

Pitkittäis- ja paneeliaineistojen analysointi

Laskuharjoitus 2

2014, Maaliskuu 26

1 Laskutehtäviä

1. Olkoon

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

missä X on $1 \times K$ -vektori ja β on $K \times 1$ -vektori. Oletetaan, että (X_i, Y_i) , $i = 1, \dots, N$, on i.i.d. otos ja $E(u_i | X_i) = 0$. Osoita, että

$$E(\hat{\beta} | X_1, \dots, X_N) = \beta,$$

missä

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

2. Olkoon

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

missä X on $1 \times K$ -vektori ja β on $K \times 1$ -vektori. Oletetaan, että (X_i, Y_i) , $i = 1, \dots, N$, on i.i.d. otos ja $E(u_i | X_i) = 0$ sekä $\text{Var}(u_i | X_i) = \sigma^2$. Osoita, että

$$\text{Var}(\hat{\beta} | X_1, \dots, X_N) = \sigma^2 \left(\sum_{i=1}^N X_i' X_i \right)^{-1},$$

missä

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

3. Havaitaan $(X_{t,i}, Y_{t,i})$, $i = 1, 2$, $t = 1, \dots, T$, jotka toteuttavat lineaarisen mallin

$$Y_{t,i} = \beta_0 + \beta_1 X_{t,i} + c_t + u_{t,i},$$

missä $E(X_{t,i}u_{t,i}) = 0$, c_t on satunnaismuuttuja jota ei havaita ja $E(c_t u_{t,i}) \neq 0$. Selvitä kuinka parametri β_1 voidaan estimoida pienimmän neliösumman menetelmällä.

4. Havaitaan $(X_{t,i}, Y_{t,i})$, $i = 1, \dots, N$, $t = 1, 2$, jotka toteuttavat lineaarisen mallin

$$Y_{t,i} = \beta_0 + \beta_1 X_{t,i} + c_i + u_{t,i}.$$

Oletetaan vahva eksogeenisyys:

$$E(u_{t,i} | X_{1,i}, X_{2,i}) = 0.$$

Osoita, että

$$E(\Delta X_i \Delta u_i) = 0,$$

missä $\Delta X_i = X_{2,i} - X_{1,i}$ ja $\Delta u_i = u_{2,i} - u_{1,i}$.

2 Tietokonetehtäviä

Tutkitaan aineistoa cornwell.raw, jossa on muuttujat

| variable name | storage type | display format | value label | variable label |
|---------------|--------------|----------------|-------------|---------------------------------|
| county | int | %9.0g | | county identifier |
| year | byte | %9.0g | | 81 to 87 |
| crm rte | float | %9.0g | | crimes committed per person |
| pr barr | float | %9.0g | | 'probability' of arrest |
| pr bconv | float | %9.0g | | 'probability' of conviction |
| pr bpris | float | %9.0g | | 'probability' of prison sentenc |
| avg sen | float | %9.0g | | avg. sentence, days |
| pol pc | float | %9.0g | | police per capita |
| density | float | %9.0g | | people per sq. mile |
| tax pc | float | %9.0g | | tax revenue per capita |
| west | byte | %9.0g | | =1 if in western N.C. |
| central | byte | %9.0g | | =1 if in central N.C. |
| urban | byte | %9.0g | | =1 if in SMSA |
| pct min80 | float | %9.0g | | perc. minority, 1980 |
| wcon | float | %9.0g | | weekly wage, construction |

| | | | |
|----------|-------|-------|---------------------------------|
| wtuc | float | %9.0g | wkly wge, trns, util, commun |
| wtrd | float | %9.0g | wkly wge, whlesle, retail trade |
| wfir | float | %9.0g | wkly wge, fin, ins, real est |
| wser | float | %9.0g | wkly wge, service industry |
| wmfg | float | %9.0g | wkly wge, manufacturing |
| wfed | float | %9.0g | wkly wge, fed employees |
| wsta | float | %9.0g | wkly wge, state employees |
| wloc | float | %9.0g | wkly wge, local gov emps |
| mix | float | %9.0g | offense mix: face-to-face/other |
| pctymle | float | %9.0g | percent young male |
| d82 | byte | %9.0g | =1 if year == 82 |
| d83 | byte | %9.0g | =1 if year == 83 |
| d84 | byte | %9.0g | =1 if year == 84 |
| d85 | byte | %9.0g | =1 if year == 85 |
| d86 | byte | %9.0g | =1 if year == 86 |
| d87 | byte | %9.0g | =1 if year == 87 |
| lcrmte | float | %9.0g | log(crmte) |
| lprbarr | float | %9.0g | log(prbarr) |
| lprbconv | float | %9.0g | log(prbconv) |
| lprbpris | float | %9.0g | log(prbpris) |
| lavgsen | float | %9.0g | log(avgsen) |
| lpolpc | float | %9.0g | log(polpc) |
| ldensity | float | %9.0g | log(density) |
| ltaxpc | float | %9.0g | log(taxpc) |
| lwcon | float | %9.0g | log(wcon) |
| lwtuc | float | %9.0g | log(wtuc) |
| lwtrd | float | %9.0g | log(wtrd) |
| lwfir | float | %9.0g | log(wfir) |
| lwser | float | %9.0g | log(wser) |
| lwmfg | float | %9.0g | log(wmfg) |
| lwfed | float | %9.0g | log(wfed) |
| lwsta | float | %9.0g | log(wsta) |
| lwloc | float | %9.0g | log(wloc) |
| lmix | float | %9.0g | log(mix) |
| lpctymle | float | %9.0g | log(pctymle) |
| lpctmin | float | %9.0g | log(pctmin) |
| clcrmte | float | %9.0g | lcrmte - lcrmte[_n-1] |
| clprbarr | float | %9.0g | lprbarr - lprbarr[_n-1] |
| clprbcon | float | %9.0g | lprbconv - lprbconv[_n-1] |
| clprbpri | float | %9.0g | lprbpri - lprbpri[t-1] |
| clavgsen | float | %9.0g | lavgsen - lavgsen[t-1] |

```

clpolpc          float   %9.0g          lpolpc - lpolpc[t-1]
cltaxpc          float   %9.0g          ltaxpc - ltaxpc[t-1]
clmix            float   %9.0g          lmix - lmix[t-1]

```

Lue data R:ään komennoilla

```

file<-"http://cc.oulu.fi/~jklemela/panel/cornwell.raw"
data<-read.table(file=file)

```

Lue data SAS:iin komennoilla (Huom. lisää tarvittavat muuttujat INPUT-riville).

```

FILENAME myurl URL 'http://cc.oulu.fi/~jklemela/panel/cornwell.raw';

```

```

DATA cornwell;
  INFILE myurl firstobs=1;
  INPUT county year crmrte prbarr prbconv prbpris avgsen;
RUN;

```

5 Estimoi regressiomalli vuoden 1987 aineistolle, jossa y -muuttuja on $\log(\text{crmrte})$ (logarithm of crimes committed per person) ja selittävät muuttujat ovat $\log(\text{prbarr})$, $\log(\text{prbconv})$, $\log(\text{prbpris})$, $\log(\text{avgsen})$.

Lisää selittäväksi muuttujaksi edellisten lisäksi $\log(\text{crmrte})$ vuodelle 1986 ja vertaa saatuja estimaatteja edelliseen.

3 Kertauskysymyksiä (eivät kuulu laskuharjoitukseen)

1. Määrittele seuraavat aineistotyypit ja anna esimerkkejä kustakin aineistotyypistä:

- (a) aikasarja-aineisto (time series data),
- (b) rinnakkaisaineisto (cross sectional data),
- (c) paneeli-aineisto (panel data).

2. Havaitaan $(X_{t,i}, Y_{t,i})$, $i = 1, \dots, N$, $t = 1, 2$, jotka toteuttavat lineaarisen mallin

$$Y_{t,i} = \beta_0 + \beta_1 X_{t,i} + c_i + u_{t,i},$$

missä $E(X_{t,i}u_{t,i}) = 0$, c_i on satunnaismuuttuja jota ei havaita ja $E(c_i u_{t,i}) \neq 0$. Selvitä kuinka parametri β_1 voidaan estimoida pienimmän neliösumman menetelmällä.