

Ekonomietrian tilastolliset menetelmät

Laskuharjoitus 5

2012, Helmikuu 16

1 Laskutehtäviä

1. Olkoon

$$\mathcal{Y} = \mathcal{X}_1\beta_1 + \mathcal{X}_2\beta_2 + \epsilon,$$

missä $\mathcal{Y} = (Y_1, \dots, Y_n)'$ ja $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ ovat $n \times 1$ -satunnaisvektoreita, \mathcal{X}_1 on $n \times K_1$ -matriisi, \mathcal{X}_2 on $n \times K_2$ -matriisi, β_1 on $K_1 \times 1$ -vektori ja β_2 on $K_2 \times 1$ -vektori.

Olkoon b_{1*} rajoitettu pienimmän neliösumman estimaattori, rajoituksella $\beta_2 = \beta_2^0$. Osoita, että

$$b_{1*} = (\mathcal{X}_1'\mathcal{X}_1)^{-1}\mathcal{X}_1'(\mathcal{Y} - \mathcal{X}_2\beta_2^0).$$

2. (Luku 7.9. Testing Nonlinear Restrictions) Olkoon

$$\mathcal{Y} = \mathcal{X}\beta + \epsilon,$$

missä $\mathcal{Y} = (Y_1, \dots, Y_n)'$ ja $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ ovat $n \times 1$ -satunnaisvektoreita, $\mathcal{X} = (X_1, \dots, X_n)'$ on $n \times K$ -matriisi ja β on $K \times 1$ -vektori. Olkoon

$$b = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\mathcal{Y}.$$

- (a) Olkoon $K = 1$ ja testataan nollahypoteesia

$$H_0 : c(\beta) = q,$$

missä $c : \mathbf{R} \rightarrow \mathbf{R}$ on epälineaarinen funktio ja $q \in \mathbf{R}$. Käytetään testisuurena yleistettyä t -testisuuretta

$$z = \frac{c(b) - q}{\widehat{\text{sd}}(c(b))},$$

missä $\widehat{\text{sd}}(c(b))$ on $c(b)$:n keskihajonnan estimaatti. Anna järkevä lauseke $\widehat{\text{sd}}(c(b))$:lle.

Ohje: käytä approksimaatiota

$$c(b) \approx c(\beta) + c'(\beta)(b - \beta),$$

missä c' on c :n derivaatta.

(b) Testataan nollahypoteesia

$$H_0 : c(\beta) = q,$$

missä $c : \mathbf{R}^K \rightarrow \mathbf{R}$ on epälineaarinen funktio ja $q \in \mathbf{R}$. Käytetään testisuurena yleistettyä t -testisuuretta

$$z = \frac{c(b) - q}{\widehat{\text{sd}}(c(b))},$$

missä $\widehat{\text{sd}}(c(b))$ on $c(b)$:n keskihajonnan estimaatti. Anna järkevä lauseke $\widehat{\text{sd}}(c(b))$:lle.

Ohje: käytä approksimaatiota

$$c(b) \approx c(\beta) + Dc(\beta)'(b - \beta),$$

missä

$$Dc(\beta) = (\partial c(\beta)/\beta_1, \dots, \partial c(\beta)/\beta_K)'$$

on c :n gradientti.

3. Olkoon

$$Y_i = \beta_1 + \beta_2 X_i + \epsilon_i, \quad i = 1, \dots, n.$$

Oletetaan, että $(X_1, Y_1), \dots, (X_n, Y_n)$ ovat i.i.d, $0 < \text{Var}(X_i) < \infty$, $0 < \text{Var}(\epsilon_i) < \infty$ ja

$$\text{Cov}(X_i, \epsilon_i) \neq 0.$$

Olkoon

$$b_2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

Osoita, että

$$b_2 \xrightarrow{p} \beta_2 + c,$$

kun $n \rightarrow \infty$, missä $c \neq 0$.

4. Olkoon

$$\mathcal{Y} = \mathcal{X}\beta + \epsilon,$$

missä $\mathcal{Y} = (Y_1, \dots, Y_n)'$ ja $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ ovat $n \times 1$ -satunnaisvektoreita, $\mathcal{X} = (X_1, \dots, X_n)'$ on $n \times K$ -matriisi ja β on $K \times 1$ -vektori. Olkoon $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d., $E(\epsilon | \mathcal{X}) = 0$ ja $E(\epsilon\epsilon' | \mathcal{X}) = \sigma^2 I$.

Oletetaan tunnetuksi, että

$$\sqrt{n}(b - \beta) \xrightarrow{d} N(0, \sigma^2 Q^{-1}),$$

missä $b = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\mathcal{Y}$ ja $Q = EX_i X_i'$. Olkoon

$$R\beta = q,$$

missä R on $J \times K$ -matriisi kun q on J -vektori, missä $1 \leq J \leq K - 1$.

(a) Totea, että

$$\sqrt{n}(Rb - q) \xrightarrow{d} N(0, \sigma^2 RQ^{-1}R').$$

(b) Totea, että

$$\sigma^{-1}(RQ^{-1}R')^{-1/2}\sqrt{n}(Rb - q) \xrightarrow{d} N(0, I).$$

(c) Totea, että

$$W \xrightarrow{d} N(0, I),$$

missä

$$W = s^{-1}(R\hat{Q}^{-1}R')^{-1/2}\sqrt{n}(Rb - q),$$

$$\hat{Q} = \frac{1}{n} \sum_{i=1}^n X_i X_i'$$

ja

$$s^2 = \frac{1}{n - K} \sum_{i=1}^n (Y_i - X_i' b)^2.$$

(d) Totea, että

$$W'W \xrightarrow{d} \chi^2(J).$$

(e) Totea, että

$$W'W = JF,$$

missä

$$F = \frac{(Rb - q)'(s^2 R(\mathcal{X}'\mathcal{X})^{-1}R')^{-1}(Rb - q)}{J}.$$

2 Tietokonetehtäviä

U.S fertility data are taken from the 1980 Census. These data were provided by Professor William Evans of the University of Maryland and were used in his paper with Joshua Angrist Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size, American Economic Review, June 1998, Vol. 88, No. 3, 450-477. The file Fertility.dta (in STATA format) contains data on 254,654 women between the age of 21 and 35. The data in Fertility are a subset of the data used in the Angrist-Evans paper. (The file Fertility_Small contains data on a 30,000 randomly selected women from the Fertility data set. This smaller dataset is provided for students with memory limitations on their computer software.)

Ks. artikkeli

http://siteresources.worldbank.org/INTPUBSERV/Resources/Angrist_and_Evans.pdf
<http://ideas.repec.org/p/nbr/nberwo/5778.html>

Datan lukeminen R:ään:

```
file<-"http://cc.oulu.fi/~jklemela/econometrics/Fertility_small.csv"  
data<-read.table(file,skip=1,sep=",")
```

Datan lukeminen SAS:iin:

```
FILENAME  
myurl URL 'http://cc.oulu.fi/~jklemela/econometrics/Fertility_small.txt';  
DATA Fertility;  
  INFILE myurl firstobs=2;  
  INPUT ind $ morekids boy1st boy2nd samesex agemom black hispan othrace  
         weeksworked;  
RUN;
```

Muuttujat:

1. V1 = morekids = 1 if mom had more than 2 children
2. V2 = boy1st = 1 if 1st child was a boy
3. V3 = boy2nd = 1 if 2nd child was a boy
4. V4 = samesex = 1 if 1st two children same sex
5. V5 = agemom= age of mom at census
6. V6 = black =1 if mom is black

7. $V7 = \text{hispan} = 1$ if mom is Hispanic
8. $V8 = \text{othrace} = 1$ if mom is not black, Hispanic or white
9. $V9 = \text{weeksworked} =$ mom's weeks worked in 1979
5. (a) Valitse weeksworked vastemuuttujaksi ja morekids selittäjäksi ja suorita pienimmän neliösumman regressio.
- (b) Suorita edellinen regressio valiten samesex instrumenttimuuttujaksi.
- (c) Lisää edelliseen regressioon muuttujat agemom, black, hispan ja othrace eksogeenisinä selittävinä muuttujina.

3 Tärkeitä tehtäviä (eivät kuulu laskuharjoitukseen)

1. Olkoon

$$Y_t = \beta_1^{(1)} + \beta_2^{(1)} X_{t2} + \dots + \beta_K^{(1)} X_{tK} + \epsilon_t, \quad t = 1, \dots, T_0$$

ja

$$Y_t = \beta_1^{(2)} + \beta_2^{(2)} X_{t2} + \dots + \beta_K^{(2)} X_{tK} + \epsilon_t, \quad t = T_0 + 1, \dots, T.$$

Selitä miten nollahypoteesia

$$H_0 : \beta_1^{(1)} = \beta_1^{(2)}, \dots, \beta_K^{(1)} = \beta_K^{(2)}$$

voidaan testata F -testillä (Chown testi).

2. Olkoon

$$\mathcal{Y} = \mathcal{X}\beta + \epsilon,$$

missä $\mathcal{Y} = (Y_1, \dots, Y_n)'$ ja $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ ovat $n \times 1$ -satunnaisvektoreita, $\mathcal{X} = (X_1, \dots, X_n)'$ on $n \times K$ -matriisi ja β on $K \times 1$ -vektori. Olkoon $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d., $E(\epsilon | \mathcal{X}) = 0$ ja $E(\epsilon\epsilon' | \mathcal{X}) = \sigma^2 I$.

Osoita, että

$$\sqrt{n}(b - \beta) \xrightarrow{d} N(0, \sigma^2 Q^{-1}),$$

kun $n \rightarrow \infty$, missä

$$b = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\mathcal{Y}$$

ja $Q = EX_i X_i'$.

3. Olkoon

$$\mathcal{Y} = \mathcal{X}\beta + \epsilon,$$

missä $\mathcal{Y} = (Y_1, \dots, Y_n)'$ ja $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ ovat $n \times 1$ -satunnaisvektoreita, $\mathcal{X} = (X_1, \dots, X_n)'$ on $n \times K$ -matriisi ja β on $K \times 1$ -vektori. Olkoon $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d., $E(\epsilon | \mathcal{X}) = 0$ ja $E(\epsilon\epsilon' | \mathcal{X}) = \sigma^2 I$.

Oletetaan tunnetuksi, että

$$\sqrt{n}(b - \beta) \xrightarrow{d} N(0, \sigma^2 Q^{-1}),$$

kun $n \rightarrow \infty$, missä

$$b = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\mathcal{Y}$$

ja $Q = EX_i X_i'$. Olkoon

$$R\beta = q,$$

missä R on $J \times K$ -matriisi ja q on J -vektori, missä $1 \leq J \leq K - 1$.

Osoita, että

$$JF \xrightarrow{d} \chi^2(J),$$

missä

$$F = \frac{(Rb - q)'(s^2 R(\mathcal{X}'\mathcal{X})^{-1} R')^{-1}(Rb - q)}{J}$$

ja

$$s^2 = \frac{1}{n - K} \sum_{i=1}^n (Y_i - X_i' b)^2.$$