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

STATISTIKA

VZORCE PRO 4ST201

verze 2.1
poslední aktualizace: 22.9.2008

Popisná statistika

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

$\tilde{x}_p$

$$n \cdot \frac{P}{100} < z_P < n \cdot \frac{P}{100} + 1 \qquad n\ p < z_p < n\ p + 1$$

$$n \frac{P}{100} + 0,5 = z_p \qquad np + 0,5 = z_p$$

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

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

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

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

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

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

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

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

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

$$s_x = \sqrt{s_x^2} \qquad V_x = \frac{s_x}{\overline{x}}$$

Pravděpodobnost

Počet pravděpodobnosti

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

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

$$P(A \cap B) = P(A)\ P(B) \qquad P(A \cap B) = P(A)\ P(B|A) = P(B)\ P(A|B)$$

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

Náhodné veličiny

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

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

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

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

$$x_P \qquad F(x_P) = P \qquad x_P = F^{-1}(P)$$

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

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

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

$X_i\ i = 1,..,n,\quad E(X_i) = \mu,\ D(X_i) = \sigma^2$ nezávislé $\quad E(\sum X_i) = n\mu, \quad D(\sum X_i) = n\sigma^2$

$$E(\frac{1}{n}\sum X_i) = \mu, \quad D(\frac{1}{n}\sum X_i) = \frac{\sigma^2}{n}$$

Pravděpodobnostní rozdělení

Alternativní rozdělení $A[\pi]$

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

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

Binomické rozdělení $Bi[n;\pi]$

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

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

Poissonovo rozdělení $Po[\lambda]$

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

Hypergeometrické rozdělení $Hy[N;M;n]$

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

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

Exponenciální rozdělení $E[A;\delta]$ $A \geq 0,\ \delta > 0$

$$F(x)=\begin{cases}0 & x\leq A\\ 1-e^{-\frac{(x-A)}{\delta}} & x>A\end{cases}\qquad E(X) = A + \delta \qquad D(X) = \delta^2$$

Normální rozdělení $N[\mu;\sigma^2]$

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

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

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

Normované normální rozdělení $N[0;1]$

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

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

Logaritmicko-normální rozdělení $LN[\mu;\sigma^2]$

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

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

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

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

Chí-kvadrát rozdělení $\chi^2[\nu]$ $x > 0$

Rozdělení t (*Studentovo*) $t[\nu]$ $-\infty < x < \infty$ $t_P = -t_{1-P}$

F - rozdělení (*Fisherovo – Snedecorovo*) $F[\nu_1;\nu_2]$ $x > 0,$ $F_P(\nu_1,\nu_2)=\frac{1}{F_{1-P}(\nu_2,\nu_1)}$

Matematická statistika

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

Odhady parametrů

střední hodnota est $\mu = \hat{\mu} = \overline{x}$ est $N\mu = N\overline{x}$

normální rozdělení

a) σ^2 známé

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

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

b) σ^2 neznámé

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

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

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

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

$$P\left(\overline{x} - u_{1-\alpha}\frac{s_x'}{\sqrt{n}} < E(X)\right) = 1-\alpha \quad P\left(E(X) < \overline{x} + u_{1-\alpha}\frac{s_x'}{\sqrt{n}}\right) = 1-\alpha$$

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

Parametr π alternativního rozdělení (odhad relativní četnosti základního souboru)

est $\pi = \hat{\pi} = p$ est $N\pi = Np$

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

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

Testování 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 $U = \dfrac{\bar{x} - \mu_0}{\sigma}\sqrt{n}$ $\quad U \sim N[0;1]$	$W_\alpha = \{U \geq u_{1-\alpha}\}$ $W_\alpha = \{U \leq -u_{1-\alpha}\}$ $W_\alpha = \{\lvert U\rvert \geq u_{1-\alpha/2}\}$
		σ^2 neznámé $t = \dfrac{\bar{x} - \mu_0}{s'_x}\sqrt{n}$ $\quad t \sim t[n-1]$	$W_\alpha = \{t \geq t_{1-\alpha}\}$ $W_\alpha = \{t \leq -t_{1-\alpha}\}$ $W_\alpha = \{\lvert t\rvert \geq t_{1-\alpha/2}\}$

Střední hodnota, obecné rozdělení, velký výběr

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

Rozptyl v normálním rozdělení

H_0	H_1	Testové kritérium	Kritický obor
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 = \dfrac{(n-1)s_x'^2}{\sigma_0^2}$ $\quad \chi^2 \sim \chi^2[n\text{-}1]$	$W_\alpha = \{\chi^2 \geq \chi^2_{1-\alpha}\}$ $W_\alpha = \{\chi^2 \leq \chi^2_{\alpha}\}$ $W_\alpha = \{\chi^2 \leq \chi^2_{\alpha/2} \cup \chi^2 \geq \chi^2_{1-\alpha/2}\}$

Parametr π alternativního rozdělení (velké výběry)

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

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 $U = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $\quad U \sim N[0;1]$	$W_\alpha = \{U \geq u_{1-\alpha}\}$ $W_\alpha = \{U \leq -u_{1-\alpha}\}$ $W_\alpha = \{\lvert U\rvert \geq u_{1-\alpha/2}\}$
		σ_1^2, σ_2^2 neznámé, ale předpokládáme, že $\sigma_1^2 = \sigma_2^2$ $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\dfrac{(n_1-1)s_1'^2 + (n_2-1)s_2'^2}{n_1 + n_2 - 2}\right)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $t \sim t[n_1 + n_2 - 2]$	$W_\alpha = \{t \geq t_{1-\alpha}\}$ $W_\alpha = \{t \leq -t_{1-\alpha}\}$ $W_\alpha = \{\lvert t\rvert \geq t_{1-\alpha/2}\}$

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

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 $$U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1'^2}{n_1} + \frac{s_2'^2}{n_2}}} \qquad U \approx N[0;1]$$	$W_\alpha=\{U \geq u_{1-\alpha}\}$ $W_\alpha=\{U \leq -u_{1-\alpha}\}$ $W_\alpha=\{\lvert U\rvert \geq u_{1-\alpha/2}\}$

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

H_0	H_1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$$t = \frac{\bar{d}}{s_d'}\sqrt{n} \qquad t \sim t[n-1]$$ $d_i = x_{1i} - x_{2i}$, $i=1,2,..,n$	$W_\alpha=\{t \geq t_{1-\alpha}\}$ $W_\alpha=\{t \leq -t_{1-\alpha}\}$ $W_\alpha=\{\lvert t\rvert \geq t_{1-\alpha/2}\}$

Rovnost rozptylů dvou normálních rozdělení

H_0	H_1	Testové kritérium	Kritický obor
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$$F = \frac{s_1'^2}{s_2'^2} \qquad F \sim F[n_1 - 1;\, n_2 - 1]$$	$W_\alpha=\{F \geq F_{1-\alpha}\}$ $W_\alpha=\{F \leq F_\alpha\}$ $W_\alpha=\{F \leq F_{\alpha/2} \cup F \geq F_{1-\alpha/2}\}$

Chí-kvadrát test dobré shody

H_0 a H_1	Testové kritérium	Kritický obor
H_0: $\pi_i = \pi_{0,i}$ $i = 1, .., k$ H_1: non H_0	$$G = \sum_{i=1}^{k} \frac{(n_i - n\pi_{0,i})^2}{n\pi_{0,i}} \qquad G \approx \chi^2[k-1]$$	$W_\alpha=\{G \geq \chi^2_{1-\alpha}\}$ $n\pi_{0,i} \geq 5$

Analýza závislostí

Kontingenční tabulka (r x s)

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

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	$G = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$ $\quad G \approx \chi^2[(r-1)(s-1)]$	$W_\alpha = \{G \geq \chi^2_{1-\alpha}\}$

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

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

Analýza rozptylu

$$S_y = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \overline{y})^2 = S_{y.m} + S_{y.v} \qquad S_{y.m} = \sum_{i=1}^{k}(\overline{y}_i - \overline{y})^2 n_i \qquad S_{y.v} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2$$

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

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

Regrese a korelace

regresní přímka $y = \beta_0 + \beta_1 x + \varepsilon$, $Y = b_0 + b_1 x$ $\quad$ $\text{minimum}_{b_0, b_1} \sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2$

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

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

$$b_0 = \frac{\sum y_i \sum x_i^2 - \sum y_i x_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2} = \overline{y} - b_{yx}\overline{x}$$

Jiné regresní funkce $Y = b_0 + b_1 x + b_2 x^2$ $Y = b_0 + b_1 x_1 + b_2 x_2 + .. + b_k x_k$

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

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

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

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

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

Test hypotézy o regresních parametrech

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

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	$F = \dfrac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}}$	$F \sim F[p-1; n-p]$	$W_\alpha = \{F \geq F_{1-\alpha}\}$

korelační koeficient

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

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

Časové řady

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

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

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

Klouzavé průměry

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

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

Dekompozice časové řady

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

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

$T_t = \beta_0 \beta_1^t$ $\qquad$ $\ln T_t = \ln(\beta_0) + \ln(\beta_1)\, t$ $\qquad$ $\ln(\hat{T}_t) = \ln(b_0) + \ln(b_1)\, t$

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

Analýza sezónní složky

1. Metoda empirických indexů (délka sezónnosti r)

$\bar{y}_t$ klouzavé průměry délky r (například $r = 4$) $\qquad$ sezónní index$_t = \frac{y_t}{\bar{y}_t}$

$$\text{průměrný sezónní index}_i = \frac{\sum_{t \text{ z } i-\text{té sezóny}} \text{sezonní index}_t}{\text{počet hodnot z } i-\text{té sezóny}} \qquad i = 1, 2, ., r \text{ (např. } r = 4)$$

$$\text{standardizovaný sezónní index} = \frac{r}{\sum_{j=1}^{r} \text{průměrný sezonní index}_j} \,.\, \text{průměrný sezonní index}_i$$

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

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

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

Indexní analýza

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

$Q = pq$

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

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

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

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

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

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

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

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

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