

## Poprawność korelacyjna zmiennych w modelu ekonometrycznym

Rozpatruje się zbiór  $Z(m)$  zmiennych wyróżnionych w modelu ekonometrycznym ze zmiennymi objaśniającymi  $X_1, \dots, X_m$  i zmienną objaśnianą  $Y$ , których wartości zaobserwowane generują macierz  $R(m) = [r_{ij}]_{m \times m}$  oraz wektor  $R_0(m) = [r_i]_{m \times 1}$ , gdzie  $r_{ij}$  jest współczynnikiem korelacji między zmiennymi  $X_i, X_j$ , zaś  $r_i$  — współczynnikiem korelacji pomiędzy zmiennymi  $X_i, Y$ ,  $i, j = 1, \dots, m$ .

Definiuje się następujące potencjały korelacyjne:

- 1) Wewnętrzny potencjał korelacyjny  $C_w(m) = \max_{1 \leq j \leq m} \sum_{i=1}^m |r_{ij}|$  zbioru  $Z(m)$  jako miarę wzajemnego skorelowania zmiennych objaśniających.
- 2) Zewnętrzny potencjał korelacyjny  $C_z(m) = \left( \sum_{i=1}^m |r_i|^2 \right)^{1/2}$  zbioru  $Z(m)$  jako miarę skorelowania zmiennych objaśniających ze zmienną objaśnianą.
- 3) Potencjał korelacyjny  $C(m) = C_w(m) + C_z(m)$  zbioru  $Z(m)$  jako miarę skorelowania wszystkich jego zmiennych.
- 4) Względny wewnętrzny potencjał korelacyjny  $D_w(m) = \frac{C_w(m)}{C(m)}$  zbioru  $Z(m)$  obrazujący udział  $C_w(m)$  w  $C(m)$ .
- 5) Względny zewnętrzny potencjał korelacyjny  $D_z(m) = \frac{C_z(m)}{C(m)}$  zbioru  $Z(m)$  obrazujący udział  $C_z(m)$  w  $C(m)$ .  
Oczywiście  $D_w(m) + D_z(m) = 1$ .

$D_w(m) \in \langle \frac{1}{2}, 1 \rangle$  i im mniejsze lub równoważnie im większe  $\frac{C_z(m)}{C_w(m)}$ , tym mniejsze  $C_w(m)$  i większe  $C_z(m)$ .  $D_z(m) \in \langle 0, \frac{1}{2} \rangle$  i im większe lub równoważnie im większe  $\frac{C_z(m)}{C_w(m)}$ , tym większe  $C_z(m)$  i mniejsze  $C_w(m)$ .

Określa się poprawność korelacyjną zmiennych zbioru  $Z(m)$  jako słabe skorelowanie zmiennych objaśniających między sobą i ich silne skorelowanie ze zmienną objaśnianą. Stopień poprawności korelacyjnej proponuje się mierzyć współczynnikiem  $K(m) = \frac{C_z(m)}{C_w(m)}$  poprawności korelacyjnej.  $K(m) \in \langle 0, 1 \rangle$  i im większe, tym wyższy stopień poprawności korelacyjnej.

Za pomocą kryterium maksymalnego współczynnika  $K(m)$  doboru zmiennych objaśniających do modelu otrzymuje się optymalny podzbiór tych zmiennych bez współliniowości.