

Standardní (též klasický) lineární regresní model

Specifikace=formulace (jednorovnicového) lineárního regresního modelu

$$(1) \quad y = X\beta + \varepsilon \quad \text{resp.} \quad y_t = \sum_{j=1}^k X_{tj}\beta_j + \varepsilon_t \quad , \text{ kde}$$

y je T-členný (sloupcový) vektor pozorování závisle (vysvětlující) proměnné (*regresandu*)

X je [Txk] matice pozorování k nezávisle proměnných (*regresorů*) – *matice plánu*

β je k-členný (sloupcový) vektor neznámých regresních koeficientů

ε je T-členný (sloupcový) vektor nepozorovatelných náhodných složek (*disturbancí*)

Ve strukturním vektorově- maticovém vyjádření

$$(1A) \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_T \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ x_{31} & x_{32} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_{T1} & x_{T2} & \dots & x_{Tk} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_T \end{pmatrix}$$

Poznámka: Pokud je v rovnici zastoupena úroňová konstanta β_1 , pak k ní příslušný jedničkový vektor $l = (1, 1, \dots, 1)'$ bude obsažen prvním sloupci matice X , tzn. že bude platit $x_{t1} = 1$ pro všechna $t = 1, 2, \dots, T$.

Předpokládáme, že **počet vysvětlujících proměnných je menší** (v krajním případě roven) než počet pozorování, tedy, že platí nerovnost
 $k < T$

Odhady parametrů nějakou vhodnou odhadovou metodou (např. **metodou nejmenších čtverců**)

$$(2) \quad b = \hat{\beta}(y, X) \quad \text{např.} \quad \hat{\beta} = (X'X)^{-1}(X'y)$$

Definice vyrovnaných hodnot závisle proměnné

$$(3) \quad \hat{y} = Xb \quad y_t = \sum_{j=1}^k X_{tj}b_j$$

Definice reziduí (odhadnutých náhodných složek)

$$(4) \quad e = \hat{\varepsilon} = y - \hat{y} \quad \text{neboli} \quad y = \hat{y} + e$$

odtud vyplývá možnost **alternativního zápisu závisle proměnné** (s odhady parametrů a s reziduy)

$$(5) \quad y = Xb + e$$

Vlastnosti proměnných lineárního regresního modelu

1. Centrovanost náhodných složek

$$E \varepsilon = 0 \quad \text{neboli} \quad E \varepsilon_t = 0 \quad \text{pro všechna } t = 1, 2, \dots, T$$

nebude-li splněno, pak střední hodnota náhodných složek (ani reziduí) nebude nulová a regresní přímka nepovede „středem“ oblasti pozorovaných hodnot, ale nad či pod ní (nepůjde o regresní přímku ve vlastním slova smyslu). Odchytky nebudou „nestranné“ a součet reziduí nebude roven 0.

2. Diagonalita kovarianční matice náhodných složek

$$\text{Var}(\varepsilon) = \text{Cov}(\varepsilon, \varepsilon') = \sigma^2 \cdot I_T \quad (\text{diagonální matice se stopou } T \cdot \sigma^2)$$

což v sobě obsahuje dvě vlastnosti, jimiž jsou

2a) homoskedasticita náhodných složek

$$\text{var } \varepsilon_t = \sigma^2 \quad (\text{rozptyl nezávislý na indexu pozorování})$$

nebude-li splněno, pak budou mít náhodné složky v různých pozorováních různý rozptyl a metoda OLS ztratí svou vydatnost (byť odhady zůstanou nestranné). Nejde však o fatální problém a parametry budou odhadnutelné. Různost rozptylů náhodných složek v různých pozorováních se nazývá *heteroskedasticita*. Ta je odstranitelná nebo zmírnitelná některými speciálními postupy, např. použitím vážené metody nejmenších čtverců WLS.

2b) nekorelovanost náhodných složek

$$\text{Cov}(\varepsilon_s, \varepsilon_t) = E(\varepsilon_s \cdot \varepsilon_t) = \delta_{st} \cdot \sigma^2, \quad \text{kde } \delta_{st} = 1 \quad \text{pro } s = t \quad \text{a také } \delta_{st} = 0 \quad \text{pro } s \neq t$$

nebude-li splněno, pak budou náhodné složky v různých pozorováních vzájemně korelované a odhady parametrů nebudou vydatné (byť zůstanou nestranné). K zajištění optimálních vlastností bude nutno uplatnit *zobecněnou metodu nejmenších čtverců GLS* (pokud budou dodatečné informace o modelu) nebo problém zmírnit některou speciální technikou připouštějící autokorelaci náhodných složek.

3. Nekorelovanost náhodných složek s nezávisle proměnnými

$$E(X' \cdot \varepsilon) = 0 \quad \text{neboli} \quad E(x_{tj} \cdot \varepsilon_t) = 0 \quad \text{pro všechna } j = 1, 2, \dots, k$$

nebude-li splněno, pak to znamená, že informace obsažená v náhodných složkách má něco společného s informací obsaženou v některé vysvětlované proměnné a že „náhodná složky“ má v sobě kousek (nebo kus) systematické informace. To je v rozporu s uvažovaným charakterem náhodné složky. Odhad parametrů nebude možno takto nijak statisticky získat (ani jinou metodou než OLS)

4. Plná hodnost matice vysvětlujících proměnných $h(X) = k$

nebude-li splněno, pak bude mít matice X hodnost menší než k , což bude znamenat, že některé její sloupce budou lineárně závislé. Jinými slovy: informace obsažená ve sloupcích matice X , tedy v jednotlivých vysvětlujících proměnných není nezávislá a vzájemně se prolíná. Z algebraického hlediska to má ten následek, že matice $X'X$ bude singulární (její determinant bude nulový) a nebude k ní existovat (jednoznačně určená inverzní matice). Odhad parametrů (metodou OLS) takto nebude možné určit, resp. při použití pseudoinverze nebude odhad určen jednoznačně.

(prostá, obyčejná) Metoda nejmenších čtverců (MNČ, OLS)

[Ordinary least squares method]

Minimalizačním kritériem je zde součet čtverců reziduí (odhadnutých náhodných složek neboli rozdílů mezi pozorovanými a vyrovnanými hodnotami) :

$Min e'e = Min (y - Xb)'(y - Xb)$ neboli v pozorovaných hodnotách

$$Min \sum_{t=1}^T e_t^2 = Min \sum_{t=1}^T (y_t - \hat{y}_t)^2 = Min \sum_{t=1}^T \left(y_t - \sum_{j=1}^K x_{tj} b_j \right) \left(y_t - \sum_{k=1}^K x_{tk} b_k \right)$$

Polohu minima (tj. bodu=odhadnutého vektoru parametrů b), ve kterém je minimalizovaný výraz nejmenší) nalezneme řešením **soustavy tzv. normálních rovnic**

$$\frac{\partial e'e}{\partial b} = \frac{\partial (y - Xb)'(y - Xb)}{\partial b} = -2X'y + 2X'Xb = 0 \quad (\text{vektorově})$$

tuto soustavu řešíme úpravami (vydělením 2, přeskupením členů) na tvar $X'Xb = X'y$, která má řešení pro b ve tvaru $b = (X'X)^{-1} X'y$, ← řešení je jednoznačné, neboť vzhledem k předpokladům $h(X) = k$, existuje (jediná) inverzní matice k matici $X'X$.

Poznámka Vyjádření minimalizace v pozorovaných hodnotách:

$$H(y, X, b) = \sum_{t=1}^T \left(y_t - \sum_{i=1}^K x_{ti} b_i \right) \left(y_t - \sum_{j=1}^K x_{tj} b_j \right) = \sum_{t=1}^T y_t^2 - \sum_{t=1}^T y_t \left(\sum_{j=1}^K x_{tj} b_j \right) - \sum_{t=1}^T y_t \left(\sum_{i=1}^K x_{ti} b_i \right) + \sum_{t=1}^T \left(\sum_{i=1}^K x_{ti} b_i \right) \left(\sum_{j=1}^K x_{tj} b_j \right)$$

$$\frac{\partial H(y, X, b)}{\partial b_j} = 0 - 2 \sum_{t=1}^T x_{tj} y_t + 2 \sum_{t=1}^T x_{tj} \sum_{i=1}^k x_{ti} b_i = 0 \quad \text{pro } j = 1, 2, \dots, k \quad \text{a odtud}$$
$$-2X'y \quad + 2X'Xb$$

$$\sum_{t=1}^T x_{tj} \sum_{i=1}^k x_{ti} b_i = \sum_{t=1}^T x_{ti} y_t \quad \text{neboli} \quad b_i = \left(\sum_{t=1}^T \sum_{l=1}^k x_{tl} x_{lj} \right)^{-1} \sum_{t=1}^T x_{ti} y_t$$

V případě konečných sumací můžeme přeskupovat členy v součtech

Příklad: Lineární regrese s jedinou vysvětlující proměnnou (+úrovňovou konstantou) ($x_{t1} = 1$)

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} T & \sum x_{t2} \\ \sum x_{t2} & \sum x_{t2}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_t \\ \sum x_{t2} y_t \end{pmatrix} = \frac{1}{D} \begin{pmatrix} \sum x_{t2}^2 & -\sum x_{t2} \\ -\sum x_{t2} & T \end{pmatrix} \begin{pmatrix} \sum y_t \\ \sum x_{t2} y_t \end{pmatrix}, \text{ kde}$$

$$D = \begin{vmatrix} T & \sum x_{t2} \\ \sum x_{t2} & \sum x_{t2}^2 \end{vmatrix} = T \cdot \sum x_{t2}^2 - (\sum x_{t2})^2. \quad \text{Odtud máme}$$

$$\beta_1 = \frac{(\sum x_{t2}^2)(\sum y_t) - (\sum x_{t2})(\sum x_{t2} y_t)}{T \cdot \sum x_{t2}^2 - (\sum x_{t2})^2}$$

hodnota úrovňové konstanty

$$\beta_2 = \frac{T \cdot (\sum x_{t2} y_t) - (\sum x_{t2})(\sum y_t)}{T \cdot \sum x_{t2}^2 - (\sum x_{t2})^2}$$

hodnota parametru sklonu regresní přímky

Vlastnosti obyčejné (prosté) metody nejmenších čtverců MNČ (OLS) v klasickém lineárním regresním modelu

poskytuje odhad ${}_{OLS}\hat{\beta}$ regresních koeficientů β ve tvaru ${}_{OLS}\hat{\beta} = (X'X)^{-1}X'y$

Věta 1 (Gauss-Markovova) Odhad regresních koeficientů ${}_{OLS}\hat{\beta}$ pořizený obyčejnou (prostou) metodou nejmenších čtverců je nejlepším nestranným lineárním odhadem vektoru parametrů β

Důkaz rozdělíme jej na několik částí.

A. odhad ${}_{OLS}\hat{\beta}$ je nestranný (pro libovolnou velikost vzorku T).

Ověření nestrannosti:

$$E(\hat{\beta}) = E(X'X)^{-1}X'y = E(X'X)^{-1}X'(X\beta + \varepsilon) = E(X'X)^{-1}X'X\beta + E(X'X)^{-1}X'\varepsilon =$$

vyjádření $y=X\beta+\varepsilon$

$$= (X'X)^{-1}X'XE\beta + (X'X)^{-1}X'E\varepsilon = E\beta + (X'X)^{-1}X'0 = \beta + 0 = \beta \quad \square.$$

nestochastičnost X $E\varepsilon=0$ $E\beta=\beta$ (β nestochastický vektor)

Důsledek A Nestranná odhadová funkce je vždy asymptoticky nestranná.

Platí-li totiž $E({}_{(T)}\hat{\beta}) = \beta$ ² pro každé konečné T , platí tentýž vztah i pro $T \rightarrow \infty$.

B. odhad ${}_{OLS}\hat{\beta}$ je konzistentní, tj. platí $plim \hat{\beta} = \beta$
 $T \rightarrow \infty$

Vlastnost platí i pro případ, že matice X je stochastická.

Ověření konzistence: lze vyvodit z následujícího tvrzení

Tvrzení Jestliže odhadová funkce je asymptoticky nestranná a kovarianční matice této odhadové funkce konverguje při $T \rightarrow \infty$ k nulové matici, pak je tato odhadová funkce konzistentní.

Důkaz a) asymptotická nestrannost OLS-odhadové funkce vyplývá z *důsledku A*.

b) konvergenci kovarianční matice OLS-odhadové funkce (její tvar viz ad D) k nulové matici ukážeme následovně

Dále ukážeme, že $Cov(\hat{\beta}) = \sigma_\varepsilon^2 (X'X)^{-1}$. Definujme matici $P = (X'X)/T$ a její ij -tý prvek označíme jako P_{ij} .³ Pro tento prvek platí

$$|P_{ij}| = \left| \frac{1}{T} \cdot \sum_{t=1}^T x_{ti} \cdot x_{tj} \right| \leq \zeta_{ij}^2, \text{ přičemž tato rovnost platí pro všechna } T$$

² Symbolem nalevo rozumíme střední hodnotu odhadnutého vektoru parametrů β spočteného na základě T pozorování.

³ Účelem je ukázat, že matice P má prvky o konečné velikosti.

Všechny prvky této matice jsou tedy (v absolutní hodnotě) shora omezeny hodnotou $\xi^2 = \max \xi_{ij}^2$. Tedy, pro všechna T má matice P konečně velké prvky a je nesingulární. Zřejmě dále platí

$$(X'X)^{-1} = P^{-1} / T \quad (\text{neboť platí } X'X = T.P)$$

a navíc matice P^{-1} má konečně velké prvky (π_{ij}) pro všechna T . Platí tedy

$$p \lim_{T \rightarrow \infty} \text{Cov}(\text{OLS} \hat{\beta}) = p \lim_{T \rightarrow \infty} \sigma_\varepsilon^2 (X'X)^{-1} = \sigma_\varepsilon^2 p \lim_{T \rightarrow \infty} P^{-1} / T = 0_k,$$

kde 0_k je (čtvercová) matice řádu k složená ze samých nul.

C. odhad $\text{OLS} \hat{\beta}$ je lineární (vzhledem k vysvětlované proměnné y), neboť je definován jako lineární forma pozorování závisle proměnné y .

Ověření linearity Lze psát $\hat{\beta} = (X'X)^{-1} X'y = C'y$, kde matice $C' = (X'X)^{-1} X'$ představuje koeficienty lineární formy, jejíž proměnné tvoří složky vektoru y .

Poznamenejme, že vždy platí $C'X = (X'X)^{-1} X'X = I_k$, kde I_k je jednotková matice řádu k .

D. kovarianční matice příslušná odhadové funkci $\hat{\beta} = (X'X)^{-1} X'y$ má následující tvar

$$\text{Cov}(\hat{\beta}) = \sigma_\varepsilon \cdot (X'X)^{-1}$$

Ověření $\text{Cov}(\hat{\beta}) = E\{[\hat{\beta} - E(\hat{\beta})][\hat{\beta} - E(\hat{\beta})]'} = E\{[\hat{\beta} - \beta][\hat{\beta} - \beta]'} =$

(dosazení $\hat{\beta} = (X'X)^{-1} X'y$)

$$= E\{[(X'X)^{-1} X'y - \beta][(X'X)^{-1} X'y - \beta]'} = E\{[(X'X)^{-1} X'(X\beta + \varepsilon) - \beta][(X'X)^{-1} X'(X\beta + \varepsilon) - \beta]'} =$$

$$E\{[\beta + (X'X)^{-1} X'\varepsilon - \beta][\beta + (X'X)^{-1} X'\varepsilon - \beta]'} = E\{[(X'X)^{-1} X'\varepsilon][(X'X)^{-1} X'\varepsilon]'} =$$

$$\stackrel{\text{(transpozice)}}{=} E\{[(X'X)^{-1} X'\varepsilon \varepsilon' X (X'X)^{-1}] = (X'X)^{-1} X' E\varepsilon \varepsilon' X (X'X)^{-1} =$$

$$\stackrel{\text{(vytknutí skaláru } \sigma_\varepsilon^2 \text{)}}{=} (X'X)^{-1} X' \sigma_\varepsilon^2 I_T X (X'X)^{-1} = \sigma_\varepsilon^2 (X'X)^{-1} X' X (X'X)^{-1} = \sigma_\varepsilon^2 (X'X)^{-1} \quad \square.$$

Odtud plyne **důsledek D1**.

Směrodatné odchylky odhadnutých regresních parametrů $\sigma_{\hat{\beta}}$ získáme jako

$$\sigma_{\hat{\beta}_j} = \sigma_\varepsilon \sqrt{\xi_{jj}}, \quad \text{kde}$$

ξ_{jj} je j -tý diagonální prvek inverzní momentové matice $(X'X)^{-1}$

σ_ε je směrodatná odchylka náhodných složek (stejná u všech ε_t).

$$\begin{aligned}
& \text{[protože } \sigma^2 \text{ je skalární hodnota]} \\
& = \sigma^2 (X'X)^{-1} + \sigma^2 G'G + \sigma^2 (X'X)^{-1} X'G + \sigma^2 G'X (X'X)^{-1} = \\
& \text{[protože } G'X = 0 \text{ a stejně } X'G = 0 \text{]} \\
& = \sigma^2 (X'X)^{-1} + \sigma^2 G'G = \text{Cov}(\hat{\beta}) + \sigma^2 G'G
\end{aligned}$$

kde $G'G$ je zřejmě pozitivně semidefinitní matice⁴ a $\sigma^2 > 0$ □

F. pro odhad rozptylu reziduí dostaneme

$$\begin{aligned}
E(e'e) &= E(\varepsilon' M M \varepsilon) = E(\varepsilon' M \varepsilon) = E \text{tr}(\varepsilon' M \varepsilon) = E \text{tr}(M \varepsilon \varepsilon') = \\
& \text{(M je idempotentní matice)} \quad \text{(skalár je současně svou stopou)} \quad \text{(tr A.B = tr B.A)} \\
& = \text{tr} E(M \varepsilon \varepsilon') = \text{tr}(M E \varepsilon \varepsilon') = \text{tr}(M \sigma_\varepsilon^2 I_T) = \sigma_\varepsilon^2 \text{tr}(M) = \sigma_\varepsilon^2 (T - k) \\
& \text{(záměna stopy a střední hodnoty)} \quad \text{(M je nestochastická)} \quad \text{(\sigma^2 je skalární hodnota)}
\end{aligned}$$

protože

$$\text{tr}(M) = \text{tr}(I_T - X(X'X)^{-1}X') = \text{tr}(I_T - X'X(X'X)^{-1}) = \text{tr}(I_T - I_K) = T - k$$

(definice M) (platí tr A.B = tr B.A) (stopa jednotkové matice je rovna její dimenzi)

G. z tvrzení F plyne, že nestranným odhadem rozptylu náhodných složek σ_ε^2

$$\text{je výraz } \hat{\sigma}_\varepsilon^2 = \frac{e'e}{T - k}, \text{ neboť zřejmě platí } E \left[\frac{\sum e_t^2}{T - k} \right] = \sigma_\varepsilon^2.$$

Z **důsledku D1** je patrné, že

odhady směrodatných odchylek odhadnutých regresních parametrů $s_{\hat{\beta}}$ získáme jako

$$s_{\hat{\beta}_j} = s_e \cdot \sqrt{\xi_{jj}}, \text{ kde}$$

ξ_{jj} je j-tý diagonální prvek inverzní momentové matice $(X'X)^{-1}$

s_e je směrodatná odchylka reziduí (stejná pro všechna $j = 1, 2, \dots, k$)

⁴ Matice $P = G'G$ (stejně jako kterákoliv jiná symetrická nenulová matice $H'H$) je pozitivně definitní, protože pro ni platí $x'Px = x'G'Gx = z'z > 0$ (skalární součin je nulový jen pro identicky nulový vektor).

Poznámka: Prostá metoda nejmenších čtverců není jedinou používanou odhadovou metodou v prostředí standardního lineárního regresního modelu. K dalším technikám patří:

Metoda maximální věrohodnosti ML (Maximum Likelihood) je založena na maximalizaci sdružené hustoty (tzv. věrohodnostní funkce) rozdělení náhodných složek. Lokalizuje se tedy poloha modusu pro β a σ^2 , v němž tato funkce nabývá maxima.

Metoda nejmenších absolutních odchylek LAD (Least Absolute Deviations) je založena na minimalizačním kritériu tvaru

$$\sum_{t=1}^T \left| y_t - \sum_{j=1}^k x_{tj} \beta_j \right|$$

Odhady pořízené metodou LAD nelze vyjádřit v explicitním tvaru, ale je nutno použít iterační postup (např. algoritmy R.L.Faira)

Zobecněná momentová metoda GMM (Generalized moment method).