Chapter 2

Økonometrisk model for kriminalitet:

$$Crime = \beta_0 + \beta_1 othinc + \beta_2 freqarr + u$$

u består af uobseverede faktorer såsom, lønnen for kriminelalitet, morale, osv.

Konstanterne β er paramternene for den økonometriske model.og beskriver retningen og styrken af forholdet mellem kriminalitet og faktorerne der bestemmer kriminalitet i modellen.

Regression analyse:

$$p_i = \beta_0 + \beta_1 m2_i + \beta_3 location + \beta_4 agentowned + \epsilon_i$$

p: observeret data
m2: husets størrelse
location: afstand til centrum
agentowned: ejet af ejendomsmægler
ϵ : fejlled
β_4 er den forventede prisforskel mellem huse der ejes af mægler og andre hus.

Datatyper

Tværsnitsdata (cross sections):

Bruges når vi har med engelkurver at gøre: y_i i=1,...,n. i kan være personer.

Regressionsmodellen for tværsnitsdata er:

$$s_i = \beta_0 + \beta_1 log(income) + \epsilon_i$$

Tidsrække data:

Det er mere makro data eller finansielle data som lønudviklingen i dk.y_t t=1,2,...,T. T er tiden

Paneldata:

Kombination af tværsnit og tidsrække; f.eks. ser vi på landes arbejdsløshed over tid: y_{it}, i=1,...,n, t=1,...,T

Definition af simpel regressions model

Simpel lineær regrression model:

$$y = \beta_0 + \beta_1 x + u$$

Hvor y er en afhængig variabel og x er uafhængig.
u står for uobserverede.
Hvis u holdes konstant har x en lineær effekt på y.

$$E(u|x) = E(u)$$

Dette fortæller ps at den gennemsnitlige værdi for u er den samme for alle stykker af populationen somer er besluttet af x. Når ovenstående holder, siger vi at u er "strengt uafhængig" på x.

Population regression function (PRF)

$$E(y|x) = \beta_0 + \beta x$$

Fortæller hvordan den gennemsnitlige værdi af y skifter med x.

Ordinary lest squares (OLS)

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Dette er ols estimater for β_0, β_1

OLS regression line eller Sample regression function (SRF)

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

hat, betyder at det er et estimat

Total sum of squares (SST)

$$SST \equiv \sum_{i=1}^{n}(y_i - \bar{y})^2$$

a measure of the total sample variation in the yi ; that is, it measures how spread out the yi are in the sample.

(SSE)

$$SSE \equiv \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

SSE measures the sample variation in the y^i (where we use the fact that $\bar{y} = \bar{y}$)

SSR)

$$SSR \equiv \sum_{i=1}^{n} \hat{u}_i^2$$

SSR measures the sample variation in the u^i

$$SST = SSE + SSR$$

Godness of fit:

$$R^2 = \frac{SSE}{SST}$$

Jo højere R^2 desto mindre varitation er der mellem punkterne i et skatter plot og desto mere bliver der forklaret af variationen.

Coefficient of determination (R-squared)

$$R^2 = SSE/SST = 1 - SSR/SST$$

Assumption SLR.5

$$Var(u|x) = \sigma^2$$

σ^2 is often called the error variance. Kvadratroden af dette er standard deviation

Sampling variances of the ols estimators

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \sigma^2 / SST_x$$

og

$$Var(\hat{\beta}_0) = \frac{\sigma^2 n^{-1} \sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Dette er standard formlerne for simpel regressions analyse. Detter er relevant for konfedensintervaller og hypotese testning

Standard error of $\hat{\beta}_1$

$$se(\hat{\beta}_1) = \hat{\sigma}/\sqrt{SST_x} = \hat{\sigma}/(\sum_{i=1}^{n}(x_i - \bar{x})^2)^{1/2}$$

Regression through the origin

$$\tilde{y} = \tilde{\beta}_1 x$$

Passes through x=0 and $\tilde{y} = 0$

Antagelser for den simple lineære regressionsmodel

SLR.1: Lineær i paremtrene

$$y = \beta_0 + \beta_1 x + u$$

SLR.2: Tilfældig udvælgelse. Oberservationerne skal være uafhængige.
SLR.3: Variation i x. x må i datasættet i antage den samme værdi for alle observationer.
SLR.4: Den betingede middelværdi af fejlledet skal være 0: $E(u|x) = 0$

Homodeskasticitet

Det betyder at variansen er konstant (ikke afhænger af x).
Det gælder at:

$$Var(u|x) = \sigma^2$$

Og at:

$$\sigma^2 = Var(u)$$

Variansen af OLS estimatoren

$$Var(\hat{\beta}_1|x) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

hvor variansen kan estimeres som:

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{SSR}{n-2}$$

Kapitel 3

$$SSR/SST + SSE/SST = 1$$

Multiple lineære regressionsmodel på matrix form:

$$Y = (\begin{matrix} y_1 \\ y_2 \\ . \\ y_n \end{matrix}), X = (\begin{matrix} 1 & x_{11} & ... & x_{1k} \\ 1 & x_{21} & ... & x_{2k} \\ . & & . & . \\ 1 & x_{n1} & ... & x_{nk} \end{matrix}), u = (\begin{matrix} u_1 \\ u_2 \\ . \\ u_n \end{matrix}), \beta = (\begin{matrix} \beta_0 \\ \beta_1 \\ . \\ \beta_k \end{matrix})$$

Omitted variable bias:

$$E(\tilde{\beta_1}) = \beta_1 + \beta_3 \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)x_{i3}}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}$$

Sampling Variances of the OLS Slope Estimator

$$var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)'}$$

, a larger s2 means larger sampling variances for the OLS estimators.

we see that the larger the total variation in xj is, the smaller is $Var(\hat{\beta}_j)$

When SSTj is small, Var1b^ j 2 can get very large, but a small SSTj is not a violation of Assumption MLR.3. Technically, as SSTj goes to zero, Var1b^ j 2 approaches infinity

R^2 close to one indicates that x2 explains much of the variation in x1 in the sample. This means that x1 and x2 are highly correlated. As R_i^2increases to one, $Var(\hat{\beta}_1)$gets larger and larger.

Variance inflation factor

$$VIF_j = 1/(1 - R_j^2)$$

MLR.1

Lineær i paremtrene og modellen kan skrives på matrixform.

MLR.2

Tilfældig udvælgelse:

Vi har tilfældigt udvalgte ovservationer fra en population. Det betyder, at observationerne er uafhængige.

MLR.3

Ingen perfekt multikolinaritet mellem de forklarende variable: I datasættet må ingen af x'erne antage den samme værdi for alle ovservationer, og der må ikke være en lineær relation mellem de forklarende variable. Dette vil medøfre, at X har fuld rang.

MLR.4

Den betingede middelværdi af fejlledet skal være lig 0: $E(u|x, ...x_k) = 0$. På matrixform skriver vi det som $E(u|X) = 0$.

Hvis MLR.4 er opfyldt er de forklarende variable eksogene. Hvis $x_j$er korreleret med u kaldes x_j for endogen.

Vi husker at $Bias = \beta_i \cdot \frac{cov(x_i,u)}{var(x_i)}$

MLR

OLS er middelret:

Teorem 3.1 OLS er middelret: Hvis antagelserne MLR.1-MLR.4 er opfyldt, så er OLS estimatoren middelret:

$$E(\hat{\beta}_j) = \beta_j$$

for alle værdier af β_j.

At den er middelret betyder at,

OLS estimatoren vil i "gennemsnit" give den rigtige værdi af parametrene.

Hvis vi kunne gentage estimation på forskellige tilfældige datasæt, vil vi i gennemsnit få den rigtige værdi.

For den enkelte realisation kan der være afvigelser.

OLS estimator

$$\hat{\beta} = (X'X)^{-1}X'y$$

Bevis'

$$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + u)$$

$$= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u$$

$$= I\beta + (X'X)^{-1}X'u$$

$$= \beta + (X'X)^{-1}X'u$$

Trin 2

$$E(\hat{\beta}|X) = \beta + E(((X'X)^{-1}X'u|X)$$

$$= \beta + (X'X)^{-1}X'E(u|X)$$

MLR.4

$$= \beta + (X'X)^{-1}X' * 0$$

$$= \beta$$

For den ubetingede middelværdi fældet det at:

$$E(\hat{\beta}_j) = \beta_j$$

Middelværdien af estimatoren

$$E(\tilde{\beta}|x) = \beta_1 + \beta_2 \frac{cov(x_2, x_1)}{Var(x_1)}$$

Summary of Bias in $\tilde{\beta_1}$ when x_2 is Omitted i estimation equation

$\beta_2 > 0$: positve bias if $Corr(x_1, x_2) > 0$ and negative $Corr(x_1, x_2) < 0$

$\beta_2 < 0$: positve bias if $Corr(x_1, x_2) < 0$ and negative $Corr(x_1, x_2) > 0$

Størrelsen på bias er beslutted ud fra stærrelsen af $\beta_2$og $\tilde{\delta}_1$

MLR.5

Variansen af fejlleddet er konstant

$$Var(u|x_1, ..., x_k) = \sigma^2$$

Variansen af OLS estimatoren

$$Var(\hat{\beta}|X) = \sigma^2(X'X)^{-1}$$

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$$

$$\hat{\beta} = \beta + (X'X)^{-1}X'u$$

bevis for varienansen af OLS estimatoren

$$Var(\hat{\beta}|X) = Var((X'X)^{-1}X'u|X)$$

$$= AVar(u|X)A'$$

MLR.2
MLR-5

$$(X'X)^{-1}X'\sigma^2 IX(X'X)^{-1}$$

$$\sigma^2(X'X)^{-1}X'IX(X'X)^{-1}$$

$$\sigma^2(X'X)^{-1}X'X(X'X)^{-1}$$

$$X'X(X'X)^{-1} = I$$

$$\sigma^2(X'X)^{-1}$$

$$B = X'X$$

$$B' = X'X''$$

Estimation af σ^2

$$\hat{\sigma}^2 = \frac{SSR}{n-k-1}$$

antallet af oberservationer minus antallet af obeservereve variationer (k) minus 1 er nævneren.

Ovenstående ligning er gældende hvis antagelserne MLR.1-MLR.5 holder.

$$E(\hat{\sigma}^2|X) = \sigma^2$$

OLS estimatoren

$$\hat{\beta}_1 = \sum_{i=1}^{n} w_i y_i$$

OLS estimatoren kan skrives som en vægtet som af y'erne hvor $w_i = \frac{x_i - \bar{x}}{SST_1}$

For vægtene gælder det

$\sum_{i=1}^{n} w_i = 0$

$\sum_{i=1}^{n} w_i x_i = 1$

Bevis for β_1 er en middelret estimator: $E(\tilde{\beta}_1) = \beta_1$

Ingredienser:

- MLR 1: $y = \beta_0 + \beta_1 x + u$
- MLR 3 (særlig variason)

-MLR 4: $E(u|X) = 0$

- Den alternative $\tilde{\beta} = \frac{y_1 - y_n}{x_1 - x_n}$

Bevis

- Trin 1: Udtryk $\tilde{\beta}$ som en sum af β_1 og et led som afhænger af u og x. Anvende MLR.1.
- Trin 2: Udregn $E(\tilde{\beta}_1|x)$ og anvend MLR.4. og regnereglen for betingede middelværdier

$$\tilde{\beta}_1 = \frac{y_1 - y_n}{x_1 - x_n}$$

$$\frac{\beta_0 + \beta_1 x_1 + u_1 - \beta_0 + \beta_1 x_n + u_n}{x_1 - x_n}$$

$$= \beta_1 \frac{u_1 - u_n}{x_1 - x_n}$$

Trin 2

$$E(\tilde{\beta}|x) = \beta_1 + E[\frac{u_1 - u_n}{x_1 - x_n}|x] = \beta_1$$

Gauss-Markov Teoremet: Bevis

1. Rengergler

Gauss-Markov Teoremet:

Ingredienser til bevis

1. Regneregler (A og B matricer, x stok. vektor, c skalar)
    1. $cAB = AcB = ABc$
    2. $(AB)' = B'A'$ og $(A')' = A$
    3. $AA^{-1} = I$ og $A^{-1}A = I$
    4. $Var(Ax) = AVar(x)A'$
    5. $Var(y)" > "Var(x)$ hvis $Var(y) - Var(x)$ er positiv semi definit (p.s.d.)
    6. Def: En kvadratisk symmetrisk matrix A er positiv semi definit hvis $x'Ax \geq 0$
    7. Hvis B har fuld rang så er $B'B$ og BB' positiv semi definit (p.s.d.)
2. Antagelser
    1. MLR.1: Model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$
    2. MLR.3: $\mathbf{X}$ fuld rang
    3. MLR.4: $E(\mathbf{u}|\mathbf{X}) = \mathbf{0}$
    4. MLR.2 og MLR.5: $Var(\mathbf{u}|\mathbf{X}) = \sigma^2\mathbf{I}$
3. Variansen af OLS Estimator: $Var(\hat{\boldsymbol{\beta}}|\mathbf{X}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

Bevis
Trin1(opstil alternativ lineær middelret estimator $\tilde{\beta}$
Lineær estimator:

$$\tilde{\beta} = Ay = A(X + \beta + u)\text{=AX}\beta\text{+Au}$$

2. Lighedstegn følger af MLR.1

$$E(\tilde{\beta}|X) = E(AX\beta + Au|X) = AX\beta + AE(u|X) = AX\beta + A0 = AX\beta$$

4. lighedstegn følger af MLR.4
Hvis $\tilde{\beta}$ skal være middelret, må det gælde at $AX = I$(identitetsmatricen)
Trin 2 (udregn variansen af $\tilde{\beta}$)

$$Var(\beta + Au|X)$$

$$= AVar(u|X)A' = A\sigma^2 IA' = \sigma^2 AA'$$

Vi benytter regneregel (4) og (1) og MLR.2 og MLR.5
Trin 3 (sammenligning af variansen $\tilde{\beta}$ og OLS estimatoren):

$$Var(\tilde{\beta}|X) - Var(\tilde{\beta}|X) = \sigma^2 AA' - \sigma^2(X'X)^{-1}$$

$$= \sigma^2(AA' - (X'X)^{-1})$$

$$\sigma^2(AA' - (AX)(X'X)^{-1}(X'A'))$$

$$= \sigma^2 A(I - X(X'X)^{-1}X')A'$$

$$\sigma^2 AMA'$$

hvor $M = (I - X(X'X)^{-1}X')$

Vi skal vise at AMA' er en postiv semi-definit matrix jvf regneregel (5). TIl det skal vi vise at M=M'

$$M' = (I - X(X'X)^{-1}X')' = I'X''[(X'X)^{-1}]'X'$$

$$= (I - X(X'X)^{-1}X') = M$$

Vi benytter regneregel (2).
Vi skal vise at MM=M:

$$MM = (I - X(X'X)^{-1}X')(I - X(X'X)^{-1}X')$$

$$= I - 2X(X'X)^{-1}X' + X(X'X)^{-1}X' = M$$

$$Var(\tilde{\beta}|X) - Var(\tilde{\beta}|X) = \sigma^2 AMA' = \sigma^2 AMMA' = \sigma^2(AM)M'A'$$

$$= \sigma^2(AM)(AM)'$$

Sidste lighedstegn fælger af (2). Ved at benytte regneregel (7) kan vi slutte at $Var(\tilde{\beta}|X) - Var(\tilde{\beta}|X)$ er postiv semi-definit

Kapitel 4

MLR.6

Normalitet funktions: $U \sim Normal(0, \sigma^2)$
$E(u) = 0$ og $Var(u) = \sigma^2$

t statistic

$$y_{\hat{\beta}_j} \equiv \hat{\beta}_j / se(\hat{\beta}_j)$$

Testont other hypotheses about beta j

$$t = (\hat{\beta}_j - a_j) / se(\hat{\beta}_j)$$

Konfidensinterval

$$\hat{\beta}_j + -c * se(\hat{\beta}_j)$$

F statistic

$$F \equiv \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$$

R-squared form of the F statistic

$$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/df_{ur}}$$

Opsummering fra kap 3 og 4

Hvis MLR.1-4 gælder såer er OLS middelret
Hvis MLR.1-5 gælder så er OLS BLUE
HVIS MLR.1-6 gælder så er OLS normalfordelt, og vi kan lave inferens med t- og F-test.

Kapitel 5

Eksempler på multiple hypoteser

$$H_0 : \beta_3 = 1 (\text{en restriktion})$$

Antallet af lighedstegn er lig antallet af restriktioner.

Teststørrelsen for multiple hypoteser

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$$

Tælleren er altid positiv.
q er antal restriktioner
n-k-1 er antal frihedsgrader i den urestikterede model.
Hvis MLR.1-MLR.6 er opfyldt så gælder følgende:

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim F(1, n-k-1)$$

Definition af konsistens

$$P(|W_n - \theta| > \epsilon) \to 0, n \to \infty$$

Hvis dette ikke gælder sig vi at Wn er inkonsistent.

Vi kræver normalt for en estimator at den skalv være konsistent for at give mening.

Store tals lov

Se appendix C-3

$$p \lim \bar{Y}_n = \mu$$

Regnerelger for konvergens i sandsynlighed:

For en kontinuert funktion g gælder:

$$p \lim g(W_n) = g(p \lim(W_n))$$

Hvis $p \lim(G_n) = \alpha$ og $p \lim(W_n) = \beta$så
$p \lim(G_n + W_n) = \alpha + \beta$
$p \lim(G_n W_n) = \alpha\beta$
$p \lim(G_n / W_n) = \alpha / \beta$

Bevis for at OLS estimatoren er konsistent

Under antalge MLR.1-4 er OLS estimatoten $\hat{\beta}_j$en konsistent estimatore for β_j

Ingredienser:

OLS estimatoren er konsistent:

Ingredienser til bevis.

1. Regneregler:
   1. Hvis $p\lim(G_n) = \alpha$ og $p\lim(W_n) = \beta$, så gælder $p\lim(G_n/W_n) = \alpha/\beta$.
   2. Store tals lov: $p\lim(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2) = Var(Y)$ og $p\lim(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})) = cov(Y, X)$.
   3. $\sum_{i=1}^{n} x_i(x_i - \bar{x}) = \sum_{i=1}^{n}(x_i - \bar{x})^2 = SST_x$.
   4. $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.
2. Antagelser:
   1. MLR.1: Model $y = \beta_0 + \beta_1 x_1 + u$.
   2. MLR.2: Tilfældigt datasæt med n uafhængige observationer (dvs. $u'erne$ er uafhængige).
   3. MLR.3: Variation i x.
   4. MLR.4: $E(u|x) = 0. \Longrightarrow cov(u, x) = 0$.
3. OLS Estimator: $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Trin i beviset

OLS estimatoren er konsistent:

Trin i beviset:

1. Vis at $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = ... = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n} u_i(x_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
   (benyt MLR.3 og MLR.1 og regneregel 1.3 og 1.4)
2. Regn på $p\lim\hat{\beta}_1 = ... = \beta_1 + \frac{Cov(u,x)}{Var(x)}$
   (Benyt MLR.2 og regneregel 1.1 og 1.2)
3. Benyt MLR.4 til at vise at $p\lim\hat{\beta}_1 = \beta_1$

Vi starter med at vise at

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n} u_i(x_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Benytter MLR.1

$$= \frac{\sum_{i=1}^{n} y_i(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i + u_i)(x_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \frac{\beta_0 \sum_{i=1}^{n}(x_i - \bar{x}) + \beta_1 \sum_{i=1}^{n} x_1(x_i - \bar{x}) + \sum_{i=1}^{n} u_1(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Benytter RR3 og 4

$$\beta_0 0 + \beta_1 \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} + \frac{\sum_{i=1}^{n} u_1(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n} u_i(x_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Trin 2 (udregn i $\hat{\beta}_1$)

$$p\lim \hat{\beta}_1 = p\lim(\beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n} u_i(x_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2})$$

Benytter RR1

$$\beta_1 + \frac{p\lim(\frac{1}{n}\sum_{i=1}^{n} u_i(x_i - \bar{x}))}{p\lim(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)}$$

bENYTTER rr2 OG mlr.2

$$\beta_1 + \frac{Cov(u, x)}{Var(x)}$$

OLS er inkonsistent

Detter gælder hvis $E(u) \neq 0$ eller $Cov(u, x) \neq 0$

Inkonsistensen eller asymptotisk bias er givet ved

$$p\lim(\hat{\beta}_1) - \beta_1 = \frac{cov(u, x)}{Var(x)}$$

Trin 3 (viser at $p\lim \hat{\beta}_1 = \beta_1$
bruger MLR.4

$$p\lim(\hat{\beta}_1) = \beta_1 + \frac{cov(u, x)}{Var(x)}$$

$$p\lim(\hat{\beta}_1) = \beta_1 + \frac{0}{Var(x)} = \beta_1$$

OLS er asymptotisk normalfordelt

Når MLR.1-5 er opfyldt gælder

$$\sqrt{n}(\hat{\beta}_j - \beta_j) \sim^a N(0, \sigma^2/a^2)$$

Den asymptotiske varians for OLS

$$Var(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

LM -test (Lagrange Multiplier)

Test for multiple lineære restriktioner
F-test
Langrange Multiplier test.
Proceduerен er:
Estimer den restikterede model
Udregn residualerende fra den restrikterede model
Hjælpemodel med residualerne (oftest anvendes R^2 fra hjælpemodellen).

Kapitel 6

Adjusted R^2

$$R^2 = 1 - (SSR/n)/(SST/n)$$

Population R-bar-squared:

$$\rho^2 = 1 - \sigma_u^2/\sigma_y^2$$

$$\bar{R^2} = 1 - (SSR/n - k - 1)/(SST/n - 1) = 1 - \hat{\sigma}^2/(SST/(n-1))$$

$$\bar{R^2} = 1 - (1 - R^2(n-1)/(n-k-1)$$

Kapitel 7

$$F = \frac{[SSR_p - (SSR_1 - SSR_2)]}{SSR_1 + SSR_2} * \frac{[n - 2(k+1)]}{k+1}$$

$$F_{chow} = \frac{\frac{[SSR_p - (SSR_1 + SSR_2 + ... + SSR_m)]}{(m-1)(k+1)}}{\frac{SSR_1 + SSR_2 + ... + SSR_m}{n - (m(k+1))}}$$

Kapitel 8

$$Var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\sigma_i^2}{SST_x^2}$$

Heterodesktacity-robust t statistic

$$t = \frac{estimate - hypothesizedvalue}{standarderror}$$

F statistik

$$F = \frac{R_{\hat{u}}^2/k}{(1 - R_{\hat{u}^2}^2)/(n-k-1)}$$

LM statistik

$$LM = n * R_{\hat{u}^2}^2$$

The Heteroskedasticity is known up to a multiplicative constant

$$Var(u|x) = \sigma^2 h(x)$$

h(x) er en funktion for de forklarende variable der bestemmet heteskedasticity.

The Heteroskedasticity Function Must Be Estimated: FGLS

$$\hat{h}_i = exp(\hat{g}_i)$$

Prediction intervals with Heteroskedasticity

$$E(y|x) = exp(\beta_0 + x\beta + exp(\delta_0 + x\delta)/2)$$

$$\hat{y}_i = exp(\hat{logy_i} + \hat{\sigma}^2\hat{h}_i/2)$$

Linear probability model

$$Var(y|x) = p(x)[1 - p(x)]$$

$$\hat{h}_i = \hat{y}_i(1 - \hat{y}_i)$$

Weighted Least Squares (WLS)

$$\frac{y_i}{\sqrt{h_i}} = \beta_0\frac{1}{\sqrt{h_i}} + \beta_1\frac{x_{j1}}{\sqrt{h_i}} + .. + \frac{u_j}{\sqrt{h_i}}$$

Kapitel 9

$$E(X_3^*|x_1, x_2, x_3) = \delta_0 + \delta_3 x_3$$

Potential outcomes and proxy variables

$$E[v(0)|w, x] = E[v(0)|x] = (x - \eta)\beta_0$$

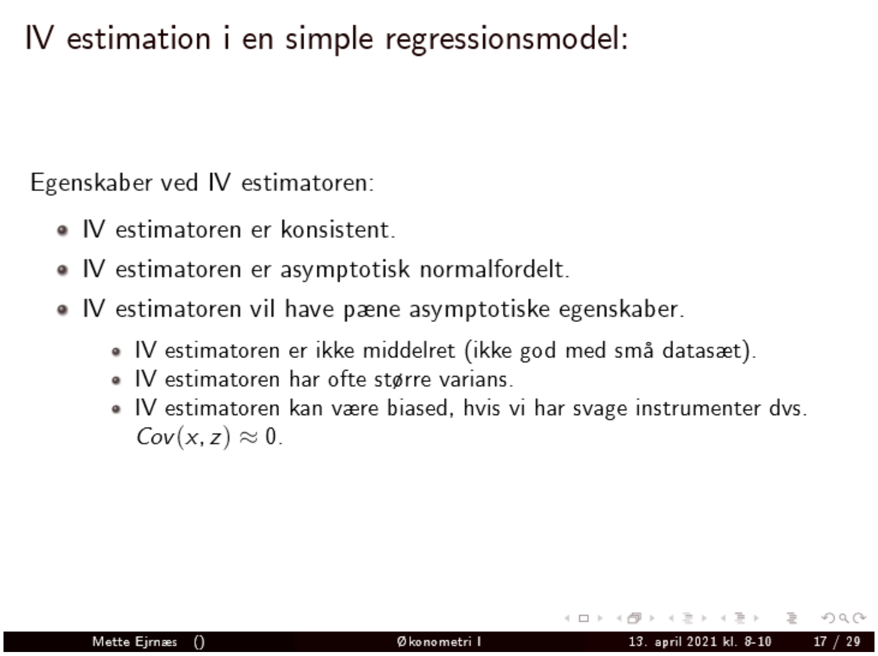
$$E[v(1)|w, x] = E[v(1)|x] = (x - \eta)\beta_1$$

Measurement Error in Explanatory Variable

$$plim(\hat{\beta}_1) = \beta_1 + \frac{Cov(x_1, u - \beta_1 e)}{Var(x_1)} = \beta_1(\frac{\sigma_{e_1}^{2*}}{\sigma_{x_1}^2 + \sigma_{e_1}^2})$$

Kapitel 15

Bevis for IV estimationen er konsistent



IV estimatoren er konsistent, hvis MLR.1-3 er opfyldt og $Cov(u, z) = 0, Cov(x, z) \neq 0 : p\lim \hat{\beta}_1^{IV} = \beta_1$

Ingredienser til bevis:

Bevis for IV estimationen er konsistent:

IV estimatoren er konsistent, hvis MLR.1-MLR.3 er opfyldt, og $Cov(u,z)=0, Cov(x,z)\neq 0 : p\lim \hat{\beta}_1^{IV} = \beta_1$

Ingredienser til bevis

- IV estimatoren: $\hat{\beta}_1^{IV} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})(z_i-\bar{z}) / \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(z_i-\bar{z})$.
- Antagelser:
  - Model: MLR.1: $y=\beta_0+\beta_1 x+u$.
  - MLR.2: uafhængige og identiske fordelte observationer.
  - Gyldigt instrument: $Cov(u,z)=0, Cov(x,z)\neq 0$.
- Regneregler:

1. $\sum_{i=1}^{n}(w_i-\bar{w})(z_i-\bar{z}) = \sum_{i=1}^{n}(w_i)(z_i-\bar{z})$.
2. Hvis $p\lim a_n = a$ og $p\lim b_n = b$ så er $p\lim(\frac{a_n}{b_n}) = \frac{a}{b}$ hvis $b\neq 0$
3. $p\lim\left[\frac{1}{n}\sum_{i=1}^{n}(w_i-\bar{w})(v_i-\bar{v})\right] = Cov(w,v)$, hvis vi har uafhængige observationer.

Mette Ejrnæs () Økonometri I 13. april 2021 kl. 8-10 18 / 29

Trin 1 bruger regneregel 1:

$$\hat{\beta_1}^{IV} = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(y_i-\bar{y})(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\bar{x})(z_i-\bar{z})} = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(y_i)(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}$$

Benytter MLR.1

$$\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(\beta_0+\beta_{\mathbf{1}}x_i+u_i)(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}$$

$$=\beta_0\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}+\beta_1\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}+\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(u_i)(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}$$

$$\hat{\beta_1}^{IV} = \beta_0 * 0 + \beta_1 * 1 + \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(u_i)(z_i-\bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i-\bar{z})}$$

$$= \beta_1 + \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(u_i)(z_i - \bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i - \bar{z})}$$

Trin 2 find grænsen i sandsynlighed

$$p\lim \hat{\beta_1}^{IV} = \beta_1 + p\lim(\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(u_i)(z_i - \bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i - \bar{z})})$$

Bruger regneregel 2

$$= \beta_1 + \frac{p\lim(\frac{1}{n}\sum\limits_{i=1}^{n}(u_i)(z_i - \bar{z}))}{p\lim(\frac{1}{n}\sum\limits_{i=1}^{n}(x_i)(z_i - \bar{z}))}$$

Bruger regneregel 3 + MLR.2

$$= \beta_1 + \frac{cov(u, z)}{cov(x, z)}$$

Antagelser om z

$$= \beta_1 + \frac{0}{cov(x, z)} = \beta_1$$

Instrumental variables (IV) estimator of β_1

$$\hat{\beta_1}^{IV} = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z})}{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}$$

Structural equation

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + u$$

Bevise af IV Estmiator

IV antagelser:

IV estimatoren er defineret og konsistent, hvis følgende antagelser gælder:

- Instrumenterne er ukorreleret med fejlleddet:

$$p\lim\left(\frac{1}{n}\sum_{i=1}^{n} z_{ij}u_i\right) = 0 \Leftrightarrow p\lim \frac{1}{n}Z'u = \mathbf{0}$$

- Instrumenterne er korreleret med de endogene variable:

$$p\lim \frac{1}{n}Z'X = \Sigma_{ZX} \text{ har fuld rang}$$

- Ingen perfekt multikollinaritet mellem instrumenterne:

$$p\lim \frac{1}{n}Z'Z = \Sigma_{ZZ} \text{ har fuld rang}$$

Udledning af IV estimator (eksakt identificeret):

- Eksakt identifikation: l instrumenter.
- Vi tager udgangspunkt i den teoretiske betingelse:

$$\begin{aligned} p\lim \frac{1}{n}Z'u &= \mathbf{0} \\ p\lim(\frac{1}{n}Z'(y - X\beta)) &= \mathbf{0} \\ p\lim(\frac{1}{n}\left(Z'y - Z'X\beta\right)) &= \mathbf{0} \end{aligned}$$

- Hvis vi erstatter de teoretiske momenter med datamomenter, får vi:

$$\begin{aligned} Z'y - Z'X\hat{\beta}^{IV} &= \mathbf{0} \\ Z'y &= Z'X\hat{\beta}^{IV} \\ \hat{\beta}^{IV} &= (Z'X)^{-1}(Z'y) \end{aligned}$$

- IV estimatoren er konsistent.

Bevis for IV er konstisten

IV er konsistent:

Bevis at IV estimatoren er konsistent:
Benyt (nogle af) følgende dele til at vise at IV estimatoren er konsistent.
Sæt delene sammen i rigtig rækkefølge:

- M $E(u|X)$
- O $p\lim$
- S $\hat{\beta}^{IV} = (Z'X)^{-1}(Z'y)$
- G $p\lim \frac{1}{n} Z'u = \mathbf{0}$
- U $Var(u|x) = \sigma^2$
- R $y = X\beta + u$
- T Σ_{ZX} har fuld rang
- N $p\lim(M_n W_n) = MW$

IV er konsistent:

Bevis for at IV estimatoren er konsistent:

Trin 1:

$$\begin{aligned}
\hat{\beta}^{IV} &= (Z'X)^{-1}(Z'y) \\
&= (Z'X)^{-1}(Z'(X\beta + u)) \\
&= (Z'X)^{-1}Z'X\beta + (Z'X)^{-1}(Z'u) \\
&= \beta + (Z'X)^{-1}(Z'u)
\end{aligned}$$

(S-

Trin 2

$$\begin{aligned}
p\lim(\hat{\beta}^{IV}) &= \beta + p\lim((Z'X)^{-1}(Z'u)) \\
&= \beta + p\lim(\frac{1}{n}(Z'X)^{-1})p\lim(\frac{1}{n}(Z'u)) \\
&= \beta + \Sigma_{ZX} \cdot \mathbf{0} \\
&= \beta
\end{aligned}$$

Bevis for at 2SLS er konsistent:

Konsistens:

2SLS er konsistent, når IV antagelserne holder.
Ingredienser til bevis:
Antagelser:

1. $p\lim \frac{1}{n} Z'u = \mathbf{0}$
2. $p\lim \frac{1}{n} Z'X = \Sigma_{ZX}$ har fuld rang.
3. $p\lim \frac{1}{n} Z'Z = \Sigma_{ZZ}$ har fuld rang.

2SLS estimatoren:

$$\hat{\beta}^{2SLS} = (X'P_Z X)^{-1}(X'P_Z y)$$

Regneregel:

1. $p\lim(A_n B_n) = AB$ når $p\lim(A_n) = A$ og $p\lim(B_n) = B$
2. $p\lim(A_n^{-1}) = A^{-1}$ når $p\lim(A_n) = A$

Mette Ejrnæs () Økonometri I 15. april 2021 kl. 10-12 13 / 29

Konsistens-Bevis:

Trin 2 (Grænse i sandsynlighed):
Først ser vi på leddene

$$\begin{aligned}
p\lim(\frac{1}{n}(X'P_Z X)) &= p\lim(\frac{1}{n}\left(X'Z(Z'Z)^{-1}Z'X\right)) \text{ (Regneregel 1 og 2)} \\
&= p\lim(\frac{1}{n}(X'Z)) \left[p\lim(\frac{1}{n}(Z'Z))\right]^{-1} p\lim(\frac{1}{n}Z'X) \\
&\qquad \text{(Antagelse 2 og 3)} \\
&= \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{XZ}'
\end{aligned}$$

$$\begin{aligned}
p\lim(\frac{1}{n}X'P_Z u) &= p\lim(\frac{1}{n}X'Z(Z'Z)^{-1}Z'u) \quad \text{(Regneregel 1 og 2)} \\
&= p\lim(\frac{1}{n}X'Z) \left[p\lim(\frac{1}{n}Z'Z)\right]^{-1} p\lim(\frac{1}{n}Z'u) \\
&\qquad \text{(Antagelse 1, 2 og 3)} \\
&= \Sigma_{XZ}\Sigma_{ZZ}^{-1}\mathbf{0} = \mathbf{0}
\end{aligned}$$

Mette Ejrnæs () Økonometri I 15. april 2021 kl. 10-12 16 / 29 W

Konsistens-Bevis:

Så følger at

$$p\lim(\hat{\beta}^{2SLS}) = \beta + \left(\Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{XZ}'\right)^{-1}\mathbf{0} = \beta$$

Mette Ejrnæs () Økonometri I 15. april 2021 kl. 10-12 17 / 29

W.10-11

Definition for svag stationaritet

En tidsrække er svagt stationær hvis:

$$E(y_t) = \mu$$

$$v(y_t) = E((y_t - \mu)^2) = \gamma_0$$

$$Cov(y_t, y_{t-h}) = E(y_t - \mu)(y_{t-h} - \mu) = \gamma_h$$

Man kan gøre en tidsrække stationær ved:

$$log(prod_t) = \delta + \gamma t + \kappa_t$$

Fjerne trenden ved:

$$\hat{\kappa}_t = log(prod_t) - \hat{\delta} - \hat{\gamma} t$$

Man kan fjerne niveauskift ved at lave en dummy variabel der er 1.

$$pandora_t = \delta + \gamma 1_{\{t<T_0\}} + \kappa_t$$

Kointergration

Er en lineær kombination af ikke-stationære tidsrækker, som er stationære

Den lineære regressionsmodel:

Antagelse 1: MLR.1

Antagel 2: Prædetermineret:
Fejlledet i modellen skal opfylde

$$E(\epsilon_t|x_t) = 0$$

Denne antagelse erstatter antagelse MLR.4

Autoregressive distributed lag model:

Dynamisk model med laggede variable af den afhængige variabel

$$y_t = \theta y_{t-1} + \epsilon_t$$

Dette kaldes en univariat model, fordi den afhængige variabel kun afhænger af sig selv.

Udledning af OLS estimatoren:

$$E(X_t\epsilon_t) = E(X_t(y_t - X_t'\beta) = 0$$

$$E(X_t y_t) - E(X_t X_t')\beta = 0$$

Hvis $E(X_t X_t')$ har fuld rang så gælder at:

$$\beta = [E(X_t X_t')]^{-1}E(X_t y_t)$$

Antaglese 3 ingen perfekt multikollinaritet:
Samme som MLR.3
Estimatoren bliver:

$$\hat{\beta} = [\frac{1}{T}\sum_{t=1}^{T} X_t x_t']^{-1}[\frac{1}{T}\sum X_t y_t]$$

Antagelse 4:
Stationaritet og svag afhængighed.
OLS estimatoren er konsistent under antaglesen 1-4 er opfyldt.

Bevis for OLS estimatoren er konsisten

OLS estimatoren:

- Sætning 1: OLS estimatoren er konsistent:
  Under antagelse af at **antagelserne 1-4 er opfyldt,** er **OLS estimatoren konsistent**:

$$p \lim_{T\to\infty} \hat{\beta} = \beta$$

- Antagelser:
  - Antagelse 1: $y_t = x_t\beta + \varepsilon_t$.
  - Antagelse 2: $E(\varepsilon_t|x_t) = 0 \Rightarrow E(x_t\varepsilon_t) = 0$.
  - Antagelse 3: Ingen perfekt multikollinaritet.
  - Antagelse 4: Stationaritet og svag afhængighed
    $\Rightarrow p\lim\left[\frac{1}{T}\sum_{t=1}^{T} x_t^2\right] = E(x_t^2)$ og $p\lim\left[\frac{1}{T}\sum_{t=1}^{T} x_t\varepsilon_t\right] = E(x_t\varepsilon_t)$.

OLS estimatoren er konsistent:

Bevis::
Trin 1: OLS estimatoren:

$$\begin{aligned}
\hat{\beta} &= \frac{\frac{1}{T}\sum x_t y_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} && \text{(Antagelse 1)}\\
&= \frac{\frac{1}{T}\sum x_t(x_t\beta + \varepsilon_t)}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} = \frac{\frac{1}{T}\sum x_t^2\beta + \frac{1}{T}\sum x_t\varepsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}\\
&= \beta + \frac{\frac{1}{T}\sum x_t\varepsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}
\end{aligned}$$

OLS estimatoren er konsistent

- Trin 2:

$$
\begin{aligned}
p\lim \hat{\beta} &= \beta + p\lim\left(\frac{\frac{1}{T}\sum x_t \varepsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}\right) && \text{(Regneregel 1)} \\
&= \beta + \frac{p\lim(\frac{1}{T}\sum x_t \varepsilon_t)}{p\lim(\frac{1}{T}\sum_{t=1}^{T} x_t^2)} && \text{(Antagelse 4)} \\
&= \beta + \frac{E(x_t \varepsilon_t)}{E(x_t^2)} && \text{(Antagelse 2)} \\
&= \beta + \frac{0}{E(x_t^2)} = \beta
\end{aligned}
$$

OLS er middel under antagelse 5

Antagelse 5: strengt eksogenitet.

$$E(\epsilon|x_1..x_t) = 0$$

Bevis for middelrethed

Middelrethed:

- Sætning 2: **OLS er middelret**.
- Hvis antagelse 1, 3, 4 og 5 er opfyldt, er OLS middelret

$$E(\hat{\beta}) = \beta.$$

- Ingredienser til bevis (simpel lineær regressionsmodel):
- OLS estimatoren:

$$\hat{\beta} = \frac{\frac{1}{T}\sum x_t y_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}$$

- Antagelser:
  - Antagelse 1: $y_t = x_t\beta + \varepsilon_t$.
  - Antagelse 3: Ingen perfekt multikollinaritet.
  - Antagelse 4: Stationaritet og svag afhængighed.
  - Antagelse 5: $E(\varepsilon_t | x_1, x_2, ... x_T) = 0$.

Bevis for middelrethed:

Trin 1: OLS estimatoren:

$$\begin{aligned}
\hat{\beta} &= \frac{\frac{1}{T}\sum x_t y_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} && \text{(Antagelse 1)} \\
&= \frac{\frac{1}{T}\sum x_t(x_t\beta + \varepsilon_t)}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} = \frac{\frac{1}{T}\sum x_t^2\beta + \frac{1}{T}\sum x_t\varepsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} \\
&= \beta + \frac{\frac{1}{T}\sum x_t\varepsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}
\end{aligned}$$

Bevis for middelrethed:

Trin 2:

$$\begin{aligned} E(\hat{\beta}|x_1, x_2, ..x_T) &= \beta + E(\frac{\frac{1}{T}\sum x_t\varepsilon_t}{\frac{1}{T}\sum_{t=1}^T x_t^2}|x_1, x_2, ..x_T) \\ &= \beta + \frac{E(\frac{1}{T}\sum x_t\varepsilon_t|x_1, x_2, .., x_T)}{\frac{1}{T}\sum_{t=1}^T x_t^2} \\ &= \beta + \frac{\frac{1}{T}\sum x_t E(\varepsilon_t|x_1, x_2, .., x_T)}{\frac{1}{T}\sum_{t=1}^T x_t^2} \text{(Antagelse 5)} \\ &= \beta + \frac{\frac{1}{T}\sum x_t 0}{E(x_t^2)} = \beta \end{aligned}$$

Så følger at

$$E(\hat{\beta}) = E(E(\hat{\beta}|x_1, x_2, ..x_T)) = \beta$$

Gentagende tværsnitsdata og paneldata

w.14.1-14.2