Statisztika I.

KÉPLETGYŰJTEMÉNY

VISZONYSZÁMOK

Láncviszonyszám: $l_i = \frac{y_i}{y_{i-1}}$ Bázisviszonyszám: $b_i = \frac{y_i}{y_b}$

Lánc- és bázisviszonyszámok összefüggése: $\frac{b_i}{b_{i-1}} = l_i$ és $l_1 l_2 \cdots l_k = b_k$

Megoszlási viszonyszám = Relatív gyakoriság: $g_i = \frac{f_i}{\sum_{i=1}^{k} f_i} = \frac{f_i}{n}$

ÉRTÉKÖSSZEGEK

Abszolút értékösszeg: $s_i = f_i x_i$ Relatív értékösszeg: $z_i = \frac{f_i \cdot x_i}{\sum_{i=1}^{k} f_i \cdot x_i}$

KÖZÉPÉRTÉKEK

Számtani átlag: $\bar{x} = \frac{\sum_{i=1}^{k} f_i \cdot x_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} g_i \cdot x_i}{\sum_{i=1}^{k} g_i}$ Mértani átlag: $\bar{X}_g = \sqrt[k]{l_1 l_2 \cdots l_k}$

Harmonikus átlag: $\bar{x}_h = \frac{\sum_{i=1}^{k} s_i}{\sum_{i=1}^{k} \frac{s_i}{x_i}} = \frac{\sum_{i=1}^{k} z_i}{\sum_{i=1}^{k} \frac{z_i}{x_i}}$ Négyzetes átlag: $\bar{x}_q = \sqrt{\frac{\sum_{i=1}^{k} f_i \cdot {x_i}^2}{\sum_{i=1}^{k} f_i}} = \sqrt{\frac{\sum_{i=1}^{k} g_i \cdot {x_i}^2}{\sum_{i=1}^{k} g_i}}$

Módusz: $Mo = x_{i0} + \frac{k_1}{k_1 + k_2} \cdot h_i$ Medián: $Me = x_{i0} + \frac{\frac{n}{2} - f'_{i-1}}{f_i} \cdot h_i$

KVARTILISEK

Alsó kvartilis: $Q_1 = x_{i0} + \frac{\frac{n}{4} - f'_{i-1}}{f_i} \cdot h_i$ Felső kvartilis: $Q_3 = x_{i0} + \frac{\frac{3 \cdot n}{4} - f'_{i-1}}{f_i} \cdot h_i$

SZÓRÓDÁSI MUTATÓK

Átlagos abszolút eltérés: $d = \frac{\sum f_i \cdot |x_i - Me|}{\sum f_i}$

Variancia, szórásnégyzet: $Var = \sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$ $\sigma^2 = \sum g_i (x_i - \bar{x})^2$

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

Szórás $\sigma = \sqrt{Var}$

Relatív szórás: $V = \frac{\sigma}{\bar{x}}$

ASZIMMETRIA

Pearson-féle mutató: $A = \frac{\bar{x} - Mo}{\sigma}$ F-mutató: $F = \frac{(Q_3 - Me) - (Me - Q_1)}{(Q_3 - Me) + (Me - Q_1)}$

ÉRTÉK-, VOLUMEN- ÉS ÁRINDEXEK

Az egyedi érték, ár és volumen összefüggése: $v_i = q_i \cdot p_i$ $q_i = \frac{v_i}{p_i}$ $p_i = \frac{v_i}{q_i}$

Értékindex: $I_v = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot \frac{q_1 \cdot p_1}{q_0 \cdot p_0}}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot i_v}{\sum q_0 \cdot p_0} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{q_1 \cdot p_1 / q_0 \cdot p_0}} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{i_v}}$

Statisztika I.

Laspeyres-féle volumenindex: $I_q{}^L = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot \frac{q_1}{q_0}}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot i_q}{\sum q_0 \cdot p_0} = \sum w_0 \cdot i_q$

ahol $w_0 = \frac{q_{i0} \cdot p_{i0}}{\sum_i q_{i0} \cdot p_{i0}}$

Paasche-féle volumenindex: $I_q{}^P = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_1} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{q_1 / q_0}} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{i_q}}$

Fisher-féle volumenindex: $I_q{}^F = \sqrt{I_q{}^L \cdot I_q{}^P}$

Laspeyres-féle árindex: $I_p{}^L = \frac{\sum q_0 \cdot p_1}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot \frac{p_1}{p_0}}{\sum q_0 \cdot p_0} = \frac{\sum q_0 \cdot p_0 \cdot i_p}{\sum q_0 \cdot p_0} = \sum w_0 \cdot i_p$

Paasche-féle árindex: $I_p{}^P = \frac{\sum q_1 \cdot p_1}{\sum q_1 \cdot p_0} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{p_1 / p_0}} = \frac{\sum q_1 \cdot p_1}{\sum \frac{q_1 \cdot p_1}{i_p}}$

Fisher-féle árindex: $I_p{}^F = \sqrt{I_p{}^L \cdot I_p{}^P}$

Indexek közötti összefüggések:

$$I_v = I_q{}^L \cdot I_p{}^P \qquad I_v = I_q{}^P \cdot I_p{}^L \qquad I_v = I_q{}^F \cdot I_p{}^F$$

ASSZOCIÁCIÓS KAPCSOLAT

Gyakoriság függetlenség esetén: $f_{ij}{}^* = \frac{f_{i.} \cdot f_{.j}}{n} = n \cdot g_{i.} \cdot g_{.j}$

Khi-négyzet: $\chi^2 = \sum_{i=1}^{s} \sum_{j=1}^{t} \frac{(f_{ij} - f_{ij}{}^*)^2}{f_{ij}{}^*}$ $\qquad s = sorok\ száma \qquad t = oszlopok\ száma$

Cramer együttható: $C = \sqrt{\frac{\chi^2}{n \cdot (s-1)}}, \quad ha \quad s \leq t$

VEGYES KAPCSOLAT

A külső, a belső és a teljes szórásnégyzet összefüggése: $\sigma^2 = \sigma^2{}_B + \sigma^2{}_K$

Belső szórásnégyzet: $\sigma_B{}^2 = \frac{\sum_j \sum_i (x_{ij} - \bar{x}_j)^2}{n} = \frac{\sum_j n_j \cdot \sigma_j{}^2}{\sum_j n_j}$

Külső szórásnégyzet: $\sigma_K{}^2 = \frac{\sum_j n_j \cdot (\bar{x}_j - \bar{x})^2}{\sum_j n_j}$

Variancahányados: $H^2 = \frac{\sigma^2{}_K}{\sigma^2} = \frac{\sigma^2{}_K}{\sigma^2{}_K + \sigma^2{}_B}$

KORRELÁCIÓS KAPCSOLAT

A lineáris regressziós egyenlet:

$$\hat{y} = b_1 x + b_0$$

A paraméterek meghatározása normálegyenletek segítségével:

$$\sum_{i=1}^{n} y_i = n \cdot b_0 + b_1 \sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} y_i \cdot x_i = b_0 \cdot \sum_{i=1}^{n} x_i + b_1 \cdot \sum_{i=1}^{n} x_i^2$$

A paraméterek meghatározása képletek segítségével: $(d_x = x - \bar{x} \qquad d_y = y - \bar{y})$

b_1 paraméter:

$$b_1 = \frac{\bar{x}\bar{y} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{\sum xy - n \cdot \bar{x} \cdot \bar{y}}{\sum x^2 - n \cdot \bar{x}^2} = \frac{\sum (x-\bar{x}) \cdot (y-\bar{y})}{\sum (x-\bar{x})^2} = \frac{\sum d_x d_y}{\sum d_x^2} = \frac{C}{\sigma_x^2} = \frac{r \cdot \sigma_y}{\sigma_x}$$

b_0 paraméter:

$$b_0 = \bar{y} - b_1 \cdot \bar{x}$$

Kovariancia:

$$C = \frac{\sum (x-\bar{x}) \cdot (y-\bar{y})}{n} = \frac{\sum d_x \cdot d_y}{n} = \frac{\sum x \cdot y}{n} - \bar{x} \cdot \bar{y}$$

Pearson-féle lineáris korrelációs együttható:

$$r = \sqrt{\frac{\frac{\sum (\hat{y}_i - \bar{y})^2}{n}}{\sigma_y^2}} = \frac{\sum (x_i-\bar{x}) \cdot (y_i-\bar{y})}{n \cdot \sigma_x \cdot \sigma_y} = \frac{\sum d_x d_y}{n \cdot \sigma_x \cdot \sigma_y} = \frac{\sum (x_i-\bar{x}) \cdot (y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \cdot \sum (y_i-\bar{y})^2}} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \cdot \sum d_y^2}}$$

$$r = \frac{C}{\sigma_x \cdot \sigma_y} = \frac{b_1 \cdot \sigma_x}{\sigma_y}$$

Determinációs együttható (lineáris esetben): $D = r^2$