

Øvelsestime 4

Monte Carlo Eksperimenter

Agenda

- 1) Introduktion til PS4
- 2) Gruppearbejde: Diskussion
- 3) Opsamling
- 4) Gruppearbejde: Stataøvelser
- 5) Opsamling

Tobias Gabel Christiansen
tgc@econ.ku.dk
15-03-2021

KØBENHAVNS UNIVERSITET



1) Planen for semesteret

Dato	Opgave
Uge 7 - 15/02	PS1
Uge 8 - 22/02	PS2
Uge 9 - 01/03	PS3
Uge 10 - 08/03	OO1
Uge 11 - 15/03	PS4
Uge 12 - 22/03	PS5 + gennemgang af OO1
Uge 13 - 29/03	PS6
Uge 14 - 05/04	Påskeferie
Uge 15 - 12/04	OO2
Uge 15 - 19/04	PS7
Uge 16 - 26/04	PS8 + gennemgang af OO2
Uge 17 - 03/05	OO3
Uge 18 - 10/05	PS9
Uge 19 - 17/05	PS10 + gennemgang af OO3
	Spørgetime

 Her er vi :-)

1) Introduktion til problem set 4

Monte Carlo eksperimenter

- Indtil videre har vi analyseret OLS-estimatorens egenskaber (fx bias og konsistens) gennem teoretiske udledninger under en række antagelser (fx MLR.1-4)
- Nogle gange er analytiske løsninger svære eller umulige at udlede.
- Som alternativ kan vi bruge Monte Carlo (MC) eksperimenter.
- I et MC-eksperiment:
 - i) Specificerer vi den data generende proces (populations modellen)
 - ii) Simulerer mange random samples
 - iii) Estimerer de samme parametre mange gange
 - iv) Undersøger estimatorens egenskaber ved at se på fordelingen over de estimerede parametre

1) Introduktion til problem set 4

Monte Carlo eksperimenter

- Fordelen ved dette er at vi kun skal specificere den datagenererende proces før vi kan analysere en estimators egenskaber
- Ulempen er at vi ikke kan generalisere disse egenskaber da vi kun kan være sikre på at de holder i det konkrete miljø vi har specificeret.
- MC-eksperimenter bruges derfor primært til enten at:
 - i) Sammenligne forskellige estimators i det samme miljø
 - ii) Undersøge egenskaberne for én estimator i forskellige miljøer

1) Introduktion til problem set 4

Monte Carlo eksperimenter

0. Specificer den datagenerende proces:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i, \quad (1)$$

$$\beta_0 = 1, \beta_1 = 2, \beta_2 = -3, \quad (2)$$

$$x_1 \sim N(25, 25), u \sim U(-50, 50), x_2^* \sim U(10, 30) \quad (3)$$

$$x_2 = \rho x_1 + x_2^*, \rho = 0.5, n = 50 \quad (4)$$

1. Simulér data. Dvs. dan et fiktiv sample med n observationer trukket tilfældigt fra den datagenerende proces.
2. Udregn relevante størrelser (estimer, standardfejl mv.) ved at bruge det sample der netop er dannet, og gem disse.
3. Gentag 1. og 2. mange gange. Gør vi fordi vi er interesseret i de generelle egenskaber for estimatorerne – ikke blot hvordan de opfører sig for ét bestemt sample.
4. Undersøg fordelingen af estimer ved brug af summary statistics (middelværdi, varians mv.)
5. Fx kan vi se om fordelingen af estimer er fordelt omkring den sande værdi (*unbiased*), hvorvidt fordelingen kollapser om den sande værdi hvis vi øger n (*konsistens*), eller hvorvidt estimatorne er mindre spredt omkring den sande værdi for én estimator en anden (*efficiens*).

1) Introduktion til problem set 4

- I PS4 bruger vi MC simuleringer til at analysere hvordan bias og varians for estimatorer påvirkes under forskellige (mis)specifikationer.
- Vi sammenholder det vi finder med teoretiske resultater

2) Gruppearbejde

To-do næste 20 minutter:

1. Discuss whether the specified data generating process (DGP) satisfies assumptions MLR.1–MLR.4.
2. How can a Monte Carlo experiment be used to compare the variances of two different unbiased estimators?

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i, \quad (1)$$

$$\beta_0 = 1, \beta_1 = 2, \beta_2 = -3, \quad (2)$$

$$x_1 \sim N(25, 25), u \sim U(-50, 50), x_2^* \sim U(10, 30) \quad (3)$$

$$x_2 = \rho x_1 + x_2^*, \rho = 0.5, n = 50 \quad (4)$$

Brug for hjælp? Skriv Jeres gruppe nummer i Padlet (link i chat)

3) Opsamling

1. Discuss whether the specified data generating process (DGP) satisfies assumptions MLR.1–MLR.4.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i, \quad (1)$$

$$\beta_0 = 1, \beta_1 = 2, \beta_2 = -3, \quad (2)$$

$$x_1 \sim N(25, 25), u \sim U(-50, 50), x_2^* \sim U(10, 30) \quad (3)$$

$$x_2 = \rho x_1 + x_2^*, \rho = 0.5, n = 50 \quad (4)$$

3) Opsamling

1. Discuss whether the specified data generating process (DGP) satisfies assumptions MLR.1–MLR.4.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i, \quad (1)$$

$$\beta_0 = 1, \beta_1 = 2, \beta_2 = -3, \quad (2)$$

$$x_1 \sim N(25, 25), u \sim U(-50, 50), x_2^* \sim U(10, 30) \quad (3)$$

$$x_2 = \rho x_1 + x_2^*, \rho = 0.5, n = 50 \quad (4)$$

MLR.1 (lineær i parametre): Opfyldt jf. ligning (1) ovenfor.

MLR.2 (tilfældig stikprøve): Opfyldt da variablene $(x_1, x_2^*$ og $u_i)$ er dannet ud fra tilfældige træk fra hhv. en normalfordeling og uniformfordeling.

MLR.3 (ingen perfekt multikollinearitet): Opfyldt da x_1 og x_2 er korreleret, men ikke perfekt korreleret.

MLR.4 ($E(u|x) = 0$): Opfyldt da x_1 og x_2^* er trukket uafhængigt af u . Formelt har vi $E(u|x) = E(u) = \frac{-50+50}{2} = 0$. De uafhængige variable er altså eskogene.

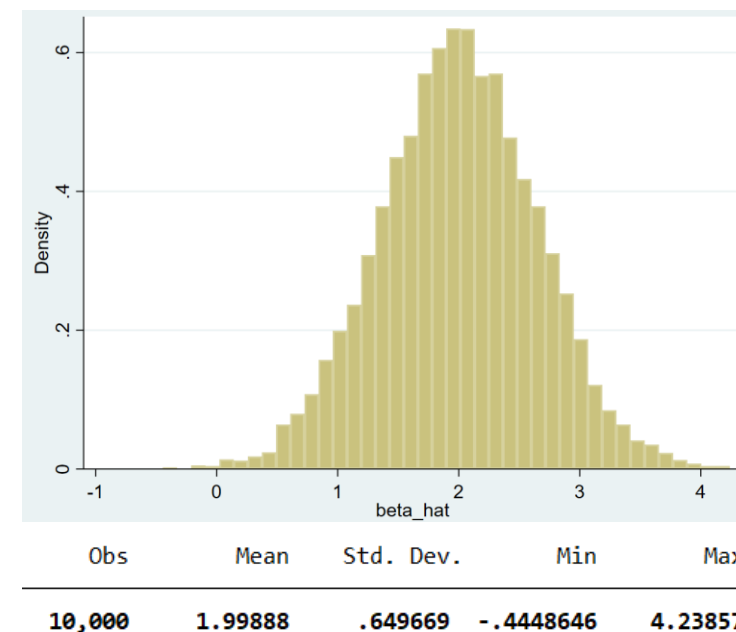
3) Opsamling

2. How can a Monte Carlo experiment be used to compare the variances of two different unbiased estimators?

3) Opsamling

2. How can a Monte Carlo experiment be used to compare the variances of two different unbiased estimators?

- I et Monte Carlo eksperiment simulerer vi data og estimerer parametrene **mange** gange.
- Med andre ord opnår vi mange forskellige estimater af fx $\widehat{\beta}_1$. Vi kan se på gennemsnittet af disse for at danne os et indtryk af hvorvidt estimatoren er unbiased eller ej.
- På samme måde kan vi istedet for at finde analytiske løsninger for estimatorens varians ($\sigma^2(X'X)^{-1}$), se på variansen for fordelingen af vores estimater.
- Det er en fordel i tilfælde af besværlige eller ikke-eksisterende udtryk for variansen.
- Hvis flere estimatorer er unbiased kan vi finde den mest efficiente ved at sammenligne deres varians.
- Hvis den ene estimator er unbiased og den anden er biased vil man skulle vurdere hvilken estimator man foretrækker ud fra et trade-off mellem størrelserne på bias og varians. *Det ser vi bort fra i dette fag, hvor vi altid foretrækker unbiased estimater.*



4) Gruppearbejde

To-do:

Arbejd på Stataøvelserne i PS4

Hints:

- 1) Koden I har fået udleveret udfører 10,000 repetitioner. For reducere tidsforbruget anbefaler jeg at i sætter det til 1,000 i stedet (reps=1000). Det ændrer ikke på konklusionerne.
- 2) Udover summarize kan I plotte data for at få et visuelt indtryk af hvad der sker. En måde at plotte to histogrammer ved siden af hinanden er:

```
twoway (histogram betahat_slr, recast(line) color(red)) ///  
       (histogram betahat_mlr, recast(line) color(blue)) ///  
       , legend(order(1 "SLR" 2 "MLR"))
```

Brug for hjælp? Skriv Jeres gruppenummer i Padlet ([link i chat](#))

5) Opsamling

1. Go to Absalon and download the STATA do-file associated with problem set 4. This do-file implements the Monte Carlo experiment outlined on the previous page using **simulate**.

- Vi begynder programmet med at give det et navn (*olsdata*). Hver gang vi kalder *olsdata* kører Stata alle kommandoer angivet i programmet.
- I det her tilfælde trækkes et sample på 50 observations baseret på den datagenerende proces. Herefter udregnes estimater for x_1 både for den simple lineære regression $\widehat{\beta}_1^{SLR}$ og den multiple lineære regression $\widehat{\beta}_1^{MLR}$.
- I den datagenerende proces trækker vi tilfældige tal fra:
rnormal(a,b): En normalfordeling med middelværdi a og standardafvigelse b
runiform(a,b): En uniform fordeling over intervallet (a, b)
- Programmet estimerer $\widehat{\beta}_1^{SLR}$ og $\widehat{\beta}_1^{MLR}$ med OLS og gemmer estimaterne med *return scalar b1_slr=_b(x1)* hvor *x1* referer til navnet på variablen hvis koefficient vi er interesseret i. Vi vælger selv navnet til dette estimat (*b1_slr*).
- Vi slutter programmet med *end* og tester altid om alt fungerer ved at køre programmet én gang. Det gøres ved at kalde programmet med det navn vi har givet det (*olsdata*).

Andet:

- *Drop_all* fjerner al information hver gang programmet køres
- *rclass* specificerer hvordan programmets output skal gemmes (ikke så vigtigt)

```
*DEFINE PROGRAM THAT SPECIFIES THE DGP
program olsdata, rclass
    drop _all

    *SET NUMBER OF CURRENT OBSERVATIONS
    set obs 50

    *DATA GENERATING PROCESS
    generate x1 = rnormal(25,5)
    generate x2 = 0.5*x1+runiform(10,30)
    generate u = runiform(-50,50)
    generate y = 1+2*x1-3*x2+u

    *CALCULATE OLS ESTIMATES AND OLS SLOPE IN B1
    regress y x1
    return scalar b1_slr=_b[x1]
    regress y x1 x2
    return scalar b1_mlr=_b[x1]
end

*RUN PROGRAM ONE TIME
olsdata
```

5) Opsamling

2. Run the Monte Carlo experiment. Compare the results for $\tilde{\beta}_1$ and $\hat{\beta}_1$. Do the experimental results correspond to what you would expect in theory?

```
*NOW RUN PROGRAM 10000 TIMES USING SIMULATE  
simulate betahat_slr=r(b1_slr) betahat_mlr=r(b1_mlr), seed(117) reps(10000) nodots:olsdata
```

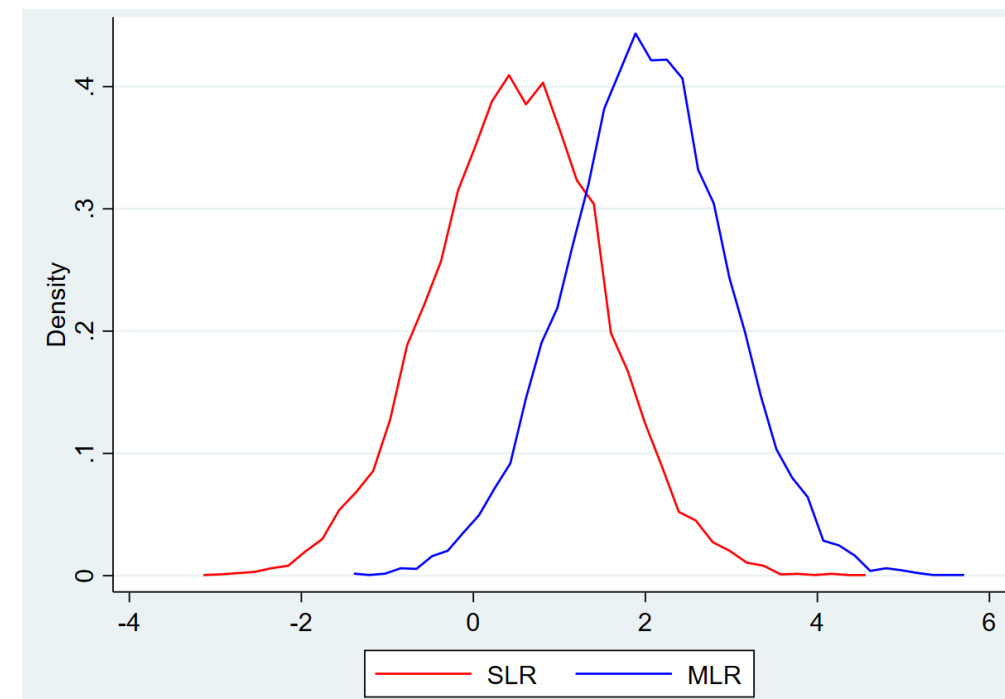
- *simulate* er en kommando der tillader os at kalde vores program et arbitrært antal gange og samtidigt gemme estimaterne fra være enkel iteration.
- Først angiver hvad vi ønsker at gemme fra hver iteration. I det her tilfælde gemmer vi de to koefficienter $\widehat{\beta}_1^{SLR}$ og $\widehat{\beta}_2^{MLR}$.
- Herefter (efter kommaet) angiver vi:
 - seed()*: Tilfældigt valgt værdi der definerer sekvensen af (pseudo) tilfældige træk i den datagenerende proces (samme seed, samme træk)
 - reps()*: Angiver hvor mange gange vi ønsker at simulere datatræk og estimere koefficienter.
 - nodots()*: Fortæller Stata at den ikke skal printe en prik for hver iteration.
 - :olsdata*: Specificerer at det er programmet *olsdata* som skal kaldes ved hver iteration.

5) Opsamling

2. Run the Monte Carlo experiment. Compare the results for $\tilde{\beta}_1$ and $\hat{\beta}_1$. Do the experimental results correspond to what you would expect in theory?

Simulerede egenskaber:

- Vi ser at $\widehat{\beta}_1^{MLR}$ er unbiased da estimerne i gennemsnit er lig den sande værdi ($\beta_1 = 2$)
- Omvendt er $\widehat{\beta}_1^{SLR}$ nedad biased da denne i gennemsnit lig 0,5
- Da vi er interesseret i den bedste unbiased estimator sammenligner vi ikke varianserne på de to estimatorer.



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.5066376	.9816975	-3.237781	4.658902
betahat_mlr	10,000	2.003952	.928953	-1.481017	5.798132

5) Opsamling

- Run the Monte Carlo experiment. Compare the results for $\tilde{\beta}_1$ and $\hat{\beta}_1$. Do the experimental results correspond to what you would expect in theory?

Teoretiske egenskaber (asymptotisk bias):

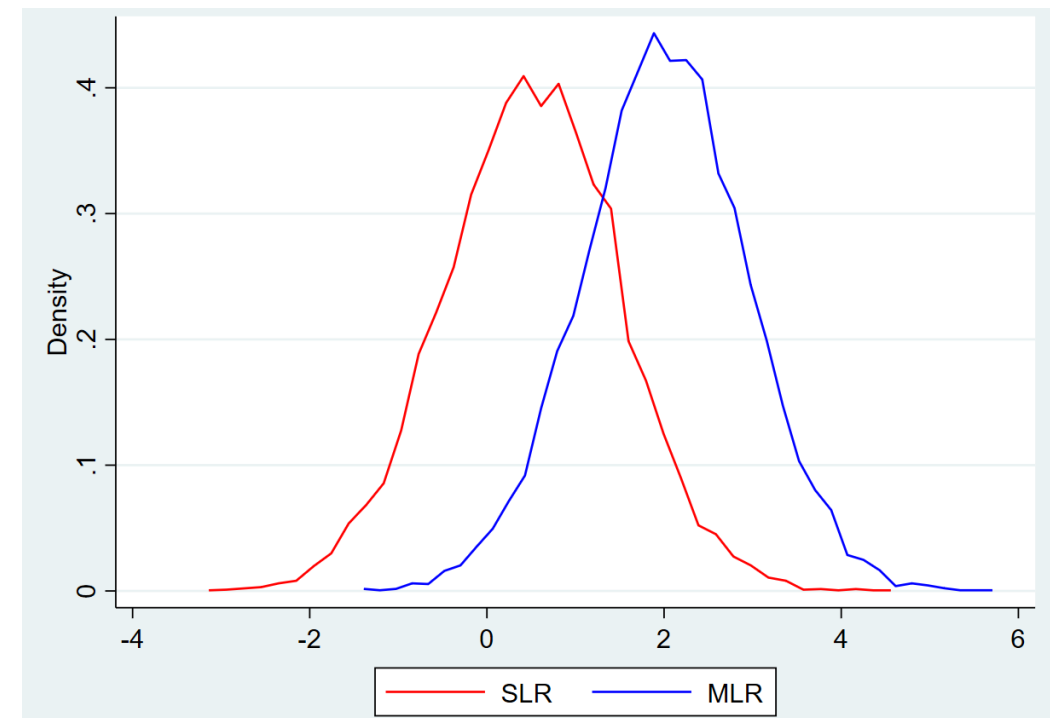
MLR unbiased da:

$$\widehat{\beta}_1^{MLR} - \beta_1 = \frac{cov(x_1, u)}{var(x_1)} = \frac{0}{var(x_1)} = 0$$

SLR biased da:

$$\begin{aligned} \widehat{\beta}_1^{SLR} - \beta_1 &= \frac{cov(x_1, \tilde{u})}{var(x_1)} = \frac{cov(x_1, \beta_2 x_2 + u)}{var(x_1)} = \frac{cov(x_1, \beta_2 \rho x_1 + \beta_2 x_2^* + u)}{var(x_1)} \\ &= \frac{\beta_2 \rho cov(x_1, x_1)}{var(x_1)} + \frac{\beta_2 cov(x_1, x_2^*)}{var(x_1)} + \frac{cov(x_1, u)}{var(x_1)} = \beta_2 \rho \frac{var(x_1)}{var(x_1)} = \beta_2 \rho \end{aligned}$$

Da $\beta_2 \rho = -3 \cdot 0,5 = -1,5$ ser vi netop at $\widehat{\beta}_1^{SLR}$ er nedad biased med -1,5 (omitted variable bias)

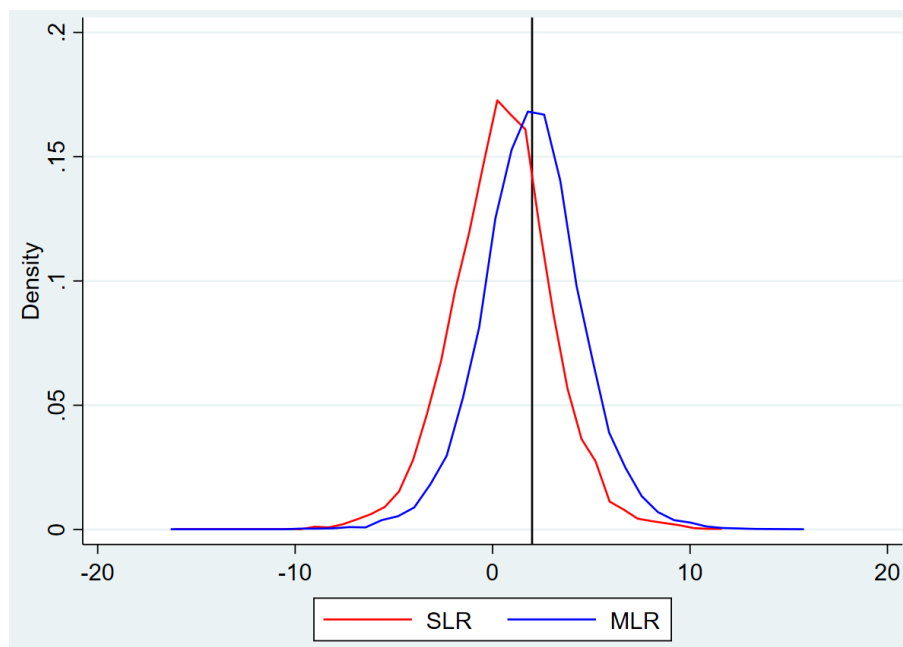


Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.5066376	.9816975	-3.237781	4.658902
betahat_mlr	10,000	2.003952	.928953	-1.481017	5.798132

5) Opsamling

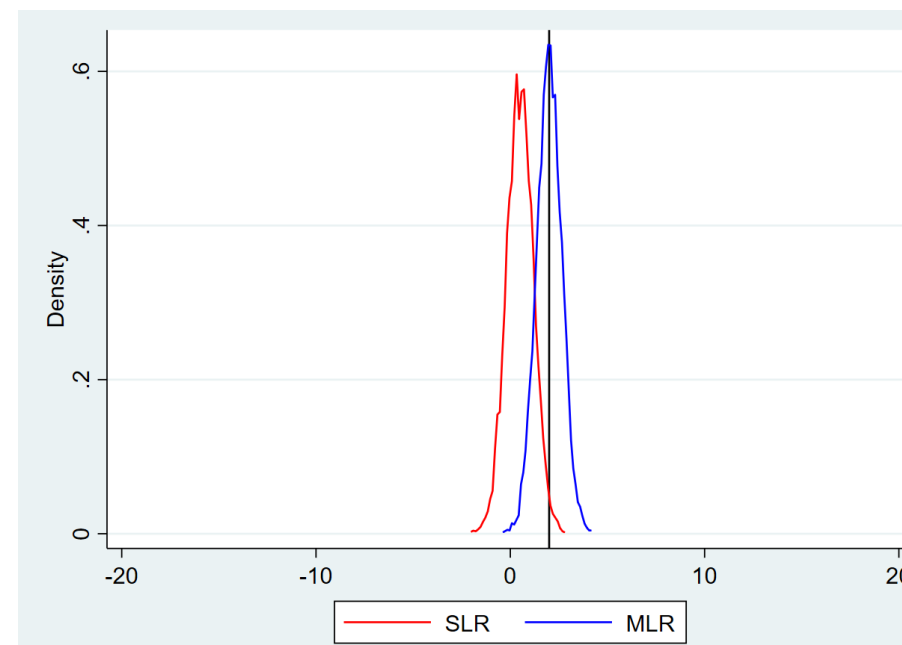
3. Run Monte Carlo experiments with $n = 10$ and $n = 100$. Compare the results for $\tilde{\beta}_1$ and $\hat{\beta}_1$. Are both $\tilde{\beta}_1$ and $\hat{\beta}_1$ consistent estimators of β_1 ?

$n = 10$ (set obs 10)



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.5046847	2.538333	-16.47147	11.97137
betahat_mlr	10,000	1.986183	2.550523	-16.71782	16.18095

$n = 100$ (set obs 100)



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.4971294	.685422	-2.090276	2.875104
betahat_mlr	10,000	1.99888	.649669	-.4448646	4.23857

5) Opsamling

3. Run Monte Carlo experiments with $n = 10$ and $n = 100$. Compare the results for $\tilde{\beta}_1$ and $\hat{\beta}_1$. Are both $\tilde{\beta}_1$ and $\hat{\beta}_1$ consistent estimators of β_1 ?

Konsistens:

- $\widehat{\beta}_1^{MLR}$ er konsistent da den konvergerer mod den sande værdi ($\beta_1 = 2$) når n vokser (fordelingen af estimer kollapser om den sande værdi)
- $\widehat{\beta}_1^{SLR}$ er ikke konsistent da den konvergerer mod 0,5 når n vokser (fordelingen af estimer kollapser om 0,5)

Større stikprøve, lavere varians

- Standardfejlen $se(\hat{\beta})$ falder *approksimativt* med en rate der svarer til $\frac{1}{\sqrt{n}}$
- Standardfejlen på vores estimer er derfor ca. $\sqrt{\frac{100}{10}} = 3,2$ gange større når $n = 10$ end når $n = 100$. Passer nogenlunde:

$$\frac{se_{n=10}(\widehat{\beta}^{SLR})}{se_{n=100}(\widehat{\beta}^{SLR})} = \frac{2,54}{0,69} = 3,7 \quad \text{og} \quad \frac{se_{n=10}(\widehat{\beta}^{MLR})}{se_{n=100}(\widehat{\beta}^{MLR})} = \frac{2,55}{0,65} = 3,9$$

5) Opsamling

4. Run Monte Carlo experiments with $\rho = 1$ and $\rho = 0$, assuming $n = 50$. What happens to the variances of $\hat{\beta}_1$ and $\hat{\beta}_1$? Is multicollinearity a problem? Does the asymptotic bias decrease or increase — if so, how and why?

- **ρ måler samvariationen** mellem x_1 og x_2 . Når $\rho = 1$ har vi:

$$x_2 = x_1 + x_2^*, \quad x_2^* \sim U(10,30)$$

- **Multikollinearitet:** Ikke et problem da x_2 og x_1 ikke er lineært afhængige på trods af den høje samvariation. Skyldes x_2^* som varierer tilfældigt.

- **Asymptotisk bias:** Kun MLR er unbiased da,

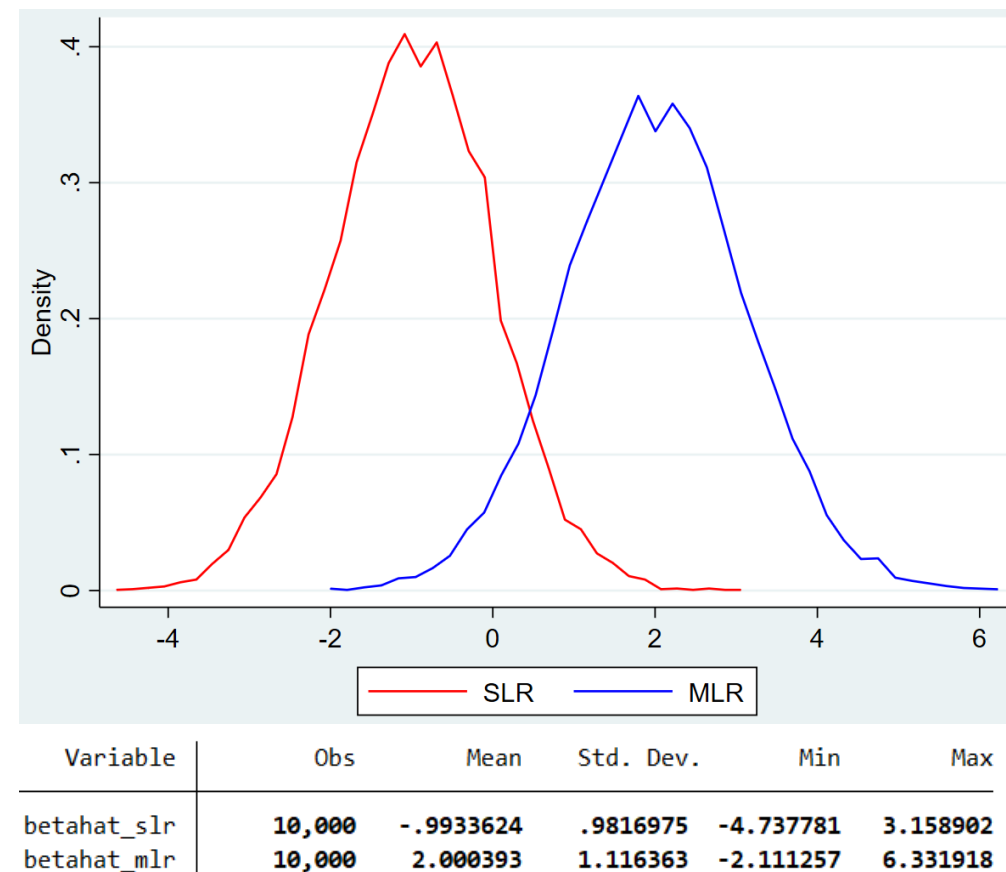
$$\widehat{\beta}_1^{MLR} - \beta_1 = 0$$

$$\widehat{\beta}_1^{SLR} - \beta_1 = \rho\beta_2 = 1(-3) = -3$$

- **Varians:** Kun $\widehat{\beta}_1^{MLR}$ er unbiased, så vi undlader at sammenligne variansen med $\widehat{\beta}_1^{SLR}$ i dette tilfælde. Til gengæld kan vi sammenligne:

$$se_{\rho=1}(\widehat{\beta}_1^{MLR}) = 1,12 \quad \text{og} \quad se_{\rho=0,5}(\widehat{\beta}_1^{MLR}) = 0,93$$

- Da samvariationen mellem x_1 og x_2 er højere når $\rho = 1$ er $se_{\rho=1}(\widehat{\beta}_1^{MLR})$ højere end $se_{\rho=0,5}(\widehat{\beta}_1^{MLR})$. Det skyldes at vi i den multiple regression isolerer den del af variationen i x_1 som ikke kan forklares af x_2 . Da de to variable har høj samvariation er den 'rene' del af variationen i x_1 mindre, hvilket øger estimatorens varians.



5) Opsamling

- Run Monte Carlo experiments with $\rho = 1$ and $\rho = 0$, assuming $n = 50$. What happens to the variances of $\hat{\beta}_1$ and $\hat{\beta}_1$? Is multicollinearity a problem? Does the asymptotic bias decrease or increase — if so, how and why?

- ρ måler samvariationen mellem x_1 og x_2 . Når $\rho = 0$ har vi:

$$x_2 = x_2^*, \quad x_2^* \sim U(10,30)$$

- Multikollinearitet:** Ikke et problem da x_2 og x_1 er uafhængige.

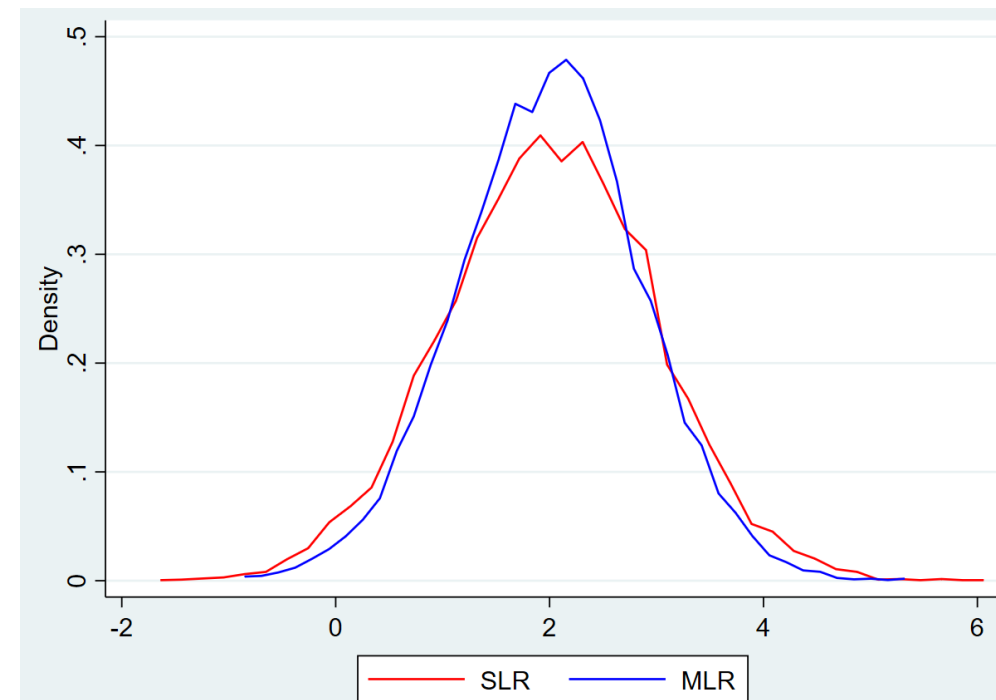
- Asymptotisk bias:** Begge estimatorer er unbiased da,

$$\widehat{\beta}_1^{MLR} - \beta_1 = 0$$

$$\widehat{\beta}_1^{SLR} - \beta_1 = \rho\beta_2 = 0(-3) = 0$$

Bemærk er x_2 stadig en del af fejlleddet i den simple model, men x_1 samvarierer ikke længere med den. Det er en general egenskab at bias kræver både $\rho \neq 0$ og $\beta_2 \neq 0$.

- Varians:** Begge estimatorer ($\widehat{\beta}_1^{MLR}$ og $\widehat{\beta}_1^{SLR}$) er unbiased. Vi kan derfor sammenligne deres varians. Da variansen for $\widehat{\beta}_1^{MLR}$ er lavest foretrækker vi denne.



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	2.006638	.9816975	-1.737781	6.158902
betahat_mlr	10,000	2.007511	.8588401	-.9324002	5.400814

5) Opsamling

Hvorfor er variansen for $\widehat{\beta}_1^{MLR}$ lavere?

- Der er to effekter på variansen ved at gå fra SLR til MLR:
 - 1)  Når vi inkluderer x_2 forklarer vi en større del af variationen i y (mindre efterladt i u). Det betyder i praksis ofte at $\hat{\sigma}^2 = \frac{SSR}{n-k-1}$ er lavere i MLR end i SLR, og dermed at estimatorens varians falder. Bemærk dog at vi ikke kan være helt sikre på dette pga. frihedsgradskorrektionen i nævneren af $\hat{\sigma}^2$ (k er større i MLR end SLR).
 - 2)  Når vi inkluderer x_2 isolerer vi den del af variationen i x_1 som ikke kan forklares af x_2 . Da de to variable har positiv samvariation er den 'rene' del af variationen i x_1 mindre end hele variationen i x_1 hvilket øger estimatorens varians.
- I tilfældet hvor $\rho = 0$ finder kun effekten fra 1) sted da x_1 og x_2 ikke samvarierer. Den analytiske varians for estimatorerne er givet ved:

$$var(\widehat{\beta}_1^{MLR}) = \frac{\hat{\sigma}_{MLR}^2}{SST_1(1-R_1^2)} \quad \text{og} \quad var(\widehat{\beta}_1^{SLR}) = \frac{\hat{\sigma}_{SLR}^2}{SST_1}$$

Her er variansen for $\widehat{\beta}_1^{MLR}$ lavere end $\widehat{\beta}_1^{SLR}$ da $\hat{\sigma}_{MLR}^2$ (ofte) er lavere end $\hat{\sigma}_{SLR}^2$ (da fejleddet i SLR-modellen indeholder $\beta_2 x_2$), samtidigt med at $R_1^2 \approx 0$ (R_1^2 er R^2 -værdien i regressionen af x_1 på x_2)

- Hvis udeladte variable (x_2) er relevante ($\beta_2 \neq 0$) og samtidigt ikke samvarierer ($\rho = 0$) med den forklarende variabel vi er interesseret i (x_1) kan vi (ofte) reducere variansen på estimatoren vi er interesseret i ($\widehat{\beta}_1$) ved at inkludere den udeladte variable.

5) Opsamling

5. Run a Monte Carlo experiment, assuming $\rho = 0.5$ and $\beta_2 = 0$. Comparing $\tilde{\beta}_1$ and $\hat{\beta}_1$, which is your preferred estimator?

Når $\beta_2 = 0$ er begge estimatorer unbiased:

$$\widehat{\beta}_1^{MLR} - \beta_1 = \frac{cov(x_1, u)}{var(x_1)} = 0$$

$$\widehat{\beta}_1^{SLR} - \beta_1 = \rho\beta_2 = 0.5 \cdot 0 = 0$$

Det skyldes at x_2 nu er irrelevant da den ikke har nogen effekt på y når $\beta_2 = 0$. $\widehat{\beta}_1^{SLR}$ er derfor unbiased på trods af at x_1 samvarierer med den udeladte variabel x_2 . Det er en general egenskab at bias kræver bade $\rho \neq 0$ og $\beta_2 \neq 0$

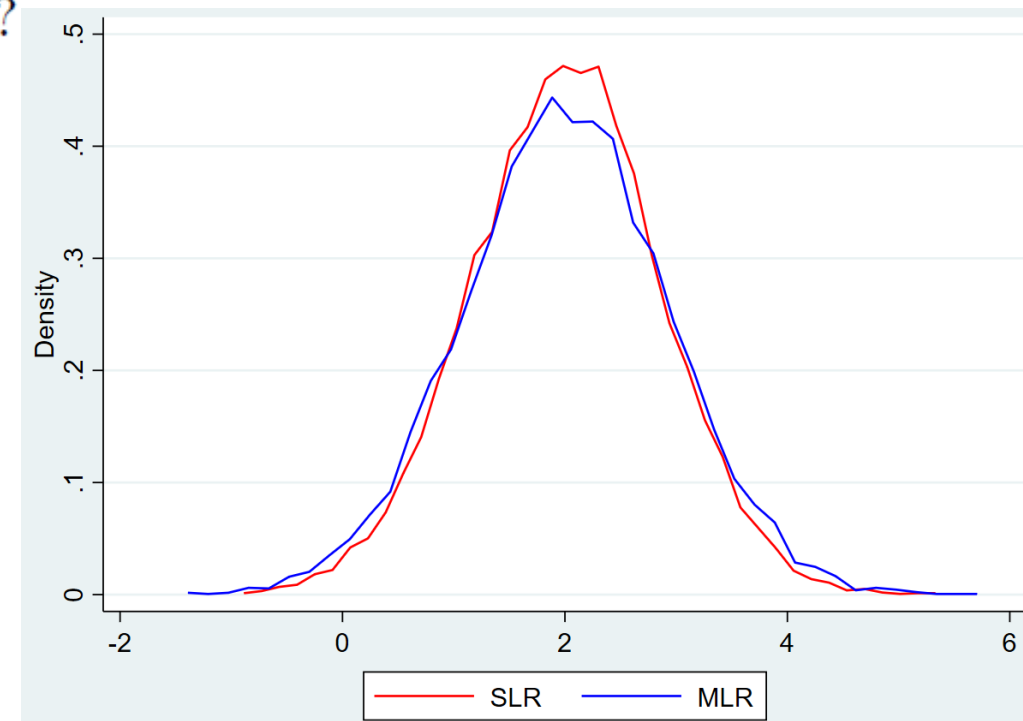
I dette tilfælde er standardfejlen for $\widehat{\beta}_1^{SLR}$ lavere end for $\widehat{\beta}_1^{MLR}$:

$$se(\widehat{\beta}_1^{SLR}) = 0,85$$

$$se(\widehat{\beta}_1^{MLR}) = 0,93$$

Skyldes at det kun er effekt 2) fra forrige slide der finder sted når $\beta_2 = 0$ og $\rho = 0,5$ og vi går fra SLR til MLR.

Generelt: Variansen på estimatoren som vi er interesseret i ($\widehat{\beta}_1$) stiger hvis vi inkluderer irrelevante ($\beta_2 = 0$) variable som samvarierer med variabelen vi interesserer os for (x_1)

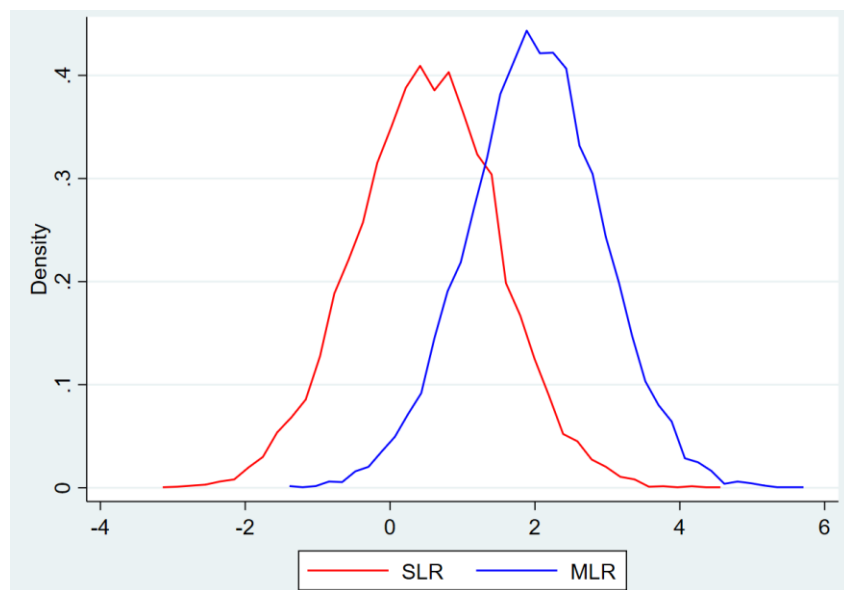


Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	2.006603	.8489552	-.9666878	5.411431
betahat_mlr	10,000	2.003952	.928953	-1.481017	5.798132

5) Opsamling

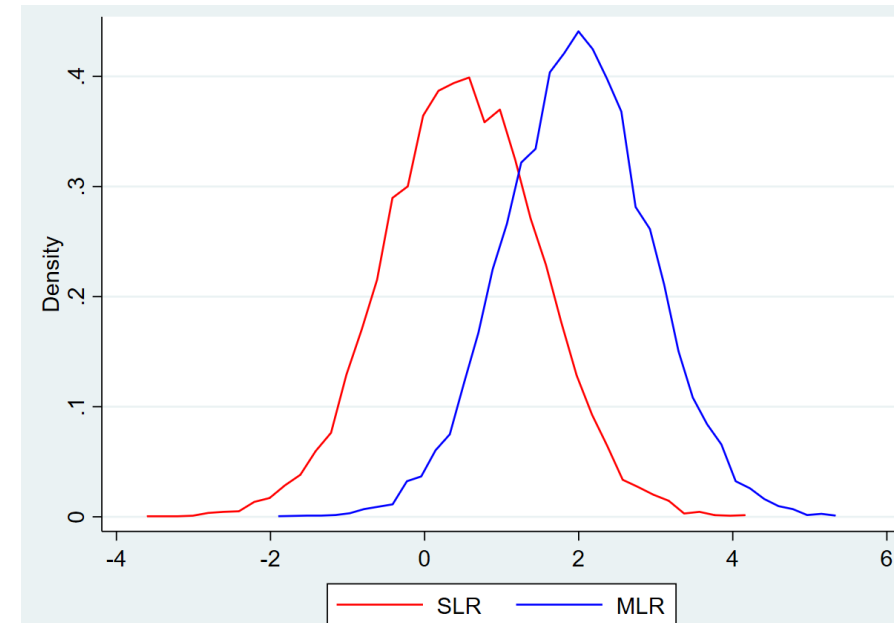
6. Run the baseline Monte Carlo experiment with a different seed number. That is, assume $\beta_2 = -3$ and set the seed number to e.g. 1479 (the founding year of University of Copenhagen). What do you conclude?

Seed=117



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.5066376	.9816975	-3.237781	4.658902
betahat_mlr	10,000	2.003952	.928953	-1.481017	5.798132

Seed=1479



Variable	Obs	Mean	Std. Dev.	Min	Max
betahat_slr	10,000	.4863054	.9909912	-3.70497	4.264863
betahat_mlr	10,000	1.983969	.9325859	-1.99193	5.428911

- Pseudo-randomness: Vores "tilfældige" træk fra fordelinger i den datagenerende proces afhænger af det valgte seed nummer. Samme seed giver samme "tilfældige" træk.
- Konklusioner ændrer sig ikke ved at skifte seed. Det er et godt tegn. Resultater bør ikke afhænge af det valgte seed nummer.
- Husk at rapportere Jeres seed nummer i opgaver/eksamenen så folk kan genskabe Jeres resultater præcist.

Tak for i dag

- Næste gang gennemgår vi obligatorisk opgave 1 (skal senest afleveres i aften) inden vi begynder på PS5
- Dagens slides bliver lagt op fredag.