

4EK211 Základy ekonometrie

ZS 2016/17 Cvičení 2: Metoda nejmenších čtverců



LENKA FIŘTOVÁ

KATEDRA EKONOMETRIE, FAKULTA INFORMATIKY A STATISTIKY

VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE

1. Doplnění a opakování z minula

Jaká je pravděpodobnost, že náhodná veličina z normovaného normálního rozdělení bude:

- 1) menší než 0,5
- 2) větší než 1
- 3) větší než 1,5 nebo menší než -1,5

Pro jakou hodnotu z normovaného normálního rozdělení platí, že:

- 1) napravo od ní leží 2,5 % hodnot
- 2) nalevo od ní leží 30 % hodnot

1. Doplnění a opakování z minula

Jaká je pravděpodobnost, že náhodná veličina z normovaného normálního rozdělení bude:

1) menší než 0,5	0,69
2) větší než 1	0,15
3) větší než 1,5 nebo menší než 1,5	0,134

Pro jakou hodnotu z normovaného normálního rozdělení platí, že:

1) napravo od ní leží 2,5 % hodnot	1,96
2) nalevo od ní leží 30 % hodnot	-0,52

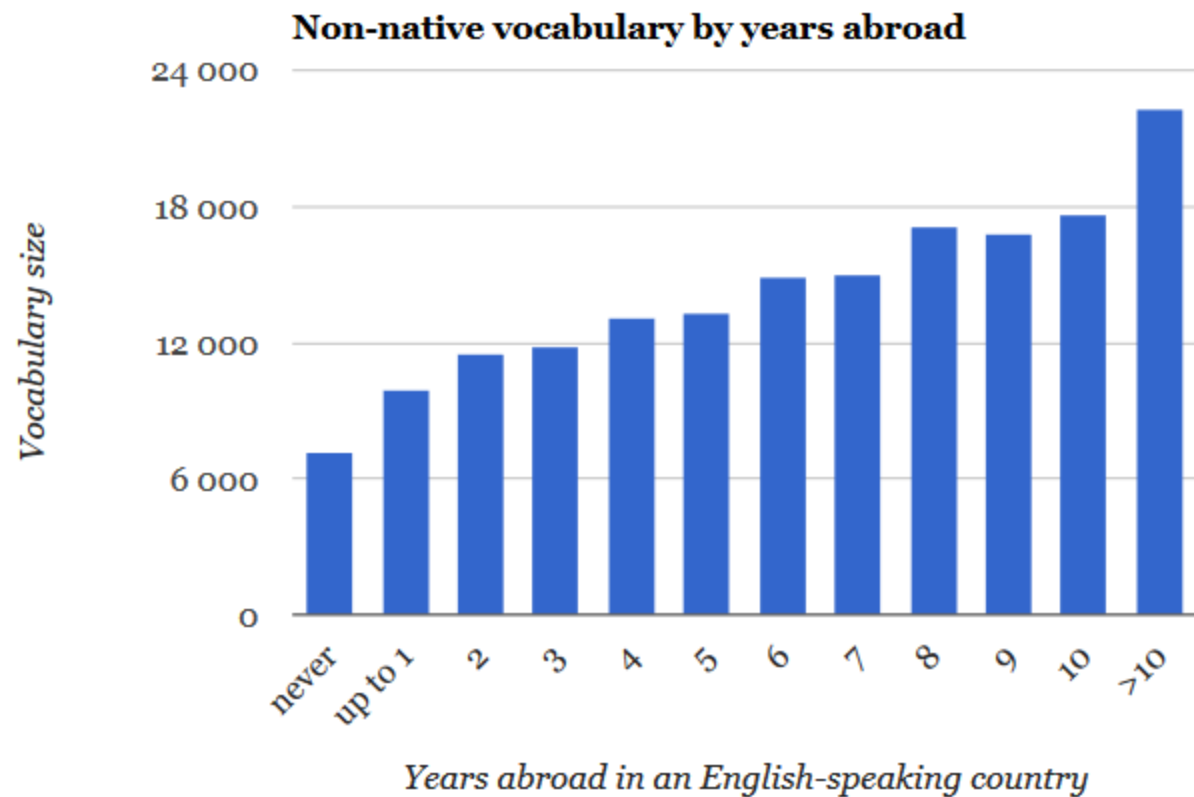
2. Metoda nejmenších čtverců

Zkuste si, kolik znáte anglických slov:

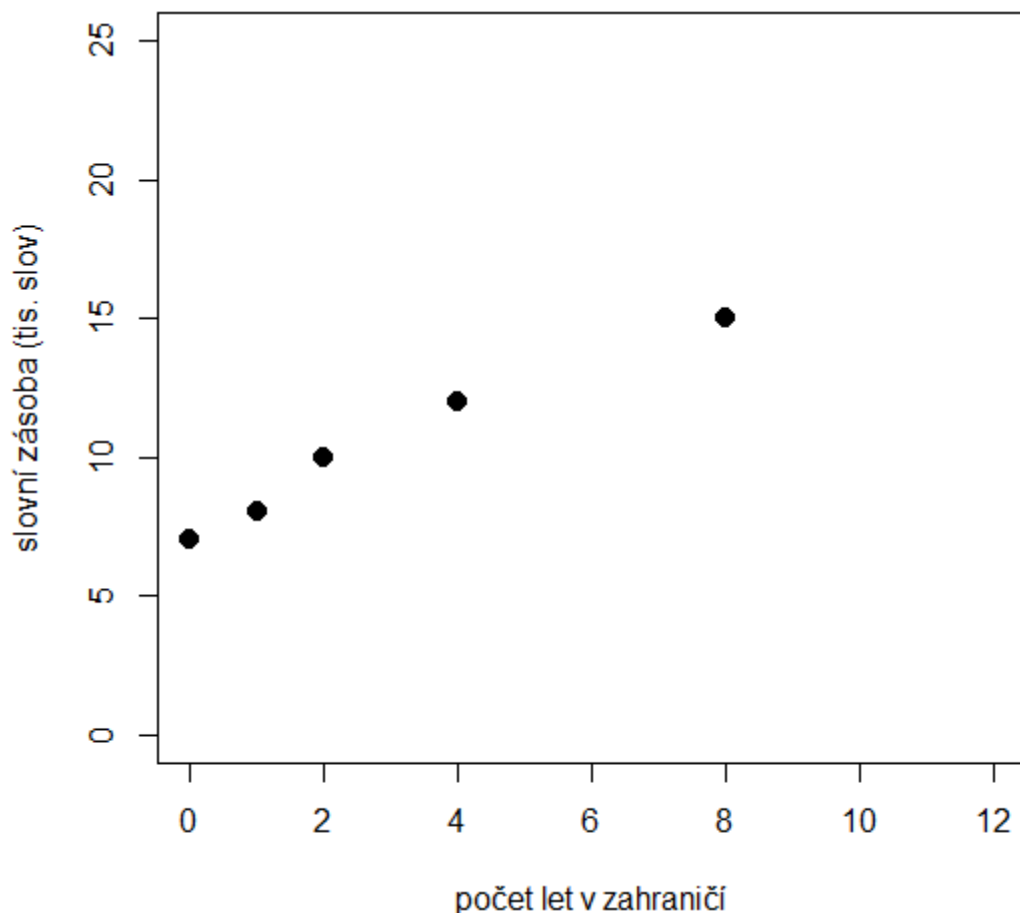
<http://testyourvocab.com>

(ale až doma).

Na tomto webu jsou i zajímavé statistiky, mimo jiné že existuje vztah mezi počtem slov, která známe, a dobou, kterou jsme strávili v zahraničí (což je vcelku očekávatelné).



2. Metoda nejmenších čtverců



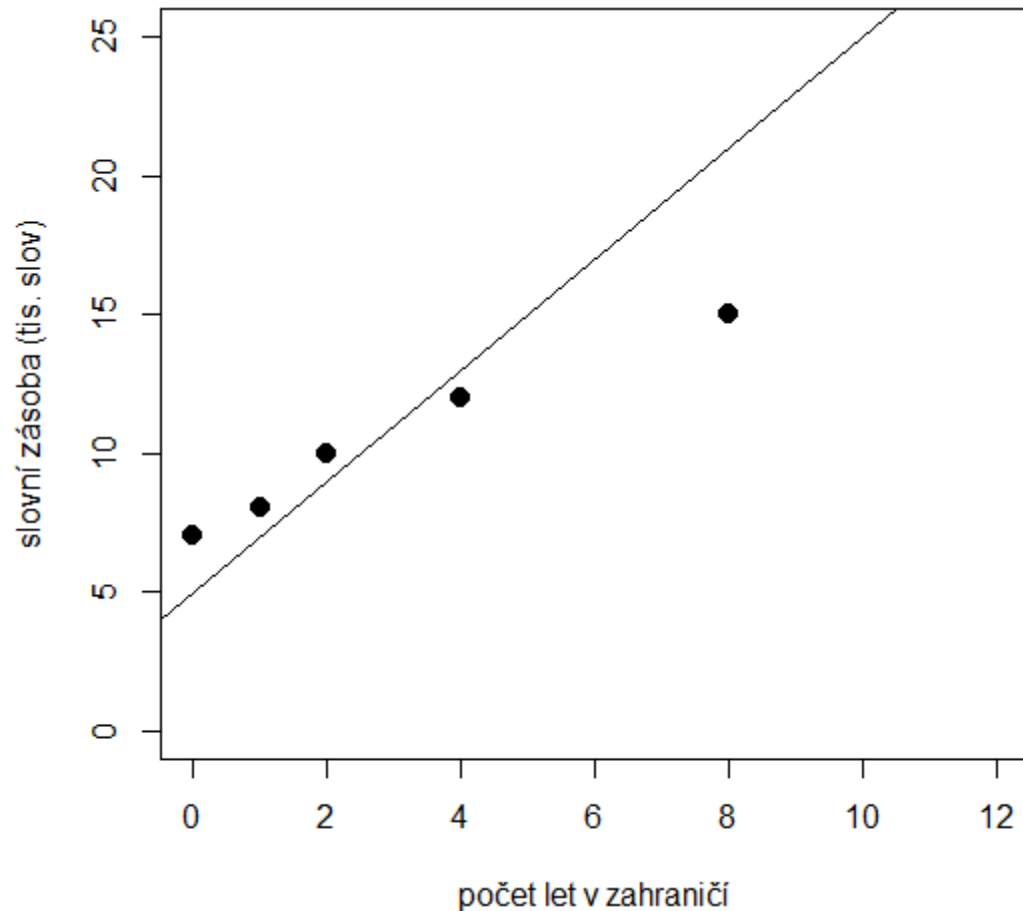
počet let v zahraničí	0	1	2	4	8
slovní zásoba (tis. slov)	7	8	10	12	15

Příklad není moc reálný (malý počet dat, fiktivní data), slouží k pochopení principu MNČ.

Chceme daty proložit přímkou, která nejlépe zachycuje vztah mezi y a x .

Jak na to?

2. Metoda nejmenších čtverců

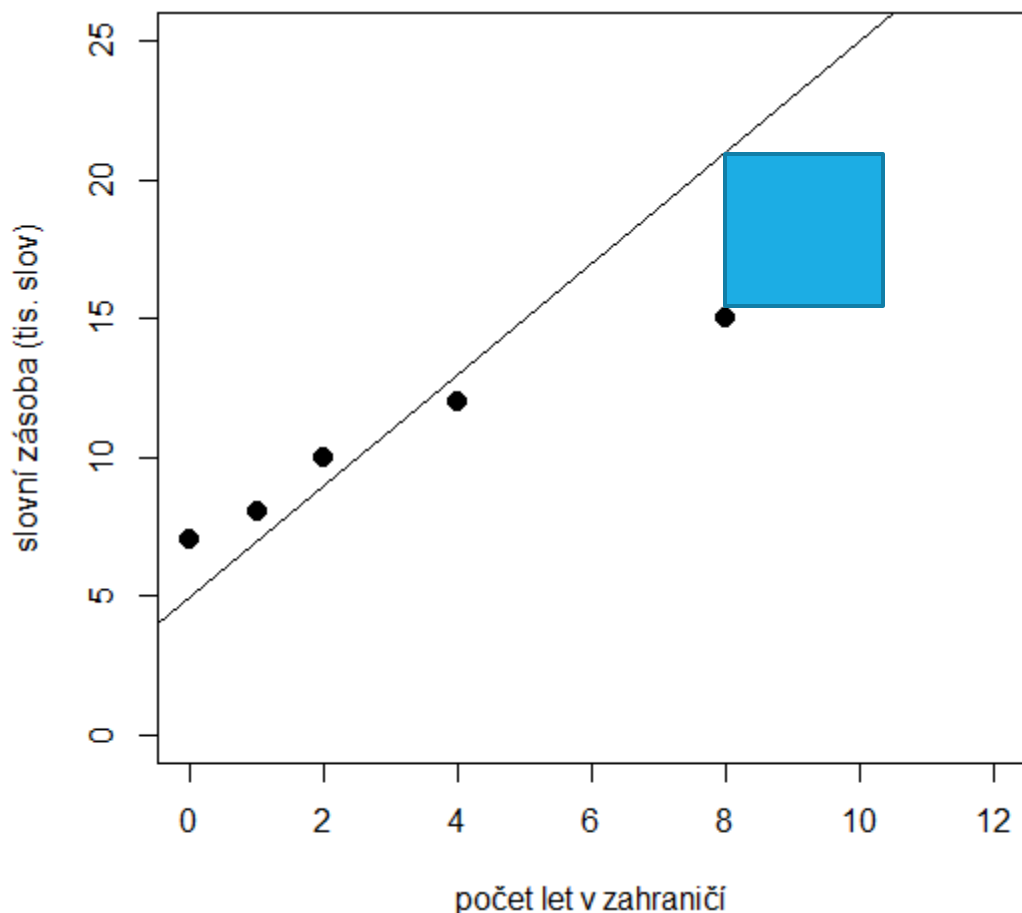


Zkusme třeba následující přímku:

$$y = 5 + 2x$$

Jak zjistíme, jestli je to „dobrá“
přímka?

2. Metoda nejmenších čtverců



Jednou z možností je sečíst druhé mocniny vzdáleností bodů od dané přímky (čtvercové chyby).

Součet čtvercových chyb by měl být co možná nejmenší.

Jaký je v tomto případě součet čtvercových chyb?

2. Metoda nejmenších čtverců

x_i	Skutečné hodnoty	Vyrovnané hodnoty	Čtvercová chyba
0	7	5	4
1	8	7	1
2	10	9	1
4	12	13	1
8	15	21	36
			43

Existuje nějaká „lepší“ přímka?

2. Metoda nejmenších čtverců

Hledáme parametry „nejlepší přímky“: $y_i = b_0 + b_1x_i$

$$SS(b_0, b_1) = (y_1 - b_0 - b_1x_1)^2 + (y_2 - b_0 - b_1x_2)^2 + (y_3 - b_0 - b_1x_3)^2 + (y_4 - b_0 - b_1x_4)^2 + (y_5 - b_0 - b_1x_5)^2 \rightarrow \text{minimalizujeme}$$

Položíme parciální derivace funkce SS podle b_0 a b_1 rovny 0 a řešíme soustavu.

2. Metoda nejmenších čtverců

Výsledek:

$$b_0 = 7,4$$

$$b_1 = 1$$

„Nejlepší“ přímka: $y_i = 7,4 + 1x_i$

Interpretujte tyto parametry.

2. Metoda nejmenších čtverců

Další úkoly:

- 1) Spočítejte vyrovnané hodnoty
- 2) Spočítejte rezidua. Co jsou to rezidua? Co je to náhodná složka?
Jaký je rozdíl mezi rezidui a náhodnou složkou?
- 3) Kolik slov bude znát v průměru člověk, který žil v zahraničí 5 let?

2. Metoda nejmenších čtverců

Spočítejte

- 1) Nevysvětlený součet čtverců
- 2) Celkový součet čtverců
- 3) Vysvětlený součet čtverců
- 4) Koeficient determinace

2. Metoda nejmenších čtverců

$$RSS = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$$

$$TSS = \sum (y_i - \bar{y})^2,$$

$$ESS = \sum (\hat{y}_i - \bar{y})^2$$

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

2. Metoda nejmenších čtverců

$$RSS = 1,2$$

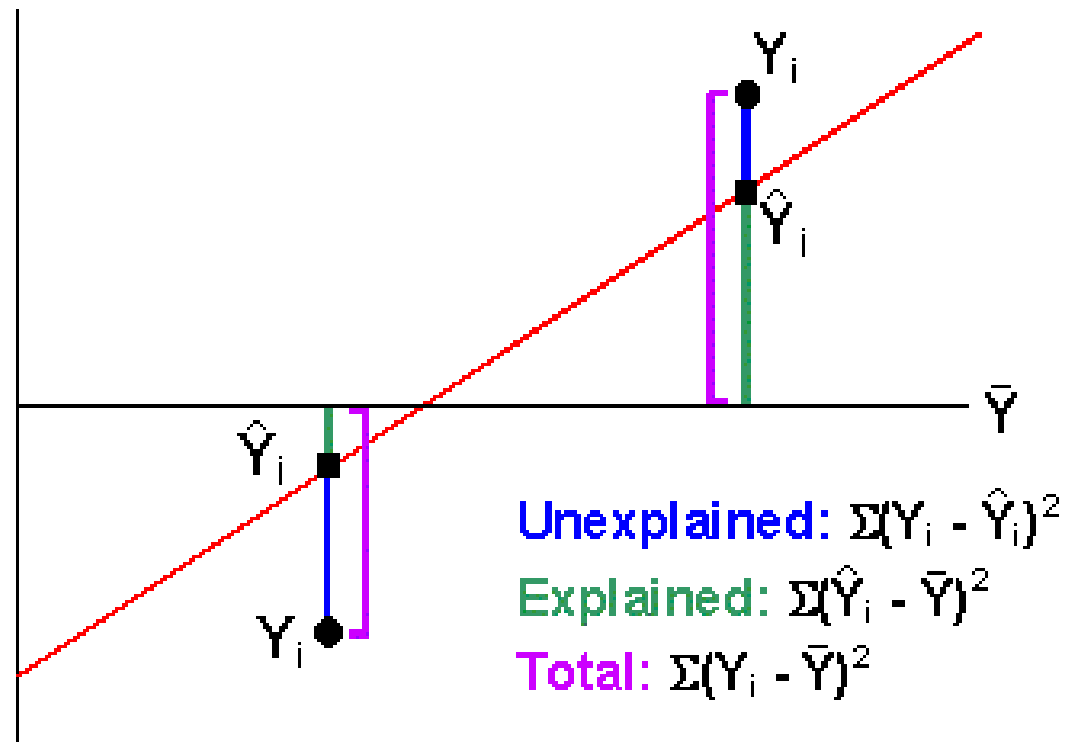
$$TSS = 41,2$$

$$ESS = 40$$

$$R^2 = 0,9709$$

Interpretujte koeficient determinace, RSS , TSS a ESS . Jakých hodnot může R^2 nabývat?

2. Metoda nejmenších čtverců



http://labs.geog.uvic.ca/geog226/images/Lab9/img_variation.gif

2. Metoda nejmenších čtverců

Zopakujte si práci s maticemi:

Uvažujte matice:

$$A = \begin{pmatrix} 1 & 3 \\ 5 & 4 \end{pmatrix}$$

$$B = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 3 & 1 \end{pmatrix}$$

Určete hodnotu uvedených matic.

Co je to regulární matice? Které z matic A , B , C jsou regulární?

Ke kterým z nich existuje inverzní matice? (jen čtvercové regulární)

Zapište matici A' .

Vynásobte, pokud to lze: AB , CA .

2. Metoda nejmenších čtverců

Najděte obecné vyjádření pro takové hodnoty parametrů b_0 , b_1 , které minimalizují celkový součet čtverců, ale použijte tentokrát maticové vyjádření.

Měli byste dojít k:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y})$$

V Excelu spočítejte pomocí tohoto vzorce odhady parametrů b_0 , b_1 .

Na doma... Co byste měli umět

Jaký je princip metody nejmenších čtverců?

Jak se interpretují odhadnuté regresní koeficienty?

Co je to vysvětlený, nevysvětlený a celkový součet čtverců a jaký je mezi nimi vztah?

Co je to koeficient determinace, jak ho získáme?

Co jsou to rezidua?