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

STATISTIKA

VZORCE PRO 4ST231

verze 1.1
poslední aktualizace: 25.9.2008

Popisná statistika

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

$\tilde{x}_p$

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

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

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

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

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

$$\bar{x}_K = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \qquad \bar{x}_K = \sqrt{\frac{\sum_{i=1}^{k} x_i^2 n_i}{\sum_{i=1}^{k} n_i}} \qquad \bar{x}_K = \sqrt{\sum_{i=1}^{k} x_i^2\, p_i}$$

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

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

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

Popisná statistika

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

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

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

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

Analýza závislostí

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

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

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

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

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

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

Jiné regresní funkce

$$Y = b_0 + b_1 x + b_2 x^2 \qquad Y = b_0 + b_1 \frac{1}{x} \qquad Y = b_0 + b_1 \log x$$

$$Y = b_0 + b_1 x_1 + b_2 x_2 + .. + b_k x_k$$

$$Y = b_o b_1^x$$

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

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

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

$$I_{yx}^2 = \frac{S_Y}{S_y} = 1 - \frac{S_{(y-Y)}}{S_y} \qquad I_{yx} = \sqrt{I_{yx}^2}$$

$$r_{i_x i_y} = 1 - \frac{6\sum \left(i_x - i_y\right)^2}{n\left(n^2 - 1\right)}$$

Časové řady

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

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

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

$$\Delta_t^2 = \Delta_t^1 - \Delta_{t-1}^1$$

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

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

Trendová složka

$m = 2p + 1$

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

$m = 2p$

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

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

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

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

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

Sezónní složka

Regresní metoda s umělými proměnnými (lineární trend, délka sezónní periody r)

$$y_{it} = T_{it} + S_{it} + \varepsilon_{it} = \beta_0 + \beta_1 t + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \dots + \alpha_{r-1} x_{r-1t} \quad i = 1, 2, \dots, r$$ ***r*** délka sezónní periody

$$t = 1, 2, \dots, n$$

$$Y_{it} = \hat{T}_{it} + \hat{S}_{ti} = b_0 + b_1 t + a_1 x_{1t} + a_2 x_{2t} + \dots + a_{r-1} x_{r-1t}$$

Časové řady

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

$$\hat{S}_{it} = \hat{S}_i \qquad \sum_{i=1}^{r} \hat{S}_i = 0$$

$$\hat{S}_i = a_i - \bar{a} \qquad i = 1, 2, \ldots, r-1$$

$$\hat{S}_r = -\bar{a}$$

$$\bar{a} = \frac{\sum_{i=1}^{r-1} a_i}{r}$$

Metoda empirických indexů (délka sezónnosti r)

$$y_{it} = T_{it} C_{it} S_{it} \varepsilon_{it}$$

$$\hat{S}_{it} = \widehat{SI}_i \qquad \sum_{i=1}^{r} \widehat{SI}_i = r$$

$$\widehat{SI}_i = \overline{SI}_i \frac{r}{\sum_{i=1}^{r} \overline{SI}_i} \quad \text{kde} \quad \overline{SI}_i = \frac{\text{Součet } SI_{it} \text{ pro i-tou sezonu}}{\text{Počet } SI_{it} \text{ pro i-tou sezonu}}$$

$$SI_{it} = \frac{y_{it}}{\bar{y}_t}, \qquad \bar{y}_t \text{ je klouzavý průměr pro } m = r$$

Indexní analýza

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

$$Q = pq$$

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

$$\sum Q = \bar{p} \cdot \sum q$$

$$I\left(\sum Q\right) = I\bar{p} \cdot I\left(\sum q\right)$$

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

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

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

$$\bar{p} = p \cdot \frac{q}{\sum q}$$

$$I_{\bar{p}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\frac{\sum p_1 q_0}{\sum q_0}}{\frac{\sum p_0 q_0}{\sum q_0}} \cdot \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_1 q_0}{\sum q_0}} \quad \text{nebo} \quad I_{\bar{p}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\frac{\sum p_0 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} \cdot \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_1}{\sum q_1}}$$

$$I_{SS}^{\left(\frac{q_0}{\sum q_0}\right)} \cdot I_{STR}^{(p_1)} \qquad I_{\bar{p}} = I_{STR}^{(p_0)} \cdot I_{SS}^{\left(\frac{q_1}{\sum q_1}\right)}$$

Indexní analýza

$$\sum Q = \sum p \cdot q$$

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

$$\Delta(\Sigma Q) = \sum p_1 q_1 - \sum p_0 q_0$$

$$I\left(\sum Q\right) = \frac{\sum Q_1}{\sum Q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_0 q_1}{\sum p_0 q_0} \cdot \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \cdot \frac{\sum p_1 q_1}{\sum p_1 q_0}$$

$$= \quad I_q^{(L)} \cdot I_p^{(P)} \quad = \quad I_p^{(L)} \cdot I_q^{(P)}$$

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

$$Ip^{(P)} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{Ip}} = \frac{\sum Q_1}{\sum \frac{Q_1}{Ip}}$$

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

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

$$Iq^{(P)} = \frac{\sum p_1 q_1}{\sum p_1 q_0} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{Iq}} = \frac{\sum Q_1}{\sum \frac{Q_1}{Iq}}$$

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