Vysoká škola ekonomická
FAKULTA INFORMATIKY a STATISTIKY
Katedra statistiky a pravděpodobnosti

STATISTIKA

Vzorce pro 4ST201 a 4ST210

Verze: 4.5
Poslední změna: 27.3.2025

Popisná statistika

kvantily

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

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

$$\frac{x_{(z_P)} + x_{(z_P+1)}}{2}, \; nP = z_P$$

průměry

$i = 1, 2, \ldots, n$ prostý $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ vážený $w_i \ge 0$ $\bar{x}_w = \dfrac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$

$x_j,\ j = 1, 2, \ldots, k,$ četnosti n_j $\quad p_j = \dfrac{n_j}{n} \quad \sum_{j=1}^{k} n_j = n \quad \sum_{j=1}^{k} p_j = 1$

$$\bar{x} = \frac{\sum_{j=1}^{k} n_j x_j}{n} \qquad \bar{x} = \sum_{j=1}^{k} p_j x_j$$

$$\bar{x}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \qquad \bar{x}_H = \frac{n}{\sum_{j=1}^{k} \frac{n_j}{x_j}} \qquad \bar{x}_H = \frac{1}{\sum_{j=1}^{k} \frac{p_j}{x_j}}$$

$$\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}, \; x_i > 0$$

rozpětí $r = x_{(n)} - x_{(1)}, q = \tilde{x}_{0,75} - \tilde{x}_{0,25}$

Poznámka: Excel počítá kvantil jako (funkce PERCENTIL.INC())

$\lfloor h \rfloor$ je dolní celá část, $\lceil h \rceil$ je horní celá část

$\tilde{x}_P,\ 0 < P < 1 \quad \tilde{x}_P = x_{(\lfloor h \rfloor)} + (h - \lfloor h \rfloor)(x_{(\lceil h \rceil)} - x_{(\lfloor h \rfloor)})$

rozptyl

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

$$s_x^2 = \frac{\sum_{j=1}^{k} n_j (x_j - \overline{x})^2}{n} \qquad s_x^2 = \overline{x^2} - \overline{x}^2 = \frac{\sum_{j=1}^{k} n_j x_j^2}{n} - \left(\frac{\sum_{j=1}^{k} n_j x_j}{n}\right)^2$$

$$s_x^2 = \sum_{j=1}^{k} p_j (x_j - \overline{x})^2 \qquad s_x^2 = \overline{x^2} - \overline{x}^2 = \sum_{j=1}^{k} p_j x_j^2 - \left[\sum_{j=1}^{k} p_j x_j\right]^2$$

Směrodatná odchylka, variační koeficient, *MAD*

$$s_x = \sqrt{s_x^2} \qquad v_x = \frac{s_x}{\overline{x}},\ \overline{x} > 0 \qquad MAD = median(|x_i - \tilde{x}_{0,5}|)$$

Průměr ze skupinových průměrů, *k* skupin

$$p_j = \frac{n_j}{n} \qquad \sum_{j=1}^{k} n_j = n \qquad \sum_{j=1}^{k} p_j = 1 \qquad j = 1, 2, ..., k$$

$$\overline{x} = \frac{\sum_{j=1}^{k} n_j \overline{x}_j}{n} \qquad \overline{x} = \sum_{j=1}^{k} p_j \overline{x}_j$$

Rozklad rozptylu

$$s_x^2 = \overline{s^2} + s_{\overline{x}}^2 = \frac{\sum_{j=1}^{k} n_j s_{xj}^2}{n} + \frac{\sum_{j=1}^{k} n_j (\overline{x}_j - \overline{x})^2}{n}$$

$$s_x^2 = \overline{s^2} + s_{\overline{x}}^2 = \sum_{j=1}^{k} p_j s_{xj}^2 + \sum_{j=1}^{k} p_j (\overline{x}_j - \overline{x})^2$$

Transformace

$x_1, x_2, ..., x_n \to \overline{x},\ s_x^2,\ v_x,\ c$ konstanta : $\quad y_i = c + x_i,\ i = 1,2, ..., n \to \overline{y} = c + \overline{x},\ s_y^2 = s_x^2$

$z_i = cx_i,\ i = 1, 2, ..., n \to \overline{z} = c\overline{x},\ s_z^2 = c^2 s_x^2,\ v_z = v_x$

Pravděpodobnost a náhodné veličiny

Teorie pravděpodobnosti

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

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

Neslučitelné jevy $\quad P(A \cup B) = P(A) + P(B)$

Podmíněná pravděpodobnost $\quad P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

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

Nezávislé jevy $\quad P(A \cap B) = P(A)\,P(B), \;\; P(A \mid B) = P(A), \;\; P(B \mid A) = P(B)$

Závislé jevy $\quad P(A \cap B) = P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$

Náhodná veličina (náhodná veličina *X*; hodnota této náhodné veličiny *x*)
Diskrétní náhodné veličiny

$$P(x) = P(X = x) \qquad F(x) = P(X \leq x) = \sum_{x_i \leq x} P(x_i)$$

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

Střední hodnota $\quad E(X) = \sum_x xP(x)$

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

Směrodatná odchylka $\quad \sqrt{D(X)}$

Spojité náhodné veličiny

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

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

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

Střední hodnota $\quad E(X) = \int_{-\infty}^{\infty} xf(x)dx$

Rozptyl $\quad D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} xf(x)dx\right]^2$

Poznámka: v anglické literatuře a také někdy v české se používá *Var*(*X*) místo *D*(*X*).

Pravděpodobnostní rozdělení

Pravděpodobnostní rozdělení diskrétních náhodných veličin

Binomické rozdělení: $X \sim Bi(n,\pi)$, $n = 1,2, \ldots$

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

$$E(X) = n\pi \qquad D(X) = n\pi(1-\pi)$$

Pravděpodobnostní rozdělení spojitých náhodných veličin

Normované normální rozdělení: $U \sim N(0,1)$, $-\infty < u < \infty$

$E(U) = 0 \quad D(U) = 1$

$\Phi(u) = 1 - \Phi(-u), \qquad u_P = -u_{1-P},\ 0 < P < 1$

Normální rozdělení: $X \sim N(\mu,\sigma^2)$, $-\infty < x < \infty \qquad -\infty < \mu < \infty \qquad \sigma^2 > 0$

$E(X) = \mu \quad D(X) = \sigma^2$

$$U = \frac{X-\mu}{\sigma} \qquad u = \frac{x-\mu}{\sigma} \qquad F(x) = \Phi(u) = \Phi\left(\frac{x-\mu}{\sigma}\right), \qquad x_P = \mu + \sigma\, u_P,\ 0 < P < 1$$

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

Chí-kvadrát rozdělení: $G \sim \chi^2(\nu)$, $\nu = 1,2,\ldots$, $g > 0$

***t*- rozdělení (Studentovo rozdělení):** $T \sim t(\nu)$, $\nu = 1,2,\ldots$, $-\infty < t < \infty \quad t_P(\nu) = -t_{1-P}(\nu)$

***F*- rozdělení (Fisherovo – Snedecorovo rozdělení)**: $F \sim F(\nu_1,\nu_2)$, $\nu_1,\nu_2 = 1,2,\ldots$, $F > 0$

Poznámky: V anglické literatuře se používá *Z* místo *U*, rovněž kvantil se značí z_P místo u_P.

U *F*-rozdělení se i pro hodnoty používá velké *F*, nikoliv malé *f*.

Statistická indukce: bodový a intervalový odhad parametrů

výběrový rozptyl

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{n}{n-1}s_x^2 \quad S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{n}{n-1}}s_x$$

Bodové odhady parametrů

Střední hodnota (populační průměr) μ

$X \sim N(\mu,\sigma^2)$: $\hat{\mu} = \overline{x}$, směrodatná chyba $\frac{\hat{\sigma}}{\sqrt{n}} = \frac{S}{\sqrt{n}}$

Obecné rozdělení: $\widehat{E(X)} = \overline{x}$, směrodatná chyba $\frac{\hat{\sigma}}{\sqrt{n}} = \frac{S}{\sqrt{n}}$

Úhrn, N jednotek, $NE(X)$: $\widehat{NE(X)} = N\overline{x}$, směrodatná chyba $N\frac{\hat{\sigma}}{\sqrt{n}} = N\frac{S}{\sqrt{n}}$

(Populační) rozptyl σ^2

$X \sim N(\mu,\sigma^2)$: $\hat{\sigma}^2 = S^2$

Pravděpodobnost (populační podíl) π

$\hat{\pi} = p = \frac{m}{n}$, směrodatná chyba $\sqrt{\frac{p(1-p)}{n}}$

Úhrn (N jednotek) $N\pi$

$\widehat{N\pi} = N\hat{\pi} = Np = \frac{Nm}{n}$, směrodatná chyba $N \cdot \sqrt{\frac{p(1-p)}{n}}$

Statistická indukce: bodový a intervalový odhad parametrů

Intervalové odhady parametrů (spolehlivost intervalového odhadu $1-\alpha$)

Střední hodnota (populační průměr)

Normální rozdělení

dvoustranný interval spolehlivosti pro parametr μ:

$$\left(\bar{x}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}},\bar{x}+t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

jednostranné intervaly spolehlivosti:

dolní mez $\bar{x}-t_{1-\alpha}(n-1)\frac{S}{\sqrt{n}}$ horní mez $\bar{x}+t_{1-\alpha}(n-1)\frac{S}{\sqrt{n}}$

Úhrn, N jednotek:

dvoustranný interval spolehlivosti

$$N\left(\bar{x}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}},\ \bar{x}+t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)=\left(N\bar{x}-Nt_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}},\ N\bar{x}+Nt_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

jednostranné intervaly spolehlivosti:

dolní mez $N\left(\bar{x}-t_{1-\alpha}(n-1)\frac{S}{\sqrt{n}}\right)$ horní mez $N\left(\bar{x}+t_{1-\alpha}(n-1)\frac{S}{\sqrt{n}}\right)$

Obecné rozdělení, velký náhodný výběr

$$\left(\bar{x}-u_{1-\alpha/2}\frac{S}{\sqrt{n}},\ \bar{x}+u_{1-\alpha/2}\frac{S}{\sqrt{n}}\right)$$

jednostranné intervaly spolehlivosti:

dolní mez $\bar{x}-u_{1-\alpha}\frac{S}{\sqrt{n}}$ horní mez $\bar{x}+u_{1-\alpha}\frac{S}{\sqrt{n}}$

Úhrn, N jednotek:

dvoustranný interval spolehlivosti

$$N\left(\bar{x}-u_{1-\alpha/2}\frac{S}{\sqrt{n}},\ \bar{x}+u_{1-\alpha/2}\frac{S}{\sqrt{n}}\right)=\left(N\bar{x}-Nu_{1-\alpha/2}\frac{S}{\sqrt{n}},\ N\bar{x}+Nu_{1-\alpha/2}\frac{S}{\sqrt{n}}\right)$$

jednostranné intervaly spolehlivosti:

dolní mez $N\left(\bar{x}-u_{1-\alpha}\frac{S}{\sqrt{n}}\right)$ horní mez $N\left(\bar{x}+u_{1-\alpha}\frac{S}{\sqrt{n}}\right)$

Pravděpodobnost (populační podíl)

velký náhodný výběr

dvoustranný interval spolehlivosti:

$$\left(p - u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}\,, p + u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right)$$

jednostranné intervaly spolehlivosti:

dolní mez $p - u_{1-\alpha}\sqrt{\frac{p(1-p)}{n}}$ horní mez $p + u_{1-\alpha}\sqrt{\frac{p(1-p)}{n}}$

Úhrn (N jednotek) velký náhodný výběr

dvoustranný interval spolehlivosti:

$$N\left(p - u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}\,, p + u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right) = \left(Np - Nu_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}\,, Np + Nu_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right)$$

jednostranné intervaly spolehlivosti:

dolní mez $N\left(p - u_{1-\alpha}\sqrt{\frac{p(1-p)}{n}} \right)$ horní mez $N\left(p + u_{1-\alpha}\sqrt{\frac{p(1-p)}{n}} \right)$

Statistická indukce: testování hypotéz

Testování statistických hypotéz o parametrech

Střední hodnota (populační průměr); normální rozdělení

H_0	H_1	Testová statistika	Kritický obor
$\mu = \mu_0$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$T = \dfrac{\bar{x} - \mu_0}{S / \sqrt{n}}$ Pokud H_0 platí $T \sim t(n-1)$	$W_\alpha = \{t;\ t \geq t_{1-\alpha}\}$ $W_\alpha = \{t;\ t \leq -t_{1-\alpha}\}$ $W_\alpha = \{t;\ \lvert t\rvert \geq t_{1-\alpha/2}\}$

Střední hodnota (populační průměr); obecné rozdělení náhodné veličiny, velký náhodný výběr

H_0	H_1	Testová statistika	Kritický obor
$E(X) = E(X)_0$	$E(X) > E(X)_0$ $E(X) < E(X)_0$ $E(X) \neq E(X)_0$	$U = \dfrac{\bar{x} - E(X)_0}{S / \sqrt{n}}$ Pokud H_0 platí $U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u\rvert \geq u_{1-\alpha/2}\}$

Pravděpodobnost (populační podíl); velký náhodný výběr

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

Testování statistických hypotéz o pravděpodobnostním rozdělení v populaci

Chí-kvadrát (χ^2) test dobré shody (velký náhodný výběr, $n'_j = n\pi_{0,j}$, $n'_j \geq 5$)

H_0	H_1	Testová statistika	Kritický obor
$\pi_j = \pi_{0,j}$ $(j = 1,2, \ldots, k)$	non H_0	$G = \sum_{j=1}^{k} \dfrac{(n_j - n'_j)^2}{n'_j}$ Pokud H_0 platí $G \approx \chi^2(k-1)$	$W_\alpha = \{g;\ g \geq \chi^2_{1-\alpha}\}$

Analýza závislosti

Analýza rozptylu

rozklad čtverců: hodnoty faktoru (k možností),

počet pozorování n_j, $j = 1, 2, \ldots, k$; y_{ji}, $i = 1, 2, \ldots, n_j$ pozorování Y

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

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

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

Kontingenční tabulka (r x s)

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

Test nezávislosti v kontingenční tabulce

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

$$C = \sqrt{\frac{G}{n+G}} \in \left\langle 0, \sqrt{\frac{m-1}{m}} \right\rangle \qquad V = \sqrt{\frac{G}{n(m-1)}}, \; m = \min(r, s)$$

$$r = s = 2 \qquad G = n \frac{(n_{11} n_{22} - n_{12} n_{21})^2}{n_{1.} n_{2.} n_{.1} n_{.2}}$$

Regrese a korelace

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

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

$$e_i = y_i - b_0 - b_1 x_i = y_i - Y_i = y_i - \hat{y}_i$$

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

$$\hat{\beta}_1 = b_1 = b_{yx} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\overline{xy} - \bar{x}.\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{s_{xy}}{s_x^2}$$

$$\hat{\beta}_0 = b_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \bar{y} - b_{yx}\bar{x}$$

Jiné regresní funkce

Regresní parabola

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \rightarrow Y = b_0 + b_1 x + b_2 x^2, \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$$

Vícenásobná regrese

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \rightarrow Y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k, \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k$$

Rozklad čtverců

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

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

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

Test hypotézy o regresních parametrech

H_0 H_1	Testové kritérium	Kritický obor
$\beta_j = 0$ $\beta_j \neq 0$	$T = \frac{b_j}{s(b_j)}$ Pokud H_0 platí $T \sim t(n-p)$	$W_\alpha = \{t; \lvert t \rvert \geq t_{1-\alpha/2}\}$

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	$F = \dfrac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}}$ Pokud H_0 platí $F \sim F(p-1, n-p)$	$W_\alpha = \{F; F \geq F_{1-\alpha}\}$

Korelační koeficienty

Pearsonův korelační koeficient

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

Spearmanův pořadový korelační koeficient

$r_s = 1 - \dfrac{6\sum_{i=1}^{n}(\text{pořadí } x_i - \text{pořadí } y_i)^2}{n(n^2-1)}$, případně pomocí Pearsonova korelačního koeficientu proměnných pořadí x_i a pořadí y_i

Test o nulovosti korelačního koeficientu

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

Analýza časových řad

Průměrné hodnoty časové řady období $t = 1,\ 2,\ \ldots,\ n$

$$\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n} \qquad \bar{y} = \frac{\frac{1}{2} y_1 + \sum_{t=2}^{n-1} y_t + \frac{1}{2} y_n}{n-1} \qquad \bar{y} = \frac{\frac{y_1 + y_2}{2} d_1 + \frac{y_2 + y_3}{2} d_2 + \ \ldots\ + \frac{y_{n-1} + y_n}{2} d_{n-1}}{d_1 + d_2 + \ \ldots\ + d_{n-1}}$$

$$\Delta_t = y_t - y_{t-1} \qquad \bar{\Delta} = \frac{1}{n-1} \sum_{t=2}^{n} \Delta_t = \frac{y_n - y_1}{n-1}$$

Řetězové a bazické indexy

$$k_t = I_{t/t-1} = \frac{y_t}{y_{t-1}} \qquad \bar{k} = \sqrt[n-1]{k_2 k_3 \ldots k_n} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

$$I_{t/1} = \frac{y_t}{y_1} = I_{2/1}\ I_{3/2} \ldots I_{t/t-1} \qquad I_{t/t-1} = \frac{y_t}{y_{t-1}} = \frac{I_{t/1}}{I_{(t-1)/1}}$$

Klouzavé průměry

$m = 2p + 1$ $$\bar{y}_t = \frac{\sum_{i=-p}^{p} y_{t+i}}{m} = \frac{y_{t-p} + \ldots + y_{t-1} + y_t + y_{t+1} + \ldots + y_{t+p}}{m}$$

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

Dekompozice časové řady

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

$$T_t = \beta_0 + \beta_1 t \qquad \hat{T}_t = b_0 + b_1 t$$

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 \qquad \hat{T}_t = b_0 + b_1 t + b_2 t^2$$

$$e_t = y_t - \hat{y}_t$$

$$MSE = \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n} = \frac{\sum_{t=1}^{n} e_t^2}{n} \qquad D = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

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

$$y_t = T_t + S_t + \varepsilon_t = \beta_0 + \beta_1 t + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \alpha_3 x_{3t} + \varepsilon_t$$

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\alpha}_1 x_{1t} + \hat{\alpha}_2 x_{2t} + \hat{\alpha}_3 x_{3t} = b_0 + b_1 t + a_1 x_{1t} + a_2 x_{2t} + a_3 x_{3t}$$

$$\bar{a} = \frac{a_1 + a_2 + a_3}{4} \qquad S_{i+4j} = a_i - \bar{a},\ i = 1,2,3 \qquad S_{4+4j} = -\bar{a} \qquad \hat{T}_t = (b_0 + \bar{a}) + b_1 t$$

Indexní analýza

$Q = pq \qquad IQ = Ip \cdot Iq$

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

Složené individuální indexy

$$I(\Sigma q) = \frac{\sum q_1}{\sum q_0} = \frac{\sum Iq.q_0}{\sum q_0} = \frac{\sum q_1}{\sum \frac{q_1}{Iq}} \qquad \Delta(\Sigma q) = \sum q_1 - \sum q_0$$

$$I(\Sigma Q) = \frac{\sum Q_1}{\sum Q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum IQ.Q_0}{\sum Q_0} = \frac{\sum Q_1}{\sum \frac{Q_1}{IQ}} \qquad \Delta(\Sigma Q) = \sum Q_1 - \sum Q_0$$

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

Souhrnné indexy

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

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