Standardní (též klasický) lineární regresní model

Specifikace=formulace (jednorovnicového) **lineárního regresního modelu**

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

- y **je T-členný** (sloupcový) **vektor pozorování závisle (vysvětlující) proměnné (*regresandu*)**
- X **je [Txk] matice pozorování k nezávisle proměnných (*regresorů*) – *tzv. 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) $$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_T \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ x_{31} & x_{32} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_{T1} & x_{T2} & \dots & x_{Tk} \end{pmatrix} . \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots. \\ \varepsilon_T \end{pmatrix}$$

Poznámka: Pokud je v rovnici zastoupena úrovňová konstanta β_1**, pak k ní příslušný jedničkový vektor** $1 = (1,1,\dots,1)'$ **bude** obvykle **obsažen v 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) $$b = \hat{\beta}(y, X) \qquad \textbf{např.} \qquad \hat{\beta} = (X'X)^{-1}(X'y)$$

Definice vyrovnaných hodnot závisle proměnné

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

Definice reziduí (odhadnutých náhodných složek)

(4) $$e = \hat{\varepsilon} = y - \hat{y} \qquad \textbf{neboli} \qquad y = \hat{y} + e$$

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

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

Některé další vztahy mezi veličinami lineárního regresního modelu

Z porovnání (1) a (5) obdržíme vztah mezi ε **a** $e = \hat{\varepsilon}$:

(6) $$X\beta + \varepsilon = Xb + e$$ [1]

(pokud použijeme odhad OLS, pak pravá strana $= X(X'X)^{-1}X'y + e$ **),**
z čehož dále plyne

(7) $$e = y - Xb = y - X(X'X)^{-1}X'y = \left[I_T - X(X'X)^{-1}X'\right]y = My$$

(8) **neboli platí** $\hat{\varepsilon} = M\boldsymbol{y}$

(9)
$$e = y - Xb = X\beta + \varepsilon - Xb = X\beta + \varepsilon - X(X'X)^{-1}X'y = X\beta + \varepsilon - X(X'X)^{-1}X'[X\beta + \varepsilon] =$$
$$= X\beta + \varepsilon - X\beta - \left[X(X'X)^{1}X'\right]\varepsilon = \left[I - X(X'X)^{1}X'\right]\varepsilon = M\varepsilon$$ [2]

(10) **neboli platí** $e = M\varepsilon$

Varovná poznámka:

Ze vztahů $e = My$ **a** $e = M\varepsilon$ **nelze usuzovat, že platí** $y = \varepsilon$, **neboť** M **je vždy singulární matice. Nelze proto psát** $y = M^{-1}\varepsilon$ **, resp.** $e = M^{-1}\varepsilon$, **protože matice** M^{-1} **neexistuje.**

Hodnost matice $M = I_T - X(X'X)^{-1}X'$ **je** $T - k < T$

protože $$h(I_T) = T \qquad h\left(X(X'X)^{-1}X'\right) = h\left(X'X\,(X'X)^{-1}\right) = k$$

Na výše uvedené zjištění můžeme pohlížet také z toho hlediska, že k určení vektoru reziduí e je skrze matici M „vytahována“ stejná informace jako z vektoru závisle proměnné. Matice M jakoby „odfiltruje“ „přebytečnou“ informaci (potřebnou pro určení e) nacházející se v y, ale nikoliv v ε.

Poznámka:

Centrovanost náhodných složek se přenáší na rezidua, která mají rovněž nulovou střední hodnotu, neboť

(11) $$E(e) = E(M\varepsilon) = E\left(I_T - X(XX)^{-1}X'\right)\varepsilon = \left(I_T - X(XX)^{-1}X'\right)E(\varepsilon) = 0.$$
dle (10) nestochastičnost M centrovanost ε

---

[1] Na levé straně (6) máme dvě nepozorovatelné veličiny: *β,ε*, zatímco na pravé straně jsou jak odhadnuté parametry *b*, tak rezidua *e,* obojí určené z pozorovaných proměnných: Je ovšem třeba dodat, že výpočet *b* (a následně i *e*) obecně závisí na užité odhadové metodě.U standardního lineárního modelu regresního i jiné odhadové metody vedou k témuž výrazu, avšak u složitějších modelů tomu ale tak nemusí být.

[2] Poznamenejme, že vzorce **(7)** a **(9)** platí pouze tehdy, odhadujeme-li parametry modelu metodou OLS.

Vlastnosti proměnných lineárního regresního modelu

1. Centrovanost náhodných složek

(12) $E\,\varepsilon = 0$ **neboli** $E\,\varepsilon_t = 0$ **pro všechna** $t = 1,2,\ldots,T$

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). Odchylky nebudou „nestranné" a součet reziduí nebude roven 0.

2. Diagonalita kovarianční matice náhodných složek

(13) $Var(\varepsilon) = Cov(\varepsilon,\varepsilon') = \sigma^2 . I_T$ **(diagonální matice se stopou** $T.\sigma^2$**)**

což v sobě obsahuje dvě vlastnosti, jimiž jsou

2a) homoskedasticita náhodných složek

(13A) $var\,\varepsilon_t = \sigma^2$ **(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 modelu 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

(13B) $Cov(\varepsilon_s,\varepsilon_t) = E(\varepsilon_s.\varepsilon_t) = \delta_{st}.\sigma^2$ **, kde** $\delta_{st} = 1$ **pro** $s = t$ **a také** $\delta_{st} = 0$ **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

(14) $E(X'.\varepsilon) = 0$ **neboli** $E(x_{tj}.\varepsilon_t) = 0$ **pro všechna** $j = 1,2,\ldots,k$

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 složka" má v sobě kousek (nebo víc) systematické, nenáhodné informace. To je v rozporu s přijatým charakterem *náhodné* složky. Odhad parametrů nebude možno takto nijak statisticky získat (ani jinou metodou než OLS)

4. Plná hodnost matice vysvětlujících proměnných

(15) $h(X) = k$

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 v některých sloupcích matice X je „redundantní" a některý z nich by mohl být beztrestně vynechán. Z algebraického hlediska to má za následek to, 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 užití pseudoinverze nebude odhad určen jednoznačně. Požadavek znamená, aby každá z vysvětlujících proměnných nesla v sobě „aspoň kousek" informace neobsažené v jiných vysvětlujících proměnných.

Poznámka: Tento požadavek nezaměňujme s daleko přísnější podmínkou tzv.**ortogonality** vysvětlujících proměnných vyjádřené zápisem $x_{ti}.x_{tj} = 0\,, i, j = 1,2,...,k$, která by znamenala navíc to, že kterákoliv z vysvětlujících proměnných nemá v sobě „ani kousek" informace obsažené v jiné vysvětlující proměnné.

(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) :**

$$Min\, e'e = Min(y - Xb)'(y - Xb) \text{ neboli v pozorovaných hodnotách}$$

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

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

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

tuto soustavu řešíme úpravami (vydělením 2, přeskupením členů) **na tvar** $X'Xb = X'y$, **která má řešení pro** b **ve tvaru**

(17) $b = (X'X)^{\ 1} X'y$ **, ← řešení je jednoznačné, neboť vzhledem k předpokladů** $h(X) = k$**, existuje (jediná) inverzní matice k matici** $X'X$**.**

Poznámka **Vyjádření minimalizace v pozorovaných hodnotách:**

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

Derivací podle (kterékoliv) pevně zvolené složky vektoru parametrů b_j dostaneme[3]

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

a odtud

(19) $$\sum_{t=1}^{T} \sum_{i=1}^{k} x_{ti} x_{tj} b_j = \sum_{t=1}^{T} x_{tj} y_t \quad \text{neboli} \quad b_j = \left(\sum_{t=1}^{T} \sum_{i=1}^{k} x_{ti} x_{tj} \right)^{-1} \sum_{t=1}^{T} x_{tj} y_t \ , \ j = 1,2,...,k$$

V případě konečných sumací můžeme přeskupovat členy v součtech

---

[3] Zřejmě $\dfrac{\partial \left(\sum_{i=1}^{k} x_{ti} b_i \right)}{\partial b_j} = x_{tj}$ a podobně $\dfrac{\partial \sum_{t=1}^{T} \left(\sum_{i=1}^{k} x_{ti} b_i \right)\left(\sum_{i=1}^{k} x_{ti} b_i \right)}{\partial b_j} = 2\sum_{t=1}^{T} \sum_{i=1}^{k} x_{ti} x_{tj} b_j$

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

$$\begin{pmatrix}\hat\beta_1\\ \hat\beta_2\end{pmatrix}=\begin{pmatrix}T & \sum x_{t2}\\ \sum x_{t2} & \sum x_{t2}^2\end{pmatrix}^{-1}\cdot\begin{pmatrix}\sum y_t\\ \sum x_{t2}y_t\end{pmatrix}=\frac{1}{D}\cdot\begin{pmatrix}\sum x_{t2}^2 & -\sum x_{t2}\\ -\sum x_{t2} & T\end{pmatrix}\begin{pmatrix}\sum y_t\\ \sum x_{t2}y_t\end{pmatrix},$$ kde

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

(20) $$\hat\beta_1=\frac{(\sum x_{t2}^2)(\sum y_t)-(\sum x_{t2})(\sum x_{t2}y_t)}{T.\sum x_{t2}^2-(\sum x_{t2})^2}\qquad \hat\beta_2=\frac{T.(\sum x_{t2}y_t)-(\sum x_{t2}).(\sum y_t)}{T.\sum x_{t2}^2-(\sum x_{t2})^2}$$

hodnota úrovňové konstanty — 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 ${}_{OLS}\hat\beta$ pořízený obyčejnou (prostou) metodou nejmenších čtverců je nejlepším nestranným lineárním odhadem vektoru parametrů β .**

Důkaz rozdělíme jej na několik částí.

A. odhad ${}_{OLS}\beta^*$ 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=Xβ+ε

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

nestochastičnost X — Eε=0 — Eβ=β (β nestochastický vektor)

Důsledek A Nestranná odhadová funkce je vždy asymptoticky nestranná.
Platí-li totiž $E({}_{(T)}\hat\beta)=\beta$ [4] pro každé konečné T, platí tentýž vztah i pro $T\to\infty$.

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

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

Ověření konzistence: lze vyvodit z následujícího tvrzení

Tvrzení **Jestliže je odhadová funkce asymptoticky nestranná a kovarianční matice této odhadové funkce konverguje při $T\to\infty$ 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**

[4] Symbolem nalevo rozumíme střední hodnotu odhadnutého vektoru parametrů β spočteného na základě T pozorování.

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}.[5] Pro tento prvek platí

$$\left|P_{ij}\right| = \left|\frac{1}{T}.\sum_{t=1}^{T} x_{ti}.x_{tj}\right| \le \xi_{ij}^{\,2} \;,\; \text{přičemž tato rovnost platí pro všechna konečná } T \,.$$

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

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

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

$$\underset{T\to\infty}{p\ lim}\ Cov({}_{OLS}\hat{\beta}) = \underset{T\to\infty}{p\ lim}\ \sigma_\varepsilon^{\,2}(X'X)^{-1} = \sigma_\varepsilon^{\,2}\ \underset{T\to\infty}{p\ lim}\ P^{-1}/T = 0_k \,,$$

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

C. odhad ${}_{OLS}\hat{\beta}$ 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

$$Cov(\hat{\beta}) = \sigma_\varepsilon^{\,2}.(X'X)^{-1}$$

(nestrannost ^β)

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

(dosazení ^β= (X'X)$^{-1}$ Xy) (dosazení za y=Xβ+ε)

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

(X′X)$^{-1}$ X′X = I$_T$ (vyrušení β)

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

(transpozice) vytknutí nestoch..členů před E) Eε.ε' = $\boldsymbol{\sigma_\varepsilon^2 I_T}$

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

(vytknutí skaláru $\boldsymbol{\sigma_\varepsilon^2}$) (X′X)$^{-1}$ X′X = I$_T$

$$= (X'X)^{-1}X'\sigma_\varepsilon^{\,2}I_TX.(X'X)^{-1} = \sigma_\varepsilon^{\,2}(X'X)^{-1}X'X.(X'X)^{-1} = \sigma_\varepsilon^{\,2}(X'X)^{-1} \qquad \square\,.$$

Odtud plyne **důsledek D1.**

---

[5] Účelem je ukázat, že matice P má prvky o konečné velikosti.

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

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

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).

E. odhad $_{OLS}\hat{\beta}$ **je nejlepší** ve smyslu "minimální" kovarianční matice $Cov(\hat{\beta})$, neboť pro kovarianční matici kterékoliv jiné (*lineární*) odhadové funkce $\tilde{\beta}$ platí :

$$Cov(\tilde{\beta}) - Cov(\hat{\beta}) = \Omega$$

kde Ω je nějaká symetrická pozitivně semidefinitní matice řádu ***k*** (rozměrů k x k).

Ověření Bez újmy na obecnosti můžeme matici D' jiné lineární odhadové funkce

$\tilde{\beta} = D'y$ vyjádřit ve tvaru $D' = (X'X)^{-1}X' + G'$

Poznámka Vzhledem k požadavku na nestrannost $\tilde{\beta}$ musí s ohledem na platnost vztahu $D'X = (X'X)^{-1}X'X + G'X = 1_k$, vždy platit $G'X = 0$.

Ověření vydatnosti

$$\beta = E(\tilde{\beta}) = E(D'y) = E[D'(X\beta + \varepsilon)] = ED'X\beta + ED'\varepsilon = D'X\beta + 0 = D'X\beta$$

[matice D, X jsou nestochastické]

z čehož přímo plyne $D'X = 1_k$

Pro libovolnou jinou (*lineární* a *nestrannou*) odhadovou funkci tedy musí platit

$$Cov(\tilde{\beta}) = E\left\{[\tilde{\beta} - E\tilde{\beta}][\tilde{\beta} - E\tilde{\beta}]'\right\} = E\left\{[D'y - \beta][D'y - \beta]'\right\} =$$

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

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

[protože G´X = 0] [protože G´X = 0]

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

[protože $(X'X)^{-1}X'X = I$] [protože $(X'X)^{-1}X'X = I$]

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

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

$$= E\left\{(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1} + G'\varepsilon\varepsilon'G + (X'X)^{-1}X'\varepsilon\varepsilon'G + G'\varepsilon\varepsilon'X(X'X)^{-1}\right\} =$$

[po uplatnění operátoru střední hodnoty a protože X i G jsou nestochastické matice]

$$= (X'X)^{-1}X'E\varepsilon\varepsilon'X(X'X)^{-1} + G'E\varepsilon\varepsilon'G + (X'X)^{-1}X'E\varepsilon\varepsilon'G + G'E\varepsilon\varepsilon'X(X'X)^{-1} =$$

[protože $E\varepsilon'\varepsilon = \sigma^2 I_T$]

$$= (X'X)^{-1}X'\sigma^2 I_T X(X'X)^{-1} + G'\sigma^2 I_T G + (X'X)^{-1}X'\sigma^2 I_T G + G'\sigma^2 I_T X(X'X)^{-1} =$$

[protože σ^2 je skalární hodnota]

$$= \sigma^2(X'X)^{-1} + \sigma^2 G'G + \sigma^2(X'X)^{-1}X'G + \sigma^2 G'X(X'X)^{-1} =$$

[protože $G'X = 0$ a stejně $X'G = 0$]

$$= \sigma^2(X'X)^{-1} + \sigma^2 G'G = Cov(\hat{\beta}) + \sigma^2 G'G$$

kde $G'G$ je zřejmě pozitivně semidefinitní matice[6] a $\sigma^2 > 0$ □

F. pro odhad rozptylu reziduí dostaneme

$$E(e'e) = E(\varepsilon'M.M\varepsilon) = E(\varepsilon'M.\varepsilon) = Etr(\varepsilon'M.\varepsilon) = Etr(M\varepsilon\varepsilon') =$$

(M je idempotentní matice) (skalár je současně svou stopou) (tr A.B = tr B.A)

$$= trE(M\varepsilon\varepsilon') = tr(ME\varepsilon\varepsilon') = tr(M\sigma_\varepsilon^2 I_T) = \sigma_\varepsilon^2 tr(M) = \sigma_\varepsilon^2(T-k)$$

(záměna stopy a střední hodnoty) (M je nestochastická) (σ^2 je skalární hodnota)

protože

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

(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

je výraz $\hat{\sigma}_\varepsilon^2 = \dfrac{e\acute{e}}{T-k}$, neboť zřejmě platí $E\left[\dfrac{\sum_{}^{T} e_t^2}{T-k}\right] = \sigma_\varepsilon^2$.

Z **důsledku D1** je patrné, že
odhady směrodatných odchylek odhadnutých regresních parametrů $s_{\hat{\beta}}$ získáme jako

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

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

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

Tvrzení 1 Platí vztah $\mathbf{X'e = 0}$, tj. rezidua jsou nekorelovaná s kteroukoliv z vysvětlujících proměnných.

Ověření: zřejmě platí $\mathbf{X'e = X'y - X'\hat{y} = X'y - X'Xb = X'y - X'X(X'X)^{-1}X'y = 0}$

---

[6] 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).

Důsledek: Platí $\sum_{i=1}^{T} e_t = 0$, tzn. součet všech **reziduí je nulový, tj. regresní přímka/nadrovina prochází středem pozorovaných hodnot.**

Ověření: protože platí $X'e = 0$, stačí v matici X uvažovat jeden z jejich sloupců jako jedničkový vektor (nemusí to být nutně ten první)**. Pak máme**

$$X'e = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{T1} \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ x_{1k} & x_{2k} & \dots & x_{Tk} \end{pmatrix} \cdot \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_T \end{pmatrix} = 0$$ a je zřejmě vidět (z druhého řádku), že platí $\sum_{t=1}^{T} e_t = 0$

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).