

Časť I

1. Odhad metódou najmenších štvorcov

Odhadujeme model

$$Y = X\beta + \varepsilon,$$

pričom matica X má rozmer $n \times k$ a má lineárne nezávislé stĺpce, vektor ε má nulovú strednú hodnotu a kovariančnú maticu $\sigma^2 I$.

- Ako vypočítame odhad parametra β metódou najmenších štvorcov?
- Dokážte jeho nevychýlenosť.
- Aká je jeho kovariančná matica? Ako odhadujeme túto kovariančnú maticu?
- Aké je pravdepodobnostné rozdelenie odhadu, ak o vektore ε predpokladáme, že má normálne rozdelenie? Kde potrebujeme tento predpoklad?

2. Signifikancia parametrov

- Sformulujte nulovú hypotézu, ktorou testujeme signifikantnosť parametrov. Kedy je parameter signifikantný - ak H_0 zamietame alebo ak H_0 nezamietame?
- Akou štatistikou túto hypotézu testujeme - ako sa vypočíta a aké má rozdelenie, ak platí H_0 ?
- Aká je kritická oblasť testu, t.j. kedy H_0 zamietame?
- Odhadujeme model

$$Y = C(1) + C(2)x + C(3)x^2 + C(4)x^3 + \varepsilon,$$

pričom ε má $N(0, \sigma^2 I)$ rozdelenie. Dostaneme:

```
Dependent Variable: Y10
Method: Least Squares
Date: 03/18/06 Time: 16:08
Sample: 1 101
Included observations: 101
Y10=C(1)+C(2)*X+C(3)*X^2+C(4)*X^3
```

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	144.8358	15.22934	9.510310	0.0000
C(2)	28.00946	8.708314	3.216405	0.0018
C(3)	-44.53205	1.335532	-33.34405	0.0000
C(4)	0.102766	0.521756	0.196961	0.8443
R-squared	0.924280	Mean dependent var		-233.6867
Adjusted R-squared	0.921938	S.D. dependent var		365.1702
S.E. of regression	102.0270	Akaike info criterion		12.12715
Sum squared resid	1009722.	Schwarz criterion		12.23072
Log likelihood	-608.4211	F-statistic		394.6779
Durbin-Watson stat	2.152210	Prob(F-statistic)		0.000000

Ktoré parametre sú signifikantné na hladine významnosti $\alpha = 0,05$?

3. Signifikancia regresie

- (a) Sformulujte nulovú hypotézu, ktorou testujeme signifikantnosť regresie. Kedy je regresia signifikantná - ak H_0 zamietame alebo ak H_0 nezamietame?
- (b) Akou štatistikou túto hypotézu testujeme - ako sa vypočíta a aké má rozdelenie, ak platí H_0 ?
- (c) Aká je kritická oblasť testu, t.j. kedy H_0 zamietame?
- (d) Je regresia z príkladu 2(d) signifikantná?

4. Intervaly spoľahlivosti (intervalové odhady)

Vypočítajte 95% interval spoľahlivosti pre každý parameter z regresie z príkladu 2(d).

5. Testovanie jednej lineárnej hypotézy o parametroch

Predpokladajme, že v modeli $Y = X\beta + \varepsilon$ testujeme hypotézu $c^T\beta = r$.

- (a) Akou štatistikou túto hypotézu testujeme - ako sa vypočíta a aké má rozdelenie, ak platí H_0 ?
- (b) Aká je kritická oblasť testu, t.j. kedy H_0 zamietame?
- (c) V modeli z príkladu 2(d) testujte hypotézu, že $C(2) = 20$.

6. Testovanie lineárnych hypotéz o parametroch

Predpokladajme, že v modeli $Y = X\beta + \varepsilon$ testujeme hypotézu $R\beta = r$.

- (a) Akou štatistikou túto hypotézu testujeme - ako sa vypočíta a aké má rozdelenie, ak platí H_0 ?
- (b) Aká je kritická oblasť testu, t.j. kedy H_0 zamietame?

7. Rozhodnite, či sú nasledujúce tvrdenia pravdivé. Svoje tvrdenie zdôvodnite.

- (a) Ak do regresného modelu pridáme ďalšiu vysvetľujúcu premennú, koeficient determinácie R^2 sa nemôže zmenšiť.
- (b) Ak do regresného modelu pridáme ďalšiu vysvetľujúcu premennú, upravený koeficient determinácie \bar{R}^2 sa nemôže zmenšiť.
- (c) Parameter je signifikantný na hladine významnosti α práve vtedy, keď $(1 - \alpha)\%$ interval spoľahlivosti pre tento parameter neobsahuje nulu.
- (d) V modeli $Y = C(1) + C(2)x + \varepsilon$ je signifikancia regresie je ekvivalentná so signifikantnosťou parametra $C(2)$.

Časť II

1. Doplňte vynechané hodnoty¹:

Dependent Variable: Y23
 Method: Least Squares
 Date: 04/07/06 Time: 13:04
 Sample: 1 75
 Included observations: 75
 Y23=C(1)+C(2)*X+C(3)*X^2

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	164.9689	16.56590		0.0000
C(2)	31.91071		3.407309	0.0011
C(3)		2.889207	-15.36713	0.0000
R-squared		Mean dependent var		-159.6316
Adjusted R-squared		S.D. dependent var		385.8950
S.E. of regression	104.8582	Akaike info criterion		12.18227
Sum squared resid	791657.9	Schwarz criterion		12.27497
Log likelihood	-453.8353	F-statistic		465.1122
Durbin-Watson stat	2.001109	Prob(F-statistic)		0.000000

2. 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		9.076391	0.0000
C(2)		12.42420	19.64899	0.0000
R-squared		Mean dependent var		-292.1046
Adjusted R-squared		S.D. dependent var		377.3411
S.E. of regression	126.7786	Akaike info criterion		12.56194
Sum squared resid	771494.9	Schwarz criterion		12.63842
Log likelihood	-312.0485	F-statistic		
Durbin-Watson stat	1.354564	Prob(F-statistic)		0.000000

3. Nájdite nutnú a postačujúcu podmienku, ktorú musia spĺňať x_i , aby odhady $\hat{\beta}_0$, $\hat{\beta}_1$ parametrov modelu $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ boli nekorelované.

¹V príkladoch tohto typu môžu byť výpočty, v ktorých treba poznať význam hodnôt *Included observations*, *Coefficient*, *Std. Error*, *t-statistic*, *R squared*, *Adjusted R squared*, *F statistics* a vzťahy medzi nimi

4. (a) Dokážte, že v modeli

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

sa koeficient determinácie R^2 rovná druhej mocnine výberového korelačného koeficienta medzi x_i a Y_i .

- (b) Dokážte, že v modeli

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

sa koeficient determinácie R^2 rovná druhej mocnine výberového korelačného koeficienta medzi skutočnými hodnotami Y_i a odhadnutými hodnotami \hat{Y}_i .

5. V článku *Hiren Sarkar, M. Aynul Hasan: Impact of Corruption on the Efficiency of Investment: Evidence from a Cross - Country Analysis. Asia-Pacific Development Journal* je okrem iného uvedený aj tento výsledok:

The cross-country regression of efficiency of investment (EII) on the Corruption Perception Index (CPI) of 87 countries is estimated as follows.³

$$(1) \quad \text{EII} = -3.319920 + 1.127249 (\text{CPI}), \text{ R squared} = .06 \\ (\text{t} = 2.24)$$

The above result shows that when CPI increases or the level of corruption decreases there is a significant (with a probability of more than 95 per cent) increase in the efficiency of investment. Although a low explanatory power of a single variable cross-section regression is not unexpected because of a large number of missing variables, the present sample of 87 countries and areas needs special explanation.

Kvôli nízkej hodnote R^2 môže vzniknúť pochybnosť o signifikancii regresie. Zistite, či je regresia signifikantná.

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

$$Y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \varepsilon, \quad (1)$$

$$Y = b_0 + b_1 x^2 + \eta. \quad (2)$$

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 $a_1 = 0$, $a_3 = 0$.