

CVIČENIA Z EKONOMETRIE 2006/2007

PÍ SOMKA 1

TERMÍN ODOVZDANIA: 27.4.2007

Pokyny:

- *Riešenia napíšte na jeden papier, resp. do jedného pdf súboru).*
 - *Súbor pošlite na bs.ulohy@gmail.com so subjectom *ekonometria - písomka 1 - vase priezvisko*. Riešenie napísané na papieri môžete nechať na M204, M266 alebo na sekretariáte M270 v priečinku. Ak písomku odovzdáte na papieri, pošlite mail na túto istú adresu, v ktorom napíšete, kedy a kde ste riešenie odovzdali.*
 - *Neodpisujte. Na rozdiel od úloh, toto je písomka, takže nie sú povolené ani konzultácie so spolužiakmi. Používať môžete poznámky z prednášok a cvičení, knihy.*
 - *Pri úlohách, ktoré vyžadujú prácu na počítači, uveďte použitý softvér.*
 - *Vo všetkých modeloch predpokladáme, že pre náhodnú zložku ε platí $\varepsilon \sim N(0, \sigma^2 I)$.*
-

1. (4 body) Z dát zo súboru `domy.txt` vytvorte regresný model s aspoň dvoma vysvetľujúcimi premennými pre cenu domu a odhadnite jeho parametre metódou najmenších štvorcov.

(a) Testovanie signifikancie parametrov

- Vysvetlite, aká hypotéza sa testuje. Kedy je parameter signifikantný - ak sa táto hypotéza zamietá alebo ak sa nezamietá?
- Ako sa vypočíta testovacia štatistika?
- Aké je rozdelenie štatistiky za platnosti nulovej hypotézy a aký je obor zametnutia?
- Testujte signifikanciu parametrov vo vašom modeli.

(b) Testovanie signifikancie regresie

- Vysvetlite, aká hypotéza sa testuje. Kedy je regresia signifikantná - ak sa táto hypotéza zamietá alebo ak sa nezamietá?
- Ako sa vypočíta testovacia štatistika?
- Aké je rozdelenie štatistiky za platnosti nulovej hypotézy a aký je obor zametnutia?
- Testujte signifikanciu regresie vo vašom modeli.

(c) Sformulujte model, ktorý je submodelom vášho modelu.

- Vysvetlite, prečo sa testovanie tohto submodelu dá chápať ako testovanie lineárnej hypotézy o parametroch pôvodného modelu.
- Ako testujeme lineárnu hypotézu $R\beta = r$ o parametroch regresného modelu? Uveďte dva možné výpočty F štatistiky: pomocou R a r a pomocou reziduálnych súm štvorcov v modeli bez reštrikcie a v modeli s reštrikciou. Aké je rozdelenie štatistiky za platnosti nulovej hypotézy a aký je obor zametnutia?
- Vyberte si jeden z uvedených výpočtov a testujte F testom váš submodel.
- Aké iné kritériá okrem F testu môžeme použiť pri výbere modelu? Použite ich na výber jedného z vašich dvoch modelov - pôvodného modelu a submodelu. Pomocou toho modelu, ktorý označíte za lepší, vypočítajte odhad ceny domu s nasledovnými parametrami:

$$x_1 = 5.9894, x_2 = 1, x_3 = 5.5200, x_4 = 1.2560,$$

$$x_5 = 2, x_6 = 6, x_7 = 3, x_8 = 40, x_9 = 1.$$

2. (1 bod) Vyberte si a vyriešte jednu z nasledujúcich úloh.

(a) Z tých istých dát (je ich 60) sme odhadli dva modely:

$$\begin{aligned}y &= a_0 + a_1x + a_2x^2 + a_3x^3 + \varepsilon, \\y &= b_0 + b_1x^2 + \eta.\end{aligned}\tag{1}$$

Koeficient determinácie v prvej regresii je $R^2 = 0.9412$. V druhej regresii je $R^2 = 0.9394$. Testujte hypotézu, že v modeli (1) je $a_1 = 0$, $a_3 = 0$.

(b) Z tých istých dát (je ich 110) sme odhadli dva modely:

$$\begin{aligned}y &= a_0 + a_1x + a_2x^2 + a_3x^3 + \varepsilon, \\y &= b_0 + b_1x^2 + \eta.\end{aligned}\tag{2}$$

Reziduálna suma štvorcov v prvej regresii je $RSS = 204.54$, v druhej regresii je $RSS = 224.95$. Testujte hypotézu, že v modeli (2) je $a_1 = 0$, $a_3 = 0$.

(c) Odhadnutím regresnej priamky z 80 dát sme dostali nasledovný výstup (v zátvorkách sú uvedného odhady štandardných odchýlok):

$$\begin{aligned}\hat{y} &= 5.664 - 10.571x \\ &\quad (1.021) \quad (2.468)\end{aligned}$$

Určte koeficient determinácie R^2 v tomto modeli.

3. (1 bod) Priemer rezíduí v regresnom modeli

(a) Ukážte, že v modeli

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_{k-1} x_{k-1} + \varepsilon,$$

ktorý odhadneme metódou najmenších štvorcov je priemer rezíduí nulový.

(b) Uvažujme teraz model bez absolútneho člena, t.j.

$$y = \beta_1 x_1 + \dots + \beta_{k-1} x_{k-1} + \varepsilon.$$

a jeho odhad metódou najmenších štvorcov. Ukážte na konkrétnom príklade, že súčet rezíduí nemusí byť nulový (stačí príklad pre $k = 2$ s malým počtom dát (napr. 3 alebo 4), ktorý sa dá zrátať ručne) a vysvetlite, prečo sa dôkaz z časti (a) nedá použiť.

4. (1 bod) Doplňte vynechané hodnoty:

Dependent Variable: Y43
 Method: Least Squares
 Date: 04/07/06 Time: 12:59
 Sample: 1 50
 Included observations: 50
 Y43=C(1)+C(2)*X

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	330.4089	<input type="text"/>	9.076391	0.0000
C(2)	<input type="text"/>	12.42420	19.64899	0.0000
R-squared	<input type="text"/>	Mean dependent var	-292.1046	
Adjusted R-squared	<input type="text"/>	S.D. dependent var	377.3411	
S.E. of regression	126.7786	Akaike info criterion	12.58194	
Sum squared resid	771494.9	Schwarz criterion	12.63842	
Log likelihood	-312.0485	F-statistic	<input type="text"/>	
Durbin-Watson stat	1.354564	Prob(F-statistic)	0.000000	

5. (3 body) Uvažujme dva modely:

$$y = a_0 + a_1x_1 + \dots + a_{k-2}x_{k-2} + a_{k-1}x_{k-1} + \varepsilon \quad (3)$$

$$y = b_0 + b_1x_1 + \dots + b_{k-2}x_{k-2} + \eta \quad (4)$$

Dokážte, že model (4) má väčší upravený koeficient determinácie ako model (3) práve vtedy, keď absolútna hodnota t-štatistiky na testovanie hypotézy $a_{k-1} = 0$ v modeli (3) je menšia ako 1.