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

STATISTIKA

VZORCE

verze 2.0
poslední aktualizace: 15.11.2006

Popisná statistika

Rozdělení četností

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

Kvantily $\tilde{x}_p$ $\qquad n \cdot \frac{p}{100} < m_p < n \cdot \frac{p}{100} + 1$

Průměry

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

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

$$\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} \qquad \bar{x}_G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_k^{n_k}}$$

Rozpětí $\qquad R = x_{\max} - x_{\min} \qquad R_Q = \tilde{x}_{75} - \tilde{x}_{25}$

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

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

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

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

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

směrodatná odchylka $\qquad s_x = \sqrt{s_x^2} \qquad$ variační koeficient $\qquad v_x = \frac{s_x}{\bar{x}}$

Pravděpodobnost

Počet pravděpodobnosti

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

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

$P(A \cap B) = P(A)\,P(B)$, $\quad P(A \cap B) = P(A)\,P(B|A) = P(B)\,P(A|B)$

$$P(A) = \sum_{i=1}^{s} P(A \cap B_i) \qquad P(A) = \sum_{i=1}^{s} P(B_i)\; P(A|B_i)$$

$$P(B_i|A) = \frac{P(B_i)\; P(A|B_i)}{P(A)}$$

Náhodné veličiny

$$P(x) = P(X = x) \qquad F(x) = P(X \le x_0) = \sum_{x \le x_0} P(x)$$

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

$$F(x) = P(X \le x_0) = \int_{-\infty}^{x_0} f(x)dx \qquad f(x) = F'(x) \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

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

$x_P \quad F(x_P) = P$

$$E(X) = \sum_x x\; P(x) \qquad E(X) = \int_{-\infty}^{\infty} x\; f(x)dx$$

$$D(X) = \sum_x x^2\; P(x) - \left[\sum_x x\; P(x)\right]^2 \qquad D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} xf(x)dx\right]^2$$

$$\sigma = \sigma(X) = \sqrt{D(X)}$$

$$U = \frac{X - E(X)}{\sqrt{D(X)}}$$

Pravděpodobnostní rozdělení

Alternativní rozdělení $\quad A(\pi)$

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

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

Binomické rozdělení $\quad Bi(n,\pi)$

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

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

Poissonovo rozdělení $Po(\lambda)$

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

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

Hypergeometrické rozdělení $Hy(N,M,n)$

$$P(x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}, \qquad x = max(0, M\text{-}N\text{+}n), \dots, min(M, n), n > 0, N \geq n, M \leq N$$

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

Rovnoměrné rozdělení $R(a,b)$

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & jinak \end{cases} \qquad F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$

$$E(X) = \frac{a+b}{2} \qquad D(X) = \frac{(b-a)^2}{12}$$

Exponenciální rozdělení $E(A,\delta)$ $A \geq 0, \delta > 0$

$$f(x) = \begin{cases} \frac{1}{\delta} e^{-\frac{(x-A)}{\delta}} & x > A \\ 0 & \text{jinak} \end{cases} \qquad F(x) = \begin{cases} 0 & x \leq A \\ 1 - e^{-\frac{(x-A)}{\delta}} & x > A \end{cases}$$

$$x_P = A - \delta\, ln(1-P) \qquad E(X) = A + \delta \qquad D(X) = \delta^2$$

Normované normální rozdělení $N(0,1)$

$$U = \frac{X-\mu}{\sigma} \qquad \Phi(u) = 1 - \Phi(-u) \qquad u_P = -u_{1-P} \quad E(X) = 0 \quad D(X) = 1$$

Normální rozdělení $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \quad E(X) = \mu \qquad D(X) = \sigma^2$$

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

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

Logaritmicko-normální rozdělení $LN(\mu, \sigma^2)$

$$U = \frac{\ln X - \mu}{\sigma} \sim N(0,1) \qquad x > 0,\ -\infty < \mu < \infty,\ \sigma^2 > 0$$

$$x_P = \exp(\mu + \sigma\, u_P) \qquad E(X) = e^{\mu + \sigma^2/2} \qquad D(X) = e^{2\mu + \sigma^2}[e^{\sigma^2} - 1]$$

$$\mu = E(\ln X) = \ln(E(X)) - \sigma^2/2 \qquad \sigma^2 = D(\ln X) = \ln\left(\frac{D(X)}{(E(X))^2} + 1\right)$$

Chí-kvadrát rozdělení $\chi^2(\nu)$

$$E(\chi^2) = \nu \qquad D(\chi^2) = 2\nu$$

Rozdělení t (Studentovo) $t(\nu)$

$$E(t) = 0 \qquad t_P = -\, t_{1-P}$$

F - rozdělení (Fisherovo – Snedecorovo) $F(\nu_1, \nu_2)$

$$F = \frac{\chi_1^2}{\nu_1} / \frac{\chi_2^2}{\nu_2} \qquad F_P(\nu_1, \nu_2) = \frac{1}{F_{1-P}(\nu_2, \nu_1)}$$

Centrální limitní věty

Moivre - Laplaceova věta

a/ $X \sim Bi(n,\pi) \approx N(n\pi,\ n\pi(1 - \pi)),\quad n\pi(1 - \pi) > 9 \qquad U = \dfrac{X - n\pi}{\sqrt{n\pi(1-\pi)}} \approx N(0,1)$

b/ $p = \dfrac{X}{n} \approx N\left(\pi,\ \dfrac{\pi(1-\pi)}{n}\right) \qquad U = \dfrac{p - \pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}} \approx N(0,1)$

Lindebergova - Lévyho věta

a/ $E(X_i) = \mu,\quad D(X_i) = \sigma^2$

$$\sum X_i \approx N(n\mu,\ n\sigma^2) \qquad U = \frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}} \approx N(0,1)$$

b/ $\bar{X} \approx N\left(\mu, \dfrac{\sigma^2}{n}\right) \qquad U = \dfrac{\bar{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \approx \mathrm{N}(0,1)$

Matematická statistika

$$s_x' = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n\text{-}1}}$$

Odhady parametrů

střední hodnota $\hat{\mu} = \bar{x}$

normální rozdělení

a) σ^2 známé

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

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

$$n \geq \frac{u_{1-\alpha/2}^2\sigma^2}{\Delta^2}$$

b) σ^2 neznámé

$$P\left(\bar{x} - t_{1-\alpha/2}\frac{s_x'}{\sqrt{n}} < \mu < \bar{x} + t_{1-\alpha/2}\frac{s_x'}{\sqrt{n}}\right) = 1-\alpha \qquad t \sim t(n\text{ - }1)$$

$$P\left(\bar{x} - t_{1-\alpha}\frac{s_x'}{\sqrt{n}} < \mu\right) = 1-\alpha \qquad P\left(\mu < \bar{x} + t_{1-\alpha}\frac{s_x'}{\sqrt{n}}\right) = 1-\alpha$$

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

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

$$P\left(\bar{x} - u_{1-\alpha}\frac{s_x'}{\sqrt{n}} < \mu\right) = 1-\alpha \quad P\left(\mu < \bar{x} + u_{1-\alpha/2}\frac{s_x'}{\sqrt{n}}\right) = 1-\alpha$$

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

$$P\left(\frac{(n-1)s_x'^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s_x'^2}{\chi^2_{\alpha/2}}\right) = 1-\alpha$$

$$P\left(\frac{(n-1)s_x'^2}{\chi^2_{1-\alpha}} < \sigma^2\right) = 1-\alpha \quad P\left(\sigma^2 < \frac{(n-1)s'^2}{\chi^2_{\alpha}}\right) = 1-\alpha$$

Parametr π alternativního rozdělení (odhad relativní četnosti základního souboru) $\hat{\pi} = p$

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

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

$$n \geq \frac{u_{1-\alpha/2}^2\pi(1-\pi)}{\Delta^2}, \quad n \geq 0{,}25\frac{u_{1-\alpha/2}^2}{\Delta^2}$$

Testování 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 známé $U = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}$ $U \sim N(0,1)$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$
		σ^2 neznámé $t = \frac{\bar{x} - \mu_0}{s'_x}\sqrt{n}$ $t \sim t(n - 1)$	$W_\alpha=\{t \geq t_{1-\alpha}\}$ $W_\alpha=\{t \leq -t_{1-\alpha}\}$ $W_\alpha=\{\lvert t\rvert \geq t_{1-\alpha/2}\}$

Střední hodnota, obecné rozdělení, velký výběr

H_0	H_1	Testové kritérium	Kritický obor
$\mu = \mu_0$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	σ^2 neznámé ($n > 30$) $U = \frac{\bar{x} - \mu_0}{s'_x}\sqrt{n}$ $U \approx N(0,1)$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$

Rozptyl v normálním rozdělení

H_0	H_1	Testové kritérium	Kritický obor
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{(n-1)s'^2_x}{\sigma_0^2}$ $\chi^2 \sim \chi^2(n-1)$	$W_\alpha=\{\chi^2 \geq \chi^2_{1-\alpha}\}$ $W_\alpha=\{\chi^2 \leq \chi^2_{\alpha}\}$ $W_\alpha=\{\chi^2 \leq \chi^2_{\alpha/2} \cup \chi^2 \geq \chi^2_{1-\alpha/2}\}$

Parametr π alternativnho rozdělení (velké výběry n > 9/π(1–π))

H_0	H_1	Testové kritérium	Kritický obor
$\pi = \pi_0$	$\pi > \pi_0$ $\pi < \pi_0$ $\pi \neq \pi_0$	$U = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ $U \sim N(0,1)$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$

Rovnost středních hodnot dvou rozdělení

normální rozdělení (nezávislé náhodné výběry z normálního rozdělení)

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	a) σ_1^2, σ_2^2 známé $U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ $U \sim N(0,1)$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$
		σ_1^2, σ_2^2 neznámé, ale předpokládáme, že $\sigma_1^2 = \sigma_2^2$ $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{(n_1-1)s'^2_1 + (n_2-1)s'^2_2}{n_1+n_2-2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $t \sim t(n_1 + n_2 - 2)$	$W_\alpha=\{t \geq t_{1-\alpha}\}$ $W_\alpha=\{t \leq -t_{1-\alpha}\}$ $W_\alpha=\{\lvert t\rvert \geq t_{1-\alpha/2}\}$

	σ_1^2, σ_2^2 neznámé, ale předpokládáme, že $\sigma_1^2 \neq \sigma_2^2$ $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1'^2}{n_1} + \dfrac{s_2'^2}{n_2}}}$ $\quad t \sim t(\nu)$ $\nu = \dfrac{\left(\dfrac{s_1'^2}{n_1} + \dfrac{s_2'^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{s_1'^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{s_2'^2}{n_2}\right)^2}$	$W_\alpha=\{t \geq t_{1-\alpha}\}$ $W_\alpha=\{t \leq -t_{1-\alpha}\}$ $W_\alpha=\{\lvert t\rvert \geq t_{1-\alpha/2}\}$

velké nezávislé výběry

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	σ_1^2, σ_2^2 neznámé $U = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1'^2}{n_1} + \dfrac{s_2'^2}{n_2}}}$ $\quad U \approx N(0,1)$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$

závislé výběry z normálního rozdělení (párový t-test)

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$t = \dfrac{\bar{d}}{s_d}\sqrt{n-1}$ $\quad t \sim t(n-1)$ $d_i = x_{1i} - x_{2i},\ i=1,2,..,n$	$W_\alpha=\{t \geq t_{1-\alpha}\}$ $W_\alpha=\{t \leq -t_{1-\alpha}\}$ $W_\alpha=\{\lvert t\rvert \geq t_{1-\alpha/2}\}$

Rovnost rozptylů dvou normálních rozdělení

H_0	H_1	Testové kritérium	Kritický obor
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$F = \dfrac{s_1'^2}{s_2'^2}$ $\quad F \sim F(n_1 - 1, n_2 - 1)$	$W_\alpha=\{F \geq F_{1-\alpha}\}$ $W_\alpha=\{F \leq F_\alpha\}$ $W_\alpha=\{F \leq F_{\alpha/2} \cup F \geq F_{1-\alpha/2}\}$

Rovnost parametrů dvou alternativních rozdělení (velké výběry)

H_0	H_1	Testové kritérium	Kritický obor
$\pi_1 = \pi_2$	$\pi_1 > \pi_2$ $\pi_1 < \pi_2$ $\pi_1 \neq \pi_2$	$U = \dfrac{p_1 - p_2}{p(1-p)\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $\quad U \approx N(0,1)$ $p = \dfrac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$

Chí-kvadrát test dobré shody

H_0 a H_1	Testové kritérium	Kritický obor
H_0: $\pi_i = \pi_{0,i}$ $\quad i = 1, .., k$ H_1: non H_0	$\chi^2 = \sum_{i=1}^{k} \dfrac{(n_i - n\pi_{0,i})^2}{n\pi_{0,i}}$ $\quad \chi^2 \sim \chi^2(k\text{-}1)$ p odhadnutých parametrů $\quad \chi^2 \sim \chi^2(k - p - 1)$	$W_\alpha=\{\chi^2 \geq \chi^2_{1-\alpha}\}$

Analýza závislostí

χ^2 test nezávislosti v kontingenční tabulce (*r* x *s*)

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

H$_0$	H$_1$	Testové kritérium	Kritický obor
$\pi_{ij} = \pi_{i.}\,\pi_{.j}$ $1 \leq i \leq r$ $1 \leq j \leq s$	non H$_0$	$G = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n}$ $\quad G \sim \chi^2((r-1)(s-1))$	$W_\alpha = \{G \geq \chi^2_{1-\alpha}\}$

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

Tabulka 2 x 2 $\qquad r_{ab} = \dfrac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{n_{1.}n_{2.}n_{.1}n_{.2}}}$

Analýza rozptylu

celkový součet čtverců $S_y = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2 = S_{y.m} + S_{y.v}$

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

H$_0$	H$_1$	Testové kritérium	Kritický obor
$\mu_1 = \mu_2 = .. = \mu_k$	non H_0	$F = \dfrac{\dfrac{\sum_{i=1}^{k}(\bar{y}_i - \bar{y})^2 n_i}{k-1}}{\dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2}{n-k}} = \dfrac{\dfrac{S_{y,m}}{k-1}}{\dfrac{S_{y,v}}{n-k}}$ $\quad F \sim F(k-1, n-k)$	$W_\alpha = \{F \geq F_{1-\alpha}\}$

Regrese a korelace $\qquad y = \eta(\boldsymbol{x}, \beta_0, \beta_1, ..., \beta_{p-1}) + \varepsilon \qquad Y_i = \eta(\boldsymbol{x_i}, b_0, b_1, ..., b_{p-1})$

Regresní přímka $\quad y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2 = \min.$

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

$$b_1 = b_{yx} = \frac{\begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum y_i x_i \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}} = \frac{n\sum y_i x_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{s_{xy}}{s_x^2}$$

$$b_0 = \frac{\begin{vmatrix} \sum y_i & \sum x_i \\ \sum y_i x_i & \sum x_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}} = \frac{\sum y_i \sum x_i^2 - \sum y_i x_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2} = \bar{y} - b_{yx}\bar{x}$$

$$S_x = \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad s_{b_0} = s\sqrt{\frac{\overline{x^2}}{S_x}} \qquad s_{b_1} = s\sqrt{\frac{1}{S_x}}$$

$$P(b_0 - t_{1-\alpha/2}\, s_{b_0} < \beta_0 < b_0 + t_{1-\alpha/2}\, s_{b_0}) = 1 - \alpha$$

$$P(b_1 - t_{1-\alpha/2}\, s_{b_1} < \beta_1 < b_1 + t_{1-\alpha/2}\, s_{b_1}) = 1 - \alpha \qquad t \sim t(n-2)$$

Jiné regresní funkce

$$Y = b_0 + b_1 x + b_2 x^2 \qquad Y = b_0 b_1^{\,x}$$

$$Y = b_0 + b_1 x^{-1} \qquad Y = b_0 x^{b_1}$$

$$Y = b_0 + b_1 \ln(x) \qquad Y = b_0 + b_1 x_1 + b_2 x_2 + .. + b_k x_k$$

celkový součet čtverců $S_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$ teoretický součet čtverců $S_T = \sum_{i=1}^{n} (Y_i - \bar{y})^2$

$$s_y^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2 = \frac{S_y}{n} \qquad s_Y^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{y})^2 = \frac{S_T}{n}$$

reziduální součet čtverců $S_R = \sum_{i=1}^{n} (y_i - Y_i)^2 = \sum_{i=1}^{n} e_i^2$

$$s_{y-Y}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - Y_i)^2 = \frac{S_R}{n} \qquad s_R^2 = \frac{S_R}{n-p}$$

$$S_y = S_R + S_T \qquad s_y^2 = s_Y^2 + s_{y-Y}^2$$

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

Test hypotézy o regresních parametrech

H_0	H_1	Testové kritérium	Kritický obor
$\beta_i = 0$	$\beta_i \neq 0$	$t = \frac{b_i}{s_{b_i}}$ $\quad t \sim t(n-p)$	$W_\alpha = \{\lvert t\rvert > t_{1-\alpha/2}\}$

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	$F = \frac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}}$ $\quad F \sim F(p-1, n-p)$	$W_\alpha = \{F \geq F_{1-\alpha}\}$

korelační koeficient

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

$$b_{xy} b_{yx} = r_{xy}^2 \qquad r_{yx} = b_{yx}\frac{s_x}{s_y} = b_{xy}\frac{s_y}{s_x}$$

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

Spearmanův korelační koeficient

$$r_S = 1 - \frac{6\sum_{i=1}^{n}(i_x - i_y)^2}{n(n^2-1)}$$

H_0	H_1	Testové kritérium	Kritický obor
$\rho_S = 0$	$\rho_S \neq 0$	$t = \frac{r_S}{\sqrt{1-r_S^2}}\sqrt{n-2}$ $t \sim t(n-2)$	$W_\alpha = \{\lvert t\rvert \geq t_{1-\alpha/2}\}$

dílčí korelační koeficient

$$r_{yx_1.x_2} = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{(1-r_{yx_2}^2)(1-r_{x_1x_2}^2)}$$

H_0	H_1	Testové kritérium	Kritický obor
$\rho_{yx_1.x_2} = 0$	$\rho_{yx_1.x_2} \neq 0$	$t = \frac{r_{yx_1.x_2}\sqrt{n-3}}{\sqrt{1-r_{yx_1.x_2}^2}}$ $t \sim t(n-3)$	$W_\alpha = \{\lvert t\rvert \geq t_{1-\alpha/2}\}$

vícenásobný korelační koeficient

$$r_{y.x_1x_2} = \sqrt{\frac{r_{yx_1}^2 - 2r_{yx_1} r_{yx_2} r_{x_1x_2} + r_{yx_2}^2}{1 - r_{x_1x_2}^2}}$$

Časové řady

$$\bar{y} = \frac{\sum_{t=1}^{T} y_t}{T}$$

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

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

$$\Delta y_t = y_t - y_{t-1} \qquad \bar{\Delta} = \frac{(y_2 - y_1) + (y_3 - y_2) + \dots + (y_T - y_{T-1})}{T-1} = \frac{y_T - y_1}{T-1}$$

$$\delta_t = \frac{\Delta y_t}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1$$

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

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

Dekompozice časové řady

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

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

$T_t = \beta_0 \qquad T_t = \beta_0 + \beta_1 t + \beta_1 t^2$

$T_t = \beta_0 t^{\beta_1} \qquad \ln T_t = \ln\beta_0 + \beta_1 \ln t$

$\mathrm{T_t} = \beta_0 \beta_1^t \qquad \ln T_t = \ln\beta_0 + t \ln\beta_1$

$T_t = \gamma + \beta_0 \beta_1^t \qquad T_t = \dfrac{\gamma}{1 + \alpha\beta^t}$

$$\mathrm{ME} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{T}_t) \qquad \mathrm{MSE} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{T}_i)^2 \qquad \mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}\left|y_t - \hat{T}_t\right|$$

Exponenciální vyrovnávání (jednoduché) $\qquad Y_t = \alpha y_t + (1 - \alpha)Y_{t-1}$

Klouzavé průměry

$m = 2p + 1$

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

$m = 2p$

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

Analýza sezónní složky (sezónnost délky r, počet období m)

1.

$\hat{y}_t$ klouzavé průměry

$$\hat{SI}_t = \frac{y_t}{\hat{y}_t} \qquad \bar{s}_j = 1 + \bar{c}_j = \frac{\sum_{k=1}^{m-1} \hat{SI}_{jk}}{m-1} \qquad \hat{s}_j = \frac{r}{\sum_{i=1}^{r} \bar{s}_i} \bar{s}_j$$

2. Regresní metoda s umělými proměnnými (lineární trend)

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

$$\bar{a} = \frac{a_1 + \ldots + a_{r-1}}{r} \qquad S_{r+rj} = -\bar{a} \qquad S_{i+rj} = a_i - \bar{a} \qquad i=1,.., r-1, j=1, .., m$$

$$\hat{T}_t = (b_0 + \bar{a}) + b_1 t$$

Durbinův Watsonův test

H_0	H_1	Testové kritérium
$\rho = 0$	$\rho \neq 0$	$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$

$$r_k = \frac{\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

Indexní analýza

$$Q = pq$$

$$i_p = \frac{p_1}{p_0} \qquad \Delta p = p_1 - p_0 \qquad i_q = \frac{q_1}{q_0} \qquad \Delta q = q_1 - q_0 \qquad i_Q = \frac{Q_1}{Q_0} \qquad \Delta Q = Q_1 - Q_0$$

$$I_q = \frac{\sum q_1}{\sum q_0} \qquad \Delta q = \sum q_1 - \sum q_0 \qquad I_q = \frac{\sum q_1}{\sum q_0} = \frac{\sum i_q q_0}{\sum q_0} \qquad I_q = \frac{\sum q_1}{\sum \frac{q_1}{i_q}}$$

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

$$I_p = \frac{\overline{p}_1}{\overline{p}_0} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_1}{\sum q_1}} \cdot \frac{\frac{\sum p_0 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} \ , \qquad I_p = \frac{\overline{p}_1}{\overline{p}_0} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{\sum p_1 q_0}{\sum q_0}}{\frac{\sum p_0 q_0}{\sum q_0}} \cdot \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_1 q_0}{\sum q_0}}$$

$$_L I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0} \qquad _L I_p = \frac{\sum \frac{p_1}{p_0} p_0 q_0}{\sum p_0 q_0} = \sum i_p w_0 \qquad w_0 = \frac{p_0 q_0}{\sum p_0 q_0} = \frac{Q_0}{\sum Q_0}$$

$$_P I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} \qquad _P I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{\frac{p_1}{p_0}}} = \frac{\sum Q_1}{\sum \frac{Q_1}{i_p}}$$

$$_F I_p = \sqrt{_L I_p \cdot {_P I_p}}$$

$$_L I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} \qquad _P I_q = \frac{\sum p_1 q_1}{\sum p_1 q_0} \qquad _F I_q = \sqrt{_L I_q \cdot {_P I_q}}$$

$$I_Q = \frac{\sum Q_1}{\sum Q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \qquad \Delta Q = \sum p_1 q_1 - \sum p_0 q_0$$

$$I_Q = {_L I_p} \cdot {_P I_q} = {_P I_p} \cdot {_L I_q}$$