

Ekonomietrian tilastolliset menetelmät

Laskuharjoitus 5

2014, Helmikuu 13

1 Laskutehtäviä

1. Olkoon $Z \sim N(0, 1)$. Osoita, että

$$P(|Z| > c) = 2(1 - \Phi(c)),$$

missä $c > 0$ ja Φ on standardin normaalijakauman kertymäfunktio.

Totea, että vastaava pätee kaikille symmetrisille jakaumille.

2. Olkoon

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

missä Y ja u ovat reaaliarvoisia satunnaismuuttujia, X on $1 \times K$ -satunnaisvektori ja β on $K \times 1$ -vektori. Oletetaan, että $E(X'u) = 0$ ja $E(X'X)$ on kääntyvä matriisi. Olkoot (X_i, Y_i) , $i = 1, \dots, N$, i.i.d. havainnot (X, Y) :n jakaumasta.

Olkoon

$$\hat{\sigma}^2 = \frac{1}{N - K} \sum_{i=1}^N \hat{u}_i^2,$$

missä $\hat{u}_i = Y_i - X_i\hat{\beta}$ ja $\hat{\beta}$ on OLS-estimaattori. Osoita, että

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2$$

kun $N \rightarrow \infty$, missä $\sigma^2 = Eu^2$.

Ohje: Kirjoita $\hat{u}_i^2 = u_i^2 + 2u_iX_i(\beta - \hat{\beta}) + [X_i(\hat{\beta} - \beta)]^2$.

3. Olkoon

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

missä Y ja u ovat reaaliarvoisia satunnaismuuttujia, X on $1 \times K$ -satunnaisvektori ja β on $K \times 1$ -vektori. Oletetaan, että $E(X'u) = 0$ ja $E(X'X)$ on kääntyvä matriisi. Olkoot (X_i, Y_i) , $i = 1, \dots, N$, i.i.d. havainnot (X, Y) :n jakaumasta.

(a) Osoita, että $\text{Var}(Y) = \text{Var}(X\beta) + \text{Var}(u)$.

(b) Merkitään

$$\rho^2 = 1 - \frac{\text{Var}(u)}{\text{Var}(Y)} = \frac{\text{Var}(X\beta)}{\text{Var}(Y)}.$$

Olkoon

$$R^2 = 1 - \frac{SSR}{SST},$$

missä $SSR = \sum_{i=1}^N (Y_i - X_i\hat{\beta})^2$, $\hat{\beta}$ on OLS-estimaattori ja $SST = \sum_{i=1}^N (Y_i - \bar{Y})^2$, missä $\bar{Y} = N^{-1} \sum_{i=1}^N Y_i$. Osoita, että

$$R^2 \xrightarrow{p} \rho^2,$$

kun $N \rightarrow \infty$. Ohje: Käytä tehtävän 2 tulosta.

(c) Pitääkö seuraava väite paikkansa: "Heteroskedastisessa tapauksessa R^2 ei ole järkevä selitysasteen mitta"?

4. Tarkastellaan regressiota

$$arvos = \beta_0 + \beta_1 yliop + \beta_2 PC + u,$$

missä $arvos$ on yliopisto-opiskelijan suoritusten keskimääräinen arvosana, $yliop$ on ylioppilastodistuksen arvosana ja $PC = 1$ mikäli opiskelija omistaa tietokoneen ja 0 muuten.

(a) Miksi PC ja u saattavat olla korreloituneita?

(b) Mikäli käytetään OLS-estimaattoria, niin onko $\hat{\beta}_2$:n harha positiivinen vai negatiivinen?

(c) Mitä muuttujia voitaisiin käyttää proxy-muuttujina?

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 Thier 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 = hispan =1 if mom is Hispanic
8. V8 = othrace =1 if mom is not black, Hispanic or white
9. V9 = 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 Kertaustehtäviä (eivät kuulu laskuharjoitukseen)

1. Olkoon

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

missä Y ja u ovat reaaliarvoisia satunnaismuuttujia, X on $1 \times K$ -satunnaisvektori ja β on $K \times 1$ -vektori. Oletetaan, että $E(X'u) = 0$ ja $E(X'X)$ on kääntyvä matriisi. Olkoot (X_i, Y_i) , $i = 1, \dots, N$, i.i.d. havainnot (X, Y) :n jakaumasta. Pienimmän neliösumman estimaattori on

$$\hat{\beta} = \left(\sum_{i=1}^N X_i'X_i \right)^{-1} \sum_{i=1}^N X_i'Y_i.$$

- (a) Anna estimaattori matriisille $\sigma^2 A^{-1}$, missä $A = E(X'X)$ ja $\sigma^2 = Eu^2$.
- (b) Anna estimaattori matriisille $A^{-1}BA^{-1}$, missä $B = E(u^2X'X)$.

2. Olkoon

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K + \gamma q + v,$$

missä $E(v | X_1, \dots, X_K, q) = 0$ ja muuttujaa q ei havaita. Olkoon z proxy-muuttuja, joka toteuttaa ehdot

- (a) $E(Y | X_1, \dots, X_K, q, z) = E(Y | X_1, \dots, X_K, q)$,
- (b) $L(q | 1, X_1, \dots, X_K, z) = L(q | 1, z)$, missä L tarkoittaa lineaarista projektiota.

Perustele, että proxy-muuttujan avulla parametrit β_1, \dots, β_K voidaan estimoida tarkentuvasti.

3. Tarkastellaan lineaarista mallia

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

missä $X = (X_1, \dots, X_K)$ on $1 \times K$ -vektori, β on $K \times 1$ -vektori ja $Y, u \in \mathbf{R}$. Olkoon $Z = (X_1, \dots, X_{K-1}, Z_1)$ ja oletetaan, että $EZ'u = 0$ ja $EZ'X$ on kääntyvä matriisi. Osoita, että

$$\beta = (EZ'X)^{-1}EZ'Y.$$