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

STATISTIKA

VZORCE PRO 4ST201 a 4ST210

verze 2018
poslední aktualizace: 1. 9. 2018

DODATEČNÉ ÚPRAVY:

STATISTICKY
NEKLASICKY

©KSTP 2018

Popisná statistika

pravděpodobnost $p_i = \dfrac{n_i}{n}$ $\qquad \sum_{i=1}^{k} n_i = n$ suma četností (abs.) $\qquad \sum_{i=1}^{k} p_i = 1$ suma všech rel. četností musí být 1. $\qquad i = 1, 2, \ldots, k$

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

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

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

prostý aritmetický průměr $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ – suma hodnot

$\bar{x} = \dfrac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i}$ vážený arit. průměr

průměr pro rel. četnosti $\bar{x} = \sum_{i=1}^{k} x_i p_i$

harmonický průměr $\bar{x}_H = \dfrac{n}{\sum_{i=1}^{n} \dfrac{1}{x_i}}$

$\bar{x}_H = \dfrac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \dfrac{n_i}{x_i}}$ harmonický průměr vážený tvar

harmonický průměr rel. četnosti $\bar{x}_H = \dfrac{1}{\sum_{i=1}^{k} \dfrac{p_i}{x_i}}$

geometrický průměr $\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$ (počet: n)

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

prostý rozptyl $s_x^2 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$

prostý rozptyl (výpočtový tvar) $s_x^2 = \overline{x^2} - \bar{x}^2 = \dfrac{\sum_{i=1}^{n} x_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} x_i}{n} \right)^2$ umocněný průměr

vážený rozptyl $s_x^2 = \dfrac{\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i}{\sum_{i=1}^{k} n_i}$

vážený rozptyl (výpočtový tvar) $s_x^2 = \overline{x^2} - \bar{x}^2 = \dfrac{\sum_{i=1}^{k} x_i^2 n_i}{\sum_{i=1}^{k} n_i} - \left(\dfrac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i} \right)^2$

rozptyl rel. četnosti $s_x^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i$

$s_x^2 = \overline{x^2} - \bar{x}^2 = \sum_{i=1}^{k} x_i^2 p_i - \left(\sum_{i=1}^{k} x_i p_i \right)^2$ rozptyl rel. četnosti (výpočtový tvar)

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

$\bar{x} = \dfrac{\sum_{i=1}^{k} \bar{x}_i n_i}{\sum_{i=1}^{k} n_i}$ – "průměr průměrů"

rozklad rozptylu (rel. četnosti) $s_x^2 = \sum_{i=1}^{k} s_{ix}^2 p_i + \sum_{i=1}^{k} (\bar{x}_i - \bar{x})^2 p_i$

$\bar{x} = \sum_{i=1}^{k} \bar{x}_i p_i$ – "průměr průměrů" (rel. četnosti)

směrodatná odchylka $s_x = \sqrt{s_x^2}$

variační koeficient $v_x = \dfrac{s_x}{\bar{x}}$

STATISTICKY NEKLASICKY

Pravděpodobnost

Počet pravděpodobnosti

doplněk pravděpodobnosti (jev opačný)

m = příznivé jevy
n = celkový počet jevů
pravděpodobnost

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

$$P(A \cap B) \leq \min(P(A), P(B))$$ ← průnik ZJ

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

↖ sjednocení slučitelných jevů

$$P(A \cap B) = P(A)\,P(B)$$ ← průnik NJ

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

↑ sjednocení neslučitelných jevů

ZJ = závislé jevy
NJ = nezávislé jevy

Náhodné veličiny

nespojité náh. veličiny

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

↑ pravděpodobnostní funkce ↑ distribuční funkce

střední hodnota

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

rozptyl

spojité náh. veličiny

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

↖ hustota pravděpodobnosti ↖ distribuční funkce

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

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

střední hodnota ↘ rozptyl ↓ horní hranice, dolní hranice, směrodatná odchylka

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

Pravděpodobnostní rozdělení

Alternativní rozdělení $A(\pi)$

pravděpodobnostní funkce ↗

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

π = konstantní pravděpodobnost

střední hodnota ↗ $E(X) = \pi$ ↗ $D(X) = \pi(1-\pi)$ rozptyl

Binomické rozdělení $Bi(n,\pi)$

pravděpodobnostní funkce →

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

střední hodnota ↗ $E(X) = n\pi$ $\quad D(X) = n\pi(1-\pi)$ rozptyl

Poissonovo rozdělení $Po(\lambda)$

pravděpodobnostní funkce →

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

střední hodnota ↗ $E(X) = \lambda$ ↗ $D(X) = \lambda$ rozptyl

STATISTICKY NEKLASICKY

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

pravděpodobnostní funkce

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

střední hodnota

$$E(X) = n\frac{M}{N}$$

rozptyl

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

$\boxed{\textit{Normované}}$ *normální rozdělení* $N(0,1)$

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

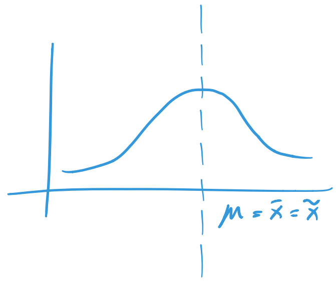
střední hodnota, rozptyl

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

Normální rozdělení $N(\mu, \sigma^2)$

střední hodnota, rozptyl

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



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

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

používáme tabulky

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

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

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

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

STATISTICKY NEKLASICKY

Matematická statistika

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

Bodové a intervalové odhady parametrů (teoretické intervaly spolehlivosti)

střední hodnota $\hat{\mu}=\bar{X}$ $\quad \hat{N\mu}=N\bar{X}$

normální rozdělení

a) σ^2 známý

směrodatná odchylka $\qquad$ hladina významnosti

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

oboustranný interval spolehlivosti

průměr $\quad$ kvantil

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

Levostranný int. spolehlivosti $\qquad$ Pravostranný int. spolehlivosti

b) σ^2 neznámý

Počet stupňů volnosti

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

oboustranný interval spolehlivosti

průměr $\quad$ Kvantil student. rozdělení

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

Levostranný int. spolehlivosti $\qquad$ Pravostranný int. spolehlivosti

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

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

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

Levostranný int. spolehlivosti $\qquad$ Pravostranný int. spolehlivosti

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

Parametr π alternativního rozdělení (odhad relativní četnosti základního souboru)

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

podíl $\quad$ Kvantil

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

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

Levostranný int. spolehlivosti $\qquad$ Pravostranný int. spolehlivosti

STATISTICKY NEKLASICKY

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 průměr — předpoklad normálního rozdělení $U = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$ $\quad U \sim N(0,1)$ σ – směr. odchylka	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ Kvantil $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u \rvert \geq u_{1-\alpha/2}\}$
hypotézy		σ^2 neznámý $T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$ $\quad T \sim t(n-1)$	$W_\alpha = \{t;\ t \geq t_{1-\alpha}\}$ $W_\alpha = \{t;\ t \leq -t_{1-\alpha}\}$ $W_\alpha = \{t;\ \lvert t \rvert \geq t_{1-\alpha/2}\}$

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 ($n > 30$) $U = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$ $\quad U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u \rvert \geq u_{1-\alpha/2}\}$

Parametr π alternativního rozdělení (velké výběry)

H$_0$	H$_1$	Testové kritérium Podíl	Kritický obor
$\pi = \pi_0$	$\pi > \pi_0$ $\pi < \pi_0$ $\pi \neq \pi_0$	$U = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ $\quad U \approx N(0,1)$	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u \rvert \geq u_{1-\alpha/2}\}$

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 průměr $U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ $\quad U \approx N(0,1)$ rozptyl, počet	$W_\alpha = \{u;\ u \geq u_{1-\alpha}\}$ $W_\alpha = \{u;\ u \leq -u_{1-\alpha}\}$ $W_\alpha = \{u;\ \lvert u \rvert \geq u_{1-\alpha/2}\}$

závislé výběry z normálního rozdělení (párový t-test)

H$_0$	H$_1$	Testové kritérium	Kritický obor
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$T = \frac{\sqrt{n}\bar{D}}{S_D}$ průměrná diference (Používáme Excel) $T \sim t(n-1)$ Počet stupňů volnosti $D_i = X_{1i} - X_{2i},\ i = 1, 2, \ldots, n$	$W_\alpha = \{t;\ t \geq t_{1-\alpha}\}$ $W_\alpha = \{t;\ t \leq -t_{1-\alpha}\}$ $W_\alpha = \{t;\ \lvert t \rvert \geq t_{1-\alpha/2}\}$

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$_0$	$G = \sum_{j=1}^{k} \frac{(n_j - n\pi_{0,j})^2}{n\pi_{0,j}}$ $\quad G \approx \chi^2(k-1)$ celkový počet, očekávaný podíl, Počet stupňů volnosti	$W_\alpha = \{g;\ g \geq \chi^2_{1-\alpha}\}$ $n\pi_j \geq 5$

STATISTICKY NEKLASICKY

Analýza závislostí

Kontingenční tabulka (r x s)

$$\sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} = \sum_{i=1}^{r} n_{i+} = \sum_{j=1}^{s} n_{+j} = n \qquad n'_{ij} = \frac{\underbrace{n_{i+}}_{\text{suma } i\text{-tého řádku}}\underbrace{n_{+j}}_{\text{suma } j\text{-tého sloupce}}}{n} \qquad n'_{ij} \geq 5$$

H0	H1	Testové kritérium	Kritický obor
znaky jsou nezávislé	non H0	$G = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n_{i+}n_{+j}/n)^2}{n_{i+}n_{+j}/n}$ $G \approx \chi^2(\underbrace{(r-1)}_{\text{počet řádků}}\underbrace{(s-1)}_{\text{počet sloupců}})$ stupně volnosti	$W_\alpha = \{g;\ g \geq \chi^2_{1-\alpha}\}$

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

Pearsonův koeficient kontingence   Cramerův koeficient kontingence

Tabulka 2 · 2 $\qquad G = \dfrac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1+}n_{+1}n_{2+}n_{+2}}$ Zjednodušený vzorec pro tabulky 2x2

Analýza rozptylu

meziskupinový součet čtverců

$$\underbrace{S_y}_{\text{celkový součet čtverců}} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ji} - \bar{y})^2 = S_{y.m} + \underbrace{S_{y.v}}_{\text{vnitroskupinový součet čtverců}} \qquad S_{y.m} = \sum_{j=1}^{k}(\bar{y}_j - \bar{y})^2 n_j \qquad S_{y.v} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ji} - \bar{y}_j)^2$$

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

H0	H1	Testové kritérium	Kritický obor
$\mu_1 = \mu_2 = \ldots = \mu_k$ rovnost středních hodnot	non H0	počet skupin; celkový počet; $F = \dfrac{\frac{S_{y.m}}{k-1}}{\frac{S_{y.v}}{n-k}}$ $F \sim F(k-1, n-k)$ stupně volnosti	$W_\alpha = \{F;\ F \geq F_{1-\alpha}\}$

Regrese a korelace

konstanta   směrnice

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

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

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

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

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

parabola

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

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

vícenásobná reg. funkce

STATISTICKY NEKLASICKY

$$S_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

celkový součet čtverců

$$S_T = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

Teoretický součet čtverců

$$S_R = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \hat{\varepsilon}_i^2$$

reziduální součet čtverců

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

$$R^2 = I^2 = \frac{S_T}{S_y}$$

Koeficient/Index determinace

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

upravený koeficient/Index determinace

Test hypotézy o regresním parametru

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

Test o modelu $\quad p = k + 1$

H_0	H_1	Testové kritérium	Kritický obor
$\beta_0 = c$ $\beta_1 = 0$ … $\beta_k = 0$	non H_0	$F = \dfrac{\dfrac{S_T}{p-1}}{\dfrac{S_R}{n-p}}$ Teoretický s. čtverců, reziduální součet čtverců; počet stupňů volnosti $F \sim F(p-1, n-p)$	$W_\alpha = \{F;\ F \geq F_{1-\alpha}\}$ Fisherovo rozdělení

korelační koeficient

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

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

test významnosti korel. koeficientu

STATISTICKY NEKLASICKY

Časové řady

aritmetický průměr

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

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

chronologický průměr

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

chronologický vážený průměr

míry dynamiky

absolutní přírůstek

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

průměrný absolutní přírůstek

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

koeficient růstu

$$k_t = \frac{y_t}{y_{t-1}}$$

průměrný koeficient růstu

$$\bar{k} = \sqrt[T-1]{k_2 \cdot k_3 \cdot \ldots \cdot k_T} = \sqrt[T-1]{\frac{y_T}{y_1}}$$

$$\delta_t = k_t - 1$$

relativní přírůstek

$$\bar{\delta} = \bar{k} - 1$$

průměrný relativní přírůstek

Dekompozice časové řady

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

$y_t = T_t \cdot C_t \cdot S_t \cdot \varepsilon_t$ multiplikativní

Modelování trendu

Trendové funkce

$T_t = \beta_0 + \beta_1 t$ přímka $\qquad \hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t$

parabola

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

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

průměrná čtvercová chyba

Klouzavé průměry

prosté $m = 2p + 1$

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

centrované $m = 2p$

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

Modelování sezónnosti

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

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

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

$$\hat{s} = \frac{\hat{\gamma}_1^* + \hat{\gamma}_2^* + \hat{\gamma}_3^*}{4}$$

$$\hat{S}_j = \hat{\gamma}_j = \hat{\gamma}_j^* - \hat{s}, \quad j = 1, 2, 3$$

j-tá sez. složka

$$\hat{S}_4 = \hat{\gamma}_4 = -\hat{s}$$

čtvrtá sezónní složka

$$\hat{\beta}_0 = \hat{\beta}_0^* + \hat{s}$$

upravená konstanta

$$\sum_{j=1}^{s} \hat{S}_j = 0$$ součet sez. složek = 0

STATISTICKY NEKLASICKY

Indexní analýza

Individuální indexy jednoduché

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

cena — Q tržba — q množství

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

cenový index — množstevní index — index tržby

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

Bázický index — Řetězový index

Individuální indexy složené

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

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

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

Souhrnné indexy

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

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

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

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

STATISTICKY NEKLASICKY