + \gamma_1^* D_{1t} + \gamma_2^* D_{2t} + \gamma_3^* D_{3t} + \varepsilon_t$

$$\hat{s} = \frac{\hat{\gamma}_1^* + \hat{\gamma}_2^* + \hat{\gamma}_3^*}{4} \qquad \hat{S}_j = \hat{\gamma}_j = \hat{\gamma}_j^* - \hat{s}, \quad j = 1, 2, 3 \qquad \hat{S}_4 = \hat{\gamma}_4 = -\hat{s} \qquad \hat{\beta}_0 = \hat{\beta}_0^* + \hat{s}$$

$$\sum_{j=1}^{s} \hat{S}_j = 0$$