

Vysoká škola ekonomická
FAKULTA INFORMATIKY a STATISTIKY
Katedra statistiky a pravděpodobnosti

STATISTIKA

Vzorce pro 4ST201 a 4ST210

Verze: 4.2
Poslední změna: 16.9.2024

©KSTP, VŠE 2024

Popisná statistika

kvantily

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

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

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

průměry

$$i = 1, 2, \dots, n \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$i = 1, 2, \dots, k, \text{ četnosti } n_i \quad p_i = \frac{n_i}{n} \quad \sum_{i=1}^k n_i = n \quad \sum_{i=1}^k p_i = 1$$

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

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

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

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

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

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

Poznámka: Excel počítá kvantil jako (funkce PERCENTILE.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

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2$$
$$s_x^2 = \frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{n} \quad s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^k n_i x_i^2}{n} - \left(\frac{\sum_{i=1}^k n_i x_i}{n} \right)^2$$
$$s_x^2 = \sum_{i=1}^k p_i (x_i - \bar{x})^2 \quad s_x^2 = \overline{x^2} - \bar{x}^2 = \sum_{i=1}^k p_i x_i^2 - \left[\sum_{i=1}^k p_i x_i \right]^2$$

Směrodatná odchylka, variační koeficient, *mad*

$$s_x = \sqrt{s_x^2} \quad v_x = \frac{s_x}{\bar{x}}, \quad \bar{x} > 0 \quad mad = median(|x_i - \tilde{x}|)$$

Průměr ze skupinových průměrů, *k* skupin

$$p_j = \frac{n_j}{n} \quad \sum_{j=1}^k n_j = n \quad \sum_{j=1}^k p_j = 1 \quad j = 1, 2, \dots, k$$
$$\bar{x} = \frac{\sum_{j=1}^k n_j \bar{x}_j}{n} \quad \bar{x} = \sum_{j=1}^k p_j \bar{x}_j$$

Pravděpodobnost a náhodné veličiny

Teorie pravděpodobnosti

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

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

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

Podmíněná pravděpodobnost $P(A|B) = \frac{P(A \cap B)}{P(B)}$

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

Nezávislé jevy $P(A \cap B) = P(A)P(B)$, $P(A|B) = P(A)$, $P(B|A) = P(B)$

Závislé jevy $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

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

Diskrétní náhodné veličiny

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

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

Střední hodnota $E(X) = \sum_x xP(x)$

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

Směrodatná odchylka $\sqrt{D(X)}$

Spojitě náhodné veličiny

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

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

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

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

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

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: $X \sim Bi(n, \pi)$

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

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

Pravděpodobnostní rozdělení spojitých náhodných veličin

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

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

$$\Phi(u) = 1 - \Phi(-u), \quad u_p = -u_{1-p}$$

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

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

$$U = \frac{X - \mu}{\sigma} \quad u = \frac{x - \mu}{\sigma} \quad F(x) = \Phi(u) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad x_p = \mu + \sigma u_p$$

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

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

t- rozdělení (Studentovo rozdělení): $T \sim t(\nu), \nu = 1, 2, \dots, -\infty < t < \infty \quad t_p(\nu) = -t_{1-p}(\nu)$

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

Poznámka: V anglické literatuře se používá Z místo U , rovněž kvantil se značí z_p místo u_p .

Statistická indukce: bodový a intervalový odhad parametrů

výběrový rozptyl

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

Bodové odhady populačních parametrů

Populační průměr (střední hodnota) μ

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

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

(Populační) rozptyl σ^2

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

Populační podíl (pravděpodobnost) π

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

Populační úhrn $N \cdot \pi$

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

Statistická indukce: bodový a intervalový odhad parametrů

Intervalové odhady populačních parametrů

(spolehlivost odhadu $1 - \alpha$)

Populační průměr (střední hodnota)

Normální rozdělení (v populaci)

dvoustranný interval spolehlivosti pro parametr μ :

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

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

Obecné rozdělení (v populaci), velký náhodný výběr

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

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

Populační podíl (pravděpodobnost)

Velký náhodný výběr

dvoustranný interval spolehlivosti:

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

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

Populační úhrn

Velký náhodný výběr

dvoustranný interval spolehlivosti:

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

$$\text{dolní mez} \quad Np - Nu_{1-\alpha} \sqrt{\frac{p(1-p)}{n}} \quad \text{horní mez} \quad Np + Nu_{1-\alpha} \sqrt{\frac{p(1-p)}{n}}$$

Statistická indukce: testování hypotéz

Testování statistických hypotéz o populačních parametrech

Populační průměr (střední hodnota); normální rozdělení v populaci

H ₀	H ₁	Testová statistika	Kritický obor
$\mu = \mu_0$	$\mu > \mu_0$	$T = \frac{\bar{x} - \mu_0}{S / \sqrt{n}}$ Pokud H ₀ platí $T \sim t(n-1)$	$W_\alpha = \{t; t \geq t_{1-\alpha}\}$
	$\mu < \mu_0$		$W_\alpha = \{t; t \leq -t_{1-\alpha}\}$
	$\mu \neq \mu_0$		$W_\alpha = \{t; t \geq t_{1-\alpha/2}\}$

Populační průměr (střední hodnota); obecné rozdělení náhodné veličiny X, velký náhodný výběr

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

Populační podíl (pravděpodobnost); předpoklad dostatečně velkého náhodného výběru

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

Testování statistických hypotéz o pravděpodobnostním rozdělení v populaci

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

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

Analýza závislosti

Rozklad rozptylu (popisná statistika)

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

Analýza rozptylu

rozklad čtverců ($j = 1, 2, \dots, k$ hodnoty faktoru (k možností), $y_{ji}, i = 1, 2, \dots, n_j$ pozorování Y)

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

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

H ₀	H ₁	Testové kritérium	Kritický obor
$\mu_1 = \mu_2 = \dots = \mu_k$	non H ₀	$F = \frac{\frac{S_{y.m}}{k-1}}{\frac{S_{y.v}}{n-k}}$ Pokud H ₀ platí $F \sim F(k-1, n-k)$	$W_\alpha = \{F; F \geq F_{1-\alpha}\}$

Kontingenční tabulka ($r \times s$)

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

Test nezávislosti v kontingenční tabulce

H ₀	H ₁	Testové kritérium	Kritický obor
znaky jsou nezávislé	non H ₀	$G = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$ Pokud H ₀ platí $G \approx \chi^2((r-1)(s-1))$	$W_\alpha = \{g; g \geq \chi^2_{1-\alpha}\}$

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

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

Regrese a korelace

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

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

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

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

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

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

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

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

Rozklad čtverců

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

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

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

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 $T \sim t(n-p)$	$W_\alpha = \{t; t \geq t_{1-\alpha/2}\}$

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	$F = \frac{\frac{S_T}{p-1}}{\frac{S_R}{n-p}}$ Pokud H ₀ platí $F \sim F(p-1, n-p)$	$W_\alpha = \{F; F \geq F_{1-\alpha}\}$

Korelační koeficienty

Pearsonův korelační koeficient

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

Spearmanův pořadový korelační koeficient

$$r_s = 1 - \frac{6 \sum_{i=1}^n (\text{pořadí } x_i - \text{pořadí } y_i)^2}{n(n^2 - 1)}$$

Test o nulovosti korelačního koeficientu

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

Analyzá časových řad

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

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

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

Řetězové a bazické indexy

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

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

Klouzavé průměry

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

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

Dekompozice časové řady

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

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

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

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

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

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

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

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

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

Vybrané funkce v Excelu

<i>Co chci spočítat?</i>	<i>Excelová funkce</i>	<i>Symbol</i>
Četnost kategorií kvalitativní proměnné	COUNTIF()	n_j
Četnost hodnot kvantitativní proměnné	ČETNOSTI()	n_j
Součet (celkem)	SUMA()	Σ
Kvantil resp. Percentil	PERCENTIL.INC()	\tilde{x}_p
Medián	MEDIAN()	\tilde{x}
Modus	MODE.MULT()	\hat{x}
Jednoduchý aritmetický průměr	PRŮMĚR()	\bar{x}
Rozptyl populační	VAR.P()	s_x^2
Směrodatná odchylka populační	SMODCH.P()	s_x
Výběrový rozptyl	VAR.S()	s^2
Výběrová směrodatná odchylka	SMODCH.VÝBĚR.S()	s
Pearsonův korelační koeficient	CORREL()	r
Pravděpodobnostní funkce, X má binomické rozdělení	BINOM.DIST()	$P(x)$
Distribuční funkce, U má normované normální rozdělení	NORM.S.DIST()	$\Phi(u)$
Kvantily pro náhodnou veličinu U , která má normované normální rozdělení	NORM.S.INV()	u_p
Distribuční funkce, X má normální rozdělení	NORM.DIST()	$F(x)$
Kvantily pro náhodnou veličinu X , která má normální rozdělení	NORM.INV()	x_p
Kvantily pro náhodnou veličinu T , která má t -rozdělení	T.INV()	t_p
Kvantily pro náhodnou veličinu G , která má rozdělení χ^2 rozdělení	CHISQ.INV()	χ_p^2
Kvantily pro náhodnou veličinu F , která má F-rozdělení	F.INV()	F_p
p-hodnota pro χ^2 test nezávislosti v kontingenční tabulce	CHISQ.TEST()	p-hodnota