QE 2021

①

ⓐ

~~No~~ Uniform density : $f_X(x) = \dfrac{1}{b-a}$ for $a \leq x \leq b$

$$f_X(x) = \frac{1}{1-(-1)} = \frac{1}{2}$$

$$\mathbb{E}[X_i^k] = \int_{-1}^{1} x^k \frac{1}{2} dx = \left[\frac{1}{k+1} x^{k+1} \frac{1}{2}\right]_{-1}^{1}$$

$$\left[\frac{1}{2(k+1)} 1^{k+1}\right] - \left[\frac{1}{2(k+1)} (-1)^{k+1}\right]$$

$$\cancel{(-1)^{k+1} = 1 \text{ if } k \text{ is odd}}$$

$$(-1)^{k+1} = \begin{cases} 1 & \text{if } k \text{ is odd} \\ -1 & \text{if } k \text{ is even.} \end{cases}$$

k is odd

$$\mathbb{E}[X_i^k] = \frac{1}{2(k+1)} - \frac{1}{2(k+1)} = 0$$

k is even

$$\mathbb{E}[X_i^k] = \frac{1}{2(k+1)} + \frac{1}{2(k+1)} = \frac{\cancel{2}}{\cancel{2}(k+1)} = \frac{1}{k+1}$$

$$\mathbb{E}[X_i^k] = \begin{cases} 1/k+1 & k = \text{even} \\ 0 & k = \text{odd.} \end{cases}$$

(b)

$$\text{var}(X_i^k) = \mathbb{E}[X_i^{2k}] - (\mathbb{E}[X_i^k])^2$$

$k = \text{odd}$ hence

$$\mathbb{E}[X_i^k] = 0 \implies (\mathbb{E}[X_i^k])^2 = 0$$

$$\text{var}(X_i^k) = \mathbb{E}[X_i^{2k}]$$

$2k = \text{even}.$

let $2k = l$

$\mathbb{E}[X_i^l] = \frac{1}{1+l}$ for $l = \text{even}.$

hence

$$\mathbb{E}[X_i^{2k}] = \frac{1}{1+2k}.$$

$$\text{var}(X_i^k) = \frac{1}{1+2k}$$

(c) Lindeberg-Levy CLT

Given that X_i^3 has finite mean: $\mathbb{E}[X_i^3] = \mu = 0$
and variance $\text{var}(X_i^3) = \frac{1}{1+2\cdot 3} = \frac{1}{7}. = \sigma^2$

then as $n \to \infty$

$$\frac{\sqrt{n}(\overline{X_i^3} - \mu)}{\sigma} \xrightarrow{D} N(0,1)$$

$$= \frac{\sqrt{n}\left(\frac{1}{n}\sum X_i^3 - \mu\right)}{\sigma} = n^{-1/2}\sum_{i=1}^{n} X_i^3 \xrightarrow{D} N(0, 1/7)$$

(2)

(a)

Unit roots:

$$\Delta W_t = \mu + \phi W_{t-1} + \gamma \Delta W_{t-1} + u_t.$$

$$W_t - W_{t-1} = \mu + (\phi + \gamma) W_{t-1} - \gamma W_{t-2} + u_t.$$

$$W_t = \mu + (\phi + \gamma + 1) W_{t-1} - \gamma W_{t-2} + u_t.$$

unit root if $\phi + \gamma + 1 - \gamma = 1$

or if $\underline{\phi = 0}$

$\therefore$ test:

$H_0 : \phi = 0$

$H_1 : \phi < 0$

ADF test constant only

(since no trend in these models)

$$t = \frac{\hat{\phi}}{s.e.(\hat{\phi})} \xrightarrow{D} DF_{cn}$$

Decision rule: reject H_0 if $t < C_\alpha$

(one sided since $\phi > 0 \Rightarrow$ explosive process which is implausible: never seen in macro data)

X_t:

t-stat: $t = -3.99$

$-3.99 < -3.43$ (DF $cv_{0.01}$)

$\therefore$ reject H_0

hence $\{X_t\}$ ~~has unit root~~.

there is evidence ~~at~~ at 1% significance that $\{X_t\}$ is stationary ~~has unit root~~

Y_t :

$t = -2.04$

$-2.04 \not< -2.86$ (DF crit. val. at 10%)

$\therefore$ do not reject H_0

~~$\Rightarrow \{Y_t\}$ is~~

evidence suggests $\{Y_t\}$ ~~is stationary~~ has unit root.

Z_t :

$t = -2.23$

$-2.23 \not< -2.86$ (DF crit. val. at 10%)

$\therefore$ do not reject H_0

$\therefore$ evidence suggests $\{Z\}$ ~~is stationary~~ has unit root.

(b)

Spurious regression : systematic tendency to find statistically significant regression relationships between $I(1)$ time series.

X_t on Y_t ?

Not spurious since $X_t \sim I(0)$ (stationary)

Z_t on Y_t ?

Both are $I(1)$ time series since we ~~reje~~ ~~accept~~ do not reject $H_0: \phi = 1$ for $\{Z_t\}$ and $\{Y_t\}$ but do reject $H_0: \phi = 1$ for $\{\Delta Z_t\}$ and $\{\Delta Y_t\}$

[t stats : $\underset{(\Delta Y_t)}{-6.18}$ and $\underset{(\Delta Z_t)}{-7.68}$ < -3.43

$\therefore$ evidence to reject null at 1%]

and if differences are stationary ($I(0)$) then levels are $I(1)$.

But Z_t & Y_t share common stochastic trend U_t

∴ co-integrated ⇒ not spurious.

③

(a)

(i)

containing 2 children $\Rightarrow$ $D_i = 1$

(3) $Y_i = \beta_0 + 0.044X_i + \cancel{\beta_2}\,1.551 D_i + 0.025 X_i D_i + \hat{u}_i$

for 2 children $D_i = 1$

$\therefore$

$$Y_i = (\beta_0 + 1.551) + (0.044 + 0.025) X_i + \hat{u}_i$$

$\therefore$ increasing income by £1 would increase expenditure on food per day by £0.069 for families with 2 kids

(ii)

for 2 kids ΔY_i w.r.t. unit ΔX is 0.069
otherwise ΔY_i w.r.t. unit ΔX is 0.044

Mean value $D_i = 0.61$

$\therefore$ 61% houses do not have 2 kids, while 39% do.

$0.61 \times 1500 = 915$ $\quad$ $0.39 \times 1500 = 585$

$$\bar{Y} = \frac{\overset{915}{\cancel{585}} \cdot 0.069 + \overset{585}{\cancel{915}} \cdot 0.044}{1500}$$

$\boxed{\cancel{\bar{Y} = 0.05375}}$ $\quad$ $\boxed{\bar{Y} = 0.05925}$

(D)

test same for 2 children + not.

F-test:

$$H_0: \hat{\beta}_D = \hat{\beta}_{XD} = 0$$

$$H_1: \hat{\beta}_I \neq 0 \quad \text{for } \exists I \in \{D, XD\}$$

unrestricted model:

$$Y_i = \hat{\beta}_{0,un} + \hat{\beta}_{X,un} X_i + \hat{\beta}_{D,un} D_i + \hat{\beta}_{XD,un} X_i D_i + \hat{u}_{i,un}$$

restricted model:

$$Y_i = \hat{\beta}_{0,RS} + \hat{\beta}_{X,RS} + \hat{u}_{i,RS}$$

$$SSR_{un} = \sum_{i=1}^{n} \hat{u}_{i,un}^2 = 202154$$

$$SSR_{RS} = \sum_{i=1}^{n} \hat{u}_{i,RS}^2 = 211685$$

$$F = \frac{SSR_{RS} - SSR_{UN}}{SSR_{UN}} \frac{n-K-1}{q} \qquad q=2 \qquad K=3$$

$$n = 1500$$

Where $F \xrightarrow{D} F_{2,\infty}$

$$F = \frac{211685 - 202154}{202154} \cdot \frac{(1500-3-1)}{2} = \cancel{23.526}\ 35.3$$

reject if $F > C_\alpha$ $\qquad C_{5\%}$ for $F_{2,\infty} = 3$

$\cancel{23.526}\ 35.2 > 3$

∴ reject H_0, there is sufficient evidence at 5% sig to reject that the relationship between exp. on food & income is the same for households w one + two children.

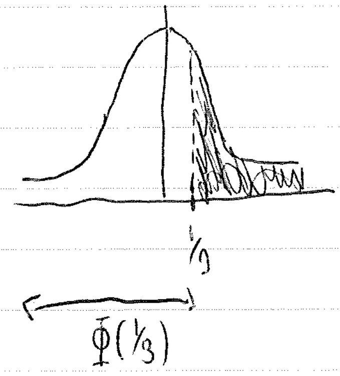
(4)

(a)

$$W \sim N(\mu, \sigma^2)$$

$$W_{i,09} \sim N(10, 3^2) \qquad \text{(by assumption)}$$

$$Z_{i,09} = \frac{W_{i,09} - 10}{3} \sim N(0,1)$$



$$Z^{11}_{i,09} = \frac{11-10}{3} = 1/3$$

$$P(Z_{i,09} > 1/3) = 1 - \Phi(1/3)$$

$$= 0.369$$

$$P(Z_{i,09} \geq 1/3) = 0.369$$

(b)

Study 1: $12.6 - 10.5 = \boxed{2.1\ (£/hr)}$ $\quad \frac{2.1}{10.5} = 0.2 \quad \therefore$ 20% increase for participants

Study 2: $12.4 - 10 = \boxed{2.4\ (£/hr)}$ $\quad \frac{2.4}{10} = 0.24 \quad \therefore$ 24% increase for participants.

(c)

Study 1:

C: control
T: treatment

$H_0: \mu_C = \mu_T \Rightarrow \mu_C - \mu_T = 0$

$H_1: \mu_C \neq \mu_T \Rightarrow \mu_C - \mu_T \neq 0$

$$t = \frac{\bar{W}_{C,11} - \bar{W}_{T,11} - 0}{\sqrt{\frac{\hat{\sigma}^2_{C,11}}{n_C} + \frac{\hat{\sigma}^2_{T,11}}{n_T}}} \xrightarrow{D} N(0,1)$$

$$t = \frac{10.5 - 12.6}{\sqrt{\frac{3.1^2}{50} + \frac{3.3^2}{50}}} = -3.279$$

Decision rule: reject if $|t| > CV_\alpha$

let $\alpha = 0.05$

$\therefore C_{0.05} = 1.96$ for $N(0,1)$

$|-3.279| > 1.96$

$\therefore$ ~~reject~~ evidence to reject H_0 at 5% significance.

$\therefore$ evidence to suggest mean earnings are affected by JTP

Study 2:

~~$H_0: \mu_{11} = \mu_{10} \Rightarrow \mu_{11} - \mu_{10} = 0$~~

~~$H_1: \mu_{11} \neq \mu_{10} \Rightarrow \mu_{11} - \mu_{10} \neq 0$~~

~~$t = \frac{\bar{W}_{11} - \bar{W}_{10} - 0}{\sqrt{\frac{\hat{\sigma}_{11}^2}{n_{11}} + \frac{\hat{\sigma}_{10}^2}{n_{10}}}} \xrightarrow{d} N(0,1)$~~

~~$H_0: \mathbb{E}[\bar{W}_{11} - \bar{W}_{10}] = 0$~~

$H_1: \mathbb{E}[\bar{W}_{11} - \bar{w}$

assumption: ~~studies~~ samples are independent!!

Study 2:

$$H_0: \ \mathbb{E}[W_{11} - W_{10}] = 0$$
$$H_1: \ \mathbb{E}[W_{11} - W_{10}] \neq 0$$

$$t = \frac{(\bar{W}_{11} - \bar{W}_{10}) - 0}{\sqrt{\frac{\hat{\sigma}^2_{W_{11}-W_{10}}}{n}}} \xrightarrow{D} N(0,1)$$

$$t = \frac{2.4}{\sqrt{\frac{4.3^2}{100}}} = 5.58$$

reject if $|t| > cv_\alpha$

$\alpha = 0.05 \ \therefore$ for $N(0,1)$ $cv_{0.05} = 1.96$

$5.58 > 1.96$

$\therefore$ sufficient evidence to reject H_0 that JTP had no effect on on mean earnings.

(d)

Strengths:

- Study 1:
  - Control group allows for comparasion between treatment & control to ensure that increase in hourly wage is not caused by a factor other than JTP
    - e.g. Economic upturn post 2008 crash could be a reason for rising wages.

- Study 2:
  - ~~St~~ Wage change recorded hence can tell how wage has changed from JTP.
  - Large sample.

Weaknesses

- Study 1:
  - No initial wage ∴ may be the case that treatment group had higher wage pre JTP than control group.

- Study 2:
  - No control ∴ Δ wages per hour could be due to ~~econor~~ state of economy / factors other than JTP.

- Would like to know:
  - Was assignment random?
    - If so then no selection bias + also negates weakness of Study 1 since RTC ⇒ no selection bias

    ∴ Study (1) would be very useful.

(e)

- estimated effects of JTP are 2.1 (£/hr) and 2.4 (£/hr) for study 1 and 2 respectively
- Further, as part (c) showed, these results are statistically significant.

HOWEVER:

- without knowing whether samples were randomly selected we cannot make the claim that the JTP will have a causal effect on (£/hr) when rolled out, since it could be the case that selection bias means that the wage increase may not be due to JTP

  [Endogeneity problem:
  - In Study 1 wage ↑ could be due to fact better workers opted for JTP
  - In study 2 wage ↑ could be due to economic upturn]

- Also not necessarily the case that JTP will work for all workers, may depend on their skills
- May not be scalable.

(5)

(a)

State dummies ~~control for~~ = proxy to control for ~~regional~~ unmeasured differences that might be linked to region, such as:

- Diet, standard of healthcare, pollution, exercise culture ...

Only include 49 to avoid problem of perfect multicollinearity.
If include all state dummies then the 49 will perfectly explain the 50th, in this case FWL fails since regressors perfectly explain one another, hence $\widetilde{\text{smoked}}_i$ from the regression

$$\text{Smoked}_i = \hat{A}_0 + \hat{A}_1 \text{age}_i + \sum_{k=1}^{50} \hat{A}_k D_{ki} + \cancel{X}\ \widetilde{\text{smoked}}$$

is very small $\therefore$ $\hat{\beta}_1 = \dfrac{\widehat{\text{cov}}(Y, \tilde{X}_1^{\wedge})}{\widehat{\text{var}}(\tilde{X}^{\wedge})}$

= low ~~$\Rightarrow$ large $\hat{\beta}_1$~~.

(b)

Causal effect: OR may not hold
$\hat{\text{cov}}(\text{smoked}_i, u_i) \neq 0$

Since other factors that ~~cause~~ effect birthweight may be correlated with smoking, such as excessive drinking or diet.
Hence $\hat{\beta}_1$ may pick up some of these effects on bweight$_i$, and hence

$\hat{\beta}_1$ does not estimate the causal effect of smoking on bweight.

(c)

Instrument may remove endogeneity issue and hence provide causal estimate of effect of smoking on bweight

Assumptions:

Z1: Relevance : $cov(smoking_i, tax_i) \neq 0$

- Plausible since tax levied will effect the price of cigarettes, which will in turn effect the amount of smoking.

Z2: Exogeneity : $cov(tax_i, u_i) = 0$

- Plausible since it is unlikely that cigarette tax is corr. with other factors that effect babies birthweight.

Z3: Exclusion : tax_i has no direct effect on $bweight_i$

- Plausible since it is hard to think of another way in which tax could effect bweight other than via smoking.

Empirical evidence?

- Regression [1] (tax on smoked)

$\hat{\beta}_{tax}$ = returned effect : $\hat{\beta}_{tax}$ -0.035
(0.005)

test : $H_0 : \beta_{tax} = 0$

$H_1 : \beta_{tax} \neq 0$

$$t = \frac{\hat{\beta}_{tax} - \beta_{tax}}{s.e(\hat{\beta}_{tax})} \xrightarrow{d} N(0,1)$$

$$t = \frac{-0.035 - 0}{0.005} = -7$$

reject H_0 if $|t| > cv_\alpha$ $\quad \alpha = 0.05$

$\therefore cv_{0.05} = 1.96$ for $N(0,1)$

$|-7| = 7 > 1.96$

$\therefore$ reject null that tax has no effect on smoking

$\therefore Cov(smoke_i, tax_i) \neq 0$

$\therefore$ Z1 is met empirically

- Can't use [4] to test Z3 since in [4] effect of tax_i on $bweight_i$ may be via $smoked_i$

~~Z3 that is:~~

~~Z3: Exclusion (tax does not enter in causal model)~~ $\neq$ $Cov(bweight_i$

- that is, assumption Z3 (exclusive) which says that tax_i cannot enter into causal model for $bweight_i$ <u>does not imply</u> that

$Cor(tax_i, bweight_i) = 0$!! (which is what we would be testing)

(d)

OLS & 2SLS return different s.e.'s due to different efficiency

- OLS is best linear unbiased estimator

$\therefore$ 2SLS is less efficient than OLS

$\Rightarrow$ returns ~~lower~~ higher s.e.()

(e)

(i)

Model [4]

$$\frac{\partial bweight}{\partial tax} = 20$$

$\therefore$ \$1 increase $\rightarrow$ 20g $\uparrow$ in bweight.

(ii)

Model [3]

$$\frac{\partial bweight}{\partial smoke} = -564$$

lower smoking by 20 p. points $\Rightarrow$ increase bweight by $0.2 \times 564 = 112.8g$.

(f)

ILS:

$$bweight_i = \beta_0 + \beta_1 smoked + \dots + u_i$$

$$smoked_i = \gamma_0 + \gamma_1 tax_i + \dots + v_i$$

$$bweight_i = \beta_0 + \beta_1\gamma_0 + \beta_1\gamma_1 tax_i + \dots + \varepsilon_i$$

let $\beta_1\gamma_1 = \pi_1$

we have:

$$\hat{\pi}_1 = \underset{(7.4)}{20}$$

$$\hat{\gamma}_1 = \underset{(0.05)}{-0.035}$$

hence

$$\hat{\beta}_1 = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{20}{-0.035} = \boxed{-571.43}$$

QE 2020

(1)

(a)

$$\cancel{Y_i = (1-p)^{1-y_i}(p)^{y_i} \quad \text{for } y \in \{0,1\}}$$

$$\cancel{\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n}(1-p)^{1-y_i}(p)^{y_i}}$$

$$\mathbb{E}[\bar{Y}] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_i]$$

$$Y_i = (1-p)^{(1-y)}(p)^{y} \quad \text{for } y \in \{0,1\}$$

$$\mathbb{E}[Y_i] = 1 \cdot p + 0 \cdot (1-p) = p$$

$$\mathbb{E}[\bar{Y}] = \frac{1}{n}\sum_{i=1}^{n} p = \frac{1}{n} np = \boxed{p}$$

$$\text{var}(\bar{Y}) = \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2}\text{var}\left(\sum_{i=1}^{n} Y_i\right)$$

$$\text{var}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n}\text{var}(Y_i) + \underbrace{2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\text{cov}(Y_i, Y_j)}_{\substack{\text{iid draws here} \\ \text{cov}(Y_i, Y_j) = 0 \ \forall i \neq j}}$$

$$\therefore \text{var}(\bar{Y}) = \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(Y_i)$$

$$\text{var}(Y_i) = \mathbb{E}[Y_i^2] - (\mathbb{E}[Y_i])^2 = \mathbb{E}[Y_i^2] - p^2$$

$$\mathbb{E}[Y_i^2] = p \cdot 1^2 + (1-p) \cdot 0^2 = p$$

$$\text{var}(Y_i) = \cancel{p(p-1)} \; p(1-p)$$

$$\text{var}(\bar{Y}) = \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{1}{n^2} n\, p(1-p)$$

$$\boxed{\text{var}(\bar{Y}) = \frac{p(1-p)}{n}}$$

(b)

$$\frac{\bar{Y} - \mathbb{E}[\bar{Y}]}{(\text{var}(\bar{Y}))^{1/2}} \qquad (\text{var}(\bar{Y}))^{1/2} = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$$

$\therefore$

$$= \frac{\sqrt{n}\,(\bar{Y} - \mathbb{E}[\bar{Y}])}{\sqrt{p(1-p)}} \qquad \sqrt{n}\,(\bar{Y} - \mathbb{E}[\bar{Y}]) \xrightarrow{D} N(0, \text{var}(\hat{Y}_i))$$

by Lindeberg-Levy CLT

$$\therefore \quad \sqrt{n}\,(\bar{Y} - \mathbb{E}[\bar{Y}]) \xrightarrow{D} N(0,\ p(1-p))$$

by Slutsky's theorem then

$$\boxed{\frac{\sqrt{n}(\bar{Y} - \mathbb{E}[\bar{Y}])}{\sqrt{p(1-p)}} \xrightarrow{D} N\left(0\,,\ \frac{p(1-p)}{p(1-p)}\right) = N(0,1)}$$

$\therefore$

$$\frac{\bar{Y} - \mathbb{E}[\bar{Y}]}{(\text{var}(\bar{Y}))^{1/2}} \xrightarrow{D} N(0,1).$$

(c)

$n = 100$

$p = 0.2$

$\cancel{|\bar{Y} - E[\bar{Y}]|}$

$$E[\bar{Y}] = 0.2$$

$$\text{var}(\bar{Y}) = \frac{(0.2)(0.8)}{100} = \frac{1}{625} \qquad \text{s.d.}(\bar{Y}) = \frac{1}{25}$$

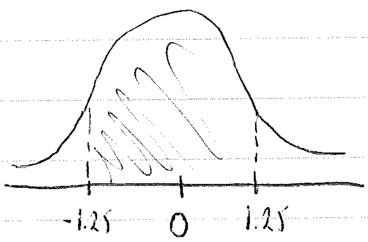
$$\bar{Y} \overset{a}{\sim} N\left(0.2\,,\ \frac{1}{625}\right)$$

$$Z = \frac{\bar{Y} - 0.2}{\sqrt{\frac{1}{625}}} \overset{a}{\sim} N(0,1)$$

$$Z = \frac{\bar{Y} - 0.2}{\frac{1}{25}} \overset{a}{\sim} N(0,1)$$

testing $\bar{Y} = \frac{15}{100} = 0.15$ and $\bar{Y} = \frac{25}{100} = 0.25$

$$Z_1 = \frac{0.15 - 0.2}{\frac{1}{25}} = -1.25 \qquad Z_2 = \frac{0.25 - 0.2}{\frac{1}{25}} = 1.25$$



$$P(0.15 \le \bar{Y} \le 0.25) = P(-1.25 \le Z \le 1.25)$$

$$\phi(1.25) - \phi(-1.25) = \boxed{0.789}$$

hence

$$\boxed{P(0.15 \le \bar{Y} \le 0.25) = 0.789}$$

②

(a)

$$MSFE\left(m(Y_t)\right) = \mathbb{E}\left[\left(Y_{t+1} - m(Y_t)\right)^2\right]$$

$$= \mathbb{E}\left[\left\{\left(Y_{t+1} - \mathbb{E}[Y_{t+1}|Y_t]\right) - \left(m(Y_t) - \mathbb{E}[Y_{t+1}|Y_t]\right)\right\}^2\right]$$

let $\varepsilon = Y_{t+1} - \mathbb{E}[Y_{t+1}|Y_t]$

$g(Y) = m(Y_t) - \mathbb{E}[Y_{t+1}|Y_t]$

$$= \mathbb{E}\left[\left(\varepsilon - G(Y_t)\right)^2\right] = \mathbb{E}[\varepsilon^2] \cancel{+} - 2\,\mathbb{E}\left[\varepsilon \cdot G(Y_t)\right] + \mathbb{E}\left[G(Y_t)^2\right]$$

$$\mathbb{E}\left[\varepsilon \cdot G(Y_t)\right] = \mathbb{E}\left[\mathbb{E}\left[\varepsilon \cdot G(Y_t) \mid Y_t\right]\right] = \mathbb{E}\left[G(Y_t) \cdot \mathbb{E}[\varepsilon \mid Y_t]\right]$$

$$\mathbb{E}[\varepsilon | Y_t] = \mathbb{E}\left[Y_{t+1} - \mathbb{E}[Y_{t+1}|Y_t] \,\Big|\, Y_t\right]$$

$$= \mathbb{E}[Y_{t+1}|Y_t] - \mathbb{E}\left[\mathbb{E}[Y_{t+1}|Y_t] \mid Y_t\right]$$

$$= \mathbb{E}[Y_{t+1}|Y_t] - \mathbb{E}[Y_{t+1}|Y_t]$$

$$= 0$$

∴

$$MSFE = \mathbb{E}\left[\left(\varepsilon - G(Y_t)\right)^2\right] = \mathbb{E}[\varepsilon^2] + \mathbb{E}\left[G(Y_t)^2\right]$$

minimised when $G(Y_t) = 0$

$\Rightarrow \quad m(Y_t) = \mathbb{E}[Y_{t+1}|Y_t]$

∴

$$\boxed{MSFE = \mathbb{E}\left[\left(Y_{t+1} - \mathbb{E}[Y_{t+1}|Y_t]\right)^2\right]}$$

(b)

$$\mathbb{E}[Y_{t+1} \mid Y_t] = \alpha_0 + Y_t$$

$$\mathbb{E} \quad Y_{t+1} = \alpha_0 + Y_t + u_{t+1}$$

where $\mathbb{E}[u_{t+1} \mid Y_t] = 0$ and $cov(Y_t, u_{t+1}) = 0$

∴ An example of this would be when Y_{t+1} is a unit root AR(1) with ~~deterministic trend~~

α_0 : trend

unit root $\Rightarrow$ δY_t where $\delta = 1$

(+ drift)

3

(a)

(i)

$$\hat{u}_i = W_i - \hat{\beta}_0 - \hat{\beta}_X X_i - \hat{\beta}_C C_i - \hat{\beta}_T T_i$$

this is found by:

$\underset{\beta_0, \beta_X, \beta_C}{\arg\min}$ $\underset{\beta_0, \beta_X, \beta_C, \beta_T}{\arg\min}$

NO EXPECTATIONS IN THE SAMPLE!

(3)

(a)

(i)

use $\sum$

$$R_i = 1 - T_i - C_i \quad \text{or} \quad T_i = 1 - C_i - R_i$$

OLS:

$$\underset{\substack{\beta_0, \beta_X, \\ \beta_C, \beta_T}}{\operatorname{argmin}} \sum_{i=1}^{n} (\cancel{\beta_0 +} W_i - \beta_0 - \beta_X X_i - \beta_C C_i - \beta_T T_i)^2$$

(relevant) focs:

$$0 = -2 \mathbb{E}\left[T_i \left(W_i - \hat{\beta}_0 - \hat{\beta}_X X_i - \hat{\beta}_C C_i - \hat{\beta}_T T_i\right)\right]$$

$$0 = \mathbb{E}[T_i \hat{u}_i]$$

$$0 = -2 \mathbb{E}\left[C_i \left(W_i - \hat{\beta}_0 - \hat{\beta}_X X_i - \hat{\beta}_C C_i - \hat{\beta}_T T_i\right)\right]$$

$$0 = \mathbb{E}[C_i \hat{u}_i]$$

$$\left(\hat{u}_i = W_i - \hat{\beta}_0 - \hat{\beta}_X X_i - \hat{\beta}_C C_i - \hat{\beta}_T T_i\right)$$

$$\widehat{Cov}(R_i, \hat{u}_i) = \mathbb{E}[R_i \hat{u}_i] - \mathbb{E}[R_i]\mathbb{E}[\hat{u}_i]$$

Sample covariance!!

$$= \mathbb{E}[(1 - C_i - T_i)\hat{u}_i] - \mathbb{E}[1 - C_i - T_i]\mathbb{E}[\hat{u}_i]$$

$$= \cancel{\mathbb{E}[\hat{u}_i]} - \mathbb{E}[C_i \hat{u}_i] - \mathbb{E}[T_i \hat{u}_i] - \cancel{\mathbb{E}[\hat{u}_i]} + \mathbb{E}[C_i]\mathbb{E}[\hat{u}_i] + \mathbb{E}[T_i]\mathbb{E}[\hat{u}_i]$$

$\mathbb{E}[C_i \hat{u}_i] = 0$ & $\mathbb{E}[T_i \hat{u}_i] = 0$ by foc's

$$\widehat{Cov}(R_i, \hat{u}_i) = \mathbb{E}[C_i]\mathbb{E}[\hat{u}_i] + \mathbb{E}[T_i]\mathbb{E}[\hat{u}_i]$$

other foc:

$$0 = -2 \mathbb{E}\left[\left(W_i - \hat{\beta}_0 - \hat{\beta}_X X_i - \hat{\beta}_C C_i - \hat{\beta}_T T_i\right)\right] = \mathbb{E}[\hat{u}_i]$$

$\therefore$

$$\Rightarrow \widehat{Cov}(R_i, \hat{u}_i) = 0$$

(ii)

By OLS:

① $W_i = \hat{\beta}_0 + \hat{\beta}_X X_i + \hat{\beta}_C C_i + \hat{\beta}_T T_i + \hat{u}_i$

② $W_i = \cancel{\beta}\hat{\gamma}_0 + \hat{\gamma}_X X_i + \hat{\gamma}_C C_i + \hat{\gamma}_R R_i + \cancel{\hat{u}_i}\hat{v}_i$

$R_i = 1 - T_i - C_i$

∴ ② : $W_i = \hat{\gamma}_0 + \hat{\gamma}_X X_i + \hat{\gamma}_C C_i + \hat{\gamma}_R - \hat{\gamma}_R T_i - \hat{\gamma}_R C_i + \hat{u}_i$

$W_i = (\hat{\gamma}_0 + \hat{\gamma}_R) + \hat{\gamma}_X X_i + (\hat{\gamma}_C - \hat{\gamma}_R) C_i - \hat{\gamma}_R T_i + \cancel{\hat{u}_i}\hat{v}_i$

② $W_i = (\hat{\gamma}_0 + \hat{\gamma}_R) + \hat{\gamma}_X X_i + (\hat{\gamma}_C - \hat{\gamma}_R) C_i - \hat{\gamma}_R T_i + \hat{v}_i$

① $W_i = \hat{\beta}_0 + \hat{\beta}_X X_i + \hat{\beta}_C C_i + \hat{\beta}_T T_i + \hat{u}_i$

let $\hat{\beta}_0 + \hat{\beta}_X X_i + \hat{\beta}_C C_i + \hat{\beta}_T T_i = \hat{\gamma}_0 + \hat{\gamma}_X X_i + \gamma_C C_i + \hat{\gamma}_R R_i$

let $\hat{\beta}_X = \hat{\gamma}_X$

then let $C_i = 1$ $\hat{\beta}_0 + \hat{\beta}_C = \hat{\gamma}_0 + \hat{\gamma}_C$.

let $R_i = 1$ $\hat{\beta}_0 = \gamma_0 + \hat{\gamma}_R$

let $T_i = 1$

$\hat{\beta}_0 + \hat{\beta}_T = \hat{\gamma}_0$

$$\boxed{\begin{aligned} \hat{\beta}_0 &= (\hat{\gamma}_0 + \hat{\gamma}_R) \\ \hat{\beta}_X &= \hat{\gamma}_X \\ \hat{\beta}_C &= (\hat{\gamma}_C - \hat{\gamma}_R) \\ \hat{\beta}_T &= -\hat{\gamma}_R \end{aligned}}$$

Omitted dummy = reference group!!

∴ (1) $\hat{\beta}_C$ = diff. between city & rural group.

(b) X

using:

$W_i = \hat{\beta}_0 + \beta_X X_i + \beta_C C_i + \beta_T T_i + u_i$

$W_i = \beta_0 + \beta_X X_i + \beta_C C_i + \beta_T T_i + \gamma_{XC} C_i X_i + \gamma_{XT} T_i X_i + u_i$

F-test:

$H_0 : \beta_C = \beta_T = 0$ $\cancel{\beta_C - \beta_T = 0}$ $H_0 = \gamma_{XC} = \gamma_{XT} = 0$

$H_1 : \cancel{\beta_C \neq \beta_T \neq 0}$ $\beta_I \neq 0$ $\exists I \in \{C, T\}$ ∴ $k = 6$, $q = 2$.

return to experience

note that $\beta_T = -\gamma_R$ ∴ if $\beta_T = 0$ then $\gamma_R = 0$

Unrestricted model : $W_i = \hat{\beta}_{0_{UN}} + \hat{\beta}_{X_{UN}} X_i + \hat{\beta}_{C_{UN}} C_i + \hat{\beta}_{T_{UN}} T_i + \hat{u}_{i,UN}$

Restricted model : $W_i = \hat{\beta}_{0,RS} + \hat{\beta}_{X,RS} + \hat{u}_{i,RS}$

$$SSR_{UN} = \sum_{i=1}^{n} \hat{u}_{i,UN}^2 \qquad SSR_{RS} = \sum_{i=1}^{n} \hat{u}_{i,RS}$$

$$F = \frac{SSR_{RS} - SSR_{UN}}{SSR_{UN}} \frac{n-K-1}{q}$$

$q =$ restrictions $= 2$

$K = \underset{(\text{in UN})}{\text{regressors}} = 3$

$$F = \frac{(SSR_{RS} - SSR_{UN})(n-1)}{2\, SSR_{UN}} \xrightarrow{D} F_{2,\infty}$$

reject H_0 if $F > CV_\alpha$

④

(a)

$\hat{\mu}_t = 3,800$ $\quad$ $\hat{sd} = 750 = \hat{\sigma}_t$ $\quad$ $n_t = \cancel{}\ 100$

$\hat{\mu}_{ut} = 3200$ $\quad$ $\hat{sd} : 750 = \hat{\sigma}_{ut}$ $\quad$ $n_{ut} = \cancel{}\ 900$

$H_0 : \mu_t = \mu_{ut} \iff \mu_t - \mu_{ut} = 0$

$H_1 : \mu_t \neq \mu_{ut} \iff \mu_t - \mu_{ut} \neq 0$

$$t_n = \frac{\hat{\mu}_t - \hat{\mu}_{ut} - 0}{\sqrt{\frac{\hat{\sigma}_t^2}{n_t} + \frac{\hat{\sigma}_{ut}^2}{n_{ut}}}} \overset{a}{\sim} N(0,1) \text{ under } H_0$$

$$t_n = \frac{3800 - 3200}{\sqrt{\frac{750^2}{100} + \frac{750^2}{900}}} = 7.589$$

reject @ $\alpha = 0.05$ if $|t| > cv_\alpha = 1.96$

$7.589 > 1.96$

$\therefore$ reject H_0

there is sufficient evidence to suggest that mean earnings of those who participated are different to those who did not.

Assumption : the samples are ~~iden~~ independent
- This likely does not hold since 100 workers opted to undertake the programme + 900 didn't
- It may be the case that the 100

who opted to participate were higher performing than those who did not, and undertook training to better themselves.

(b)

Endogeneity:
- Other causes of wage (e.g. work ethic, or competence) are likely correlated with whether or not a worker choose to participate in the ~~study~~. programme.
- More driven workers are more likely to participate

∴ Endogeneity
∴ OR doesn't hold since participation is correlated with other factors that determine wages.

(c) RCT:
- Randomly assign treatment to participants ∴ treatment is independent of other factors by design

(d)

$\hat{\mu}_t = 3625$ & $\hat{\sigma}_t = 750$ $n_T = 50$

$\hat{\mu}_{ut} = 3400$ $\hat{\sigma}_{ut} = 750$ $n_{ut} = 150$

same nulls and alternatives as before, also same decision rule

$$t = \frac{3625 - 3400 - 0}{\sqrt{\frac{750^2}{50} + \frac{750^2}{150}}} = 1.837$$

$1.837 < 1.96$

$\therefore$ not sufficient evidence to reject H_0 at 5% significance level.

(e)

(?) $\hat{\beta} = \hat{\mathbb{E}}[Y|D=1] - \hat{\mathbb{E}}[Y|D=0]$

$\hat{\beta} = 3625 - 3400 = 225$

but no causal interpretation.

(f)

$$t = \frac{3625 - 3400}{\sqrt{\frac{750^2}{80} + \frac{750^2}{120}}} = 2.07$$

$2.07 > 1.96$

$\therefore$ sufficient evidence to reject H_0.

(g)

$$\max t = \frac{3625 - 400}{\sqrt{\frac{750^2}{x} + \frac{750^2}{y}}}$$ where $x + y = A$

implies

$$\min \sqrt{\frac{750^2}{x} + \frac{750^2}{A-x}}$$

$$\min \sqrt{\frac{750^2}{x} + \frac{750^2}{A-x}}$$

$$\cancel{\frac{\partial}{\partial x}\sqrt{\frac{750^2}{x}+\frac{750^2}{A-x}} = \frac{\partial}{\partial x}\left(750^2 x^{-1} + 750^2 (A-x)^{-1}\right)^{1/2}}$$

$f(x) = x^2$ monotonic transformation for $x>0$

of course s.d. >0 always!! $\therefore x>0$!

$\therefore$ same minimum

$$\min \quad 750^2 x^{-1} + 750^2 (A-x)^{-1}$$

$$\frac{\partial}{\partial x} = -750^2 x^{-2} - (-1)\, 750^2 (A-x)^{-2}$$

$$0 = \frac{750^2}{(A-x)^2} - \frac{750^2}{x^2}$$

$$750^2 x^2 = 750^2 (A-x^2)$$

$$7\cancel{5}0\, x = 7\cancel{5}0\, (A-x)$$

$$x = A - x$$

$$\boxed{x = y} \quad \text{or} \quad \boxed{x = \frac{A}{2} = y}$$

as required

SOC:

$$\frac{\partial^2}{\partial x^2} = 2\cdot 750^2\, x^{-3} + (-1)(-2)\cdot 750^2 (A-x)^{-3}$$

$$= 2\cdot 750^2\left[\frac{1}{x^3} + \frac{1}{(A-x)^3}\right]$$

$$>0 \quad \forall\, A \geq x \geq 0$$

$\therefore$ local minimum.

(h)

predict £225 ↑ in salaries

- note RTC was conducted on new employees ∴ effect may be different for existing ones
- other companies work may not be that similar

(5)

(a)

$$V = \beta_2 X^2 + u$$

$$\mathbb{E}[V|X] = \mathbb{E}[\beta_2 X^2 + u | X]$$

$$= \beta_2 \mathbb{E}[X^2|X] + \underbrace{\mathbb{E}[u|X]}_{=0}$$

$$= \beta_2 X^2$$

$$\mathbb{E}[V|X] = 0 \quad \text{if} \quad \beta_2 = 0$$

better:

$$V = Y - \gamma_0 - \gamma_1 X$$

notice $\gamma_0 = \mathbb{E}Y - \gamma_1 \mathbb{E}[X]$

$$\mathbb{E}[V = Y - (\mathbb{E}(Y) - \gamma_1 \mathbb{E}[X]) - \gamma_1 X$$

$$= Y - \mathbb{E}[Y] - \frac{cov(X,Y)}{var(X)}(X - \mathbb{E}[X])$$

sub in for Y.

(b)

- $\hat{\beta}_0$'s of both regressions have relatively high s.e., high enough that in both cases we could not reject that ~~$\beta_0 = 0$~~ $\beta_0 = 0$

regr. Alice:

$$t = \frac{-0.01 - 0}{0.02} = -0.5$$

Bob:

$$t = \frac{0.03 - 0}{0.02} = 1.5$$

at 5% sig. reject $H_0: \beta_0 = 0$ if $|t| > 1.96$

not the case $\therefore$ assume $\beta_0 = 0$

- $\hat{\gamma}_0$ for Bob also has relatively high s.e. such that could not reject $H_0: \gamma_0 = 0$

Bob:

$$t = \frac{-0.08}{0.21} = -0.38 \quad \text{reject if } |t| > 1.96$$

- rule of thumb: estimates $= (2) \cdot (\text{s.e.})$ and all other estimates meet this $\therefore$ significant.

Hence:

Alice: $Y = 0 + 4.99X - 1.98X^2 + u$

$$\boxed{\begin{array}{ll} \text{3A} & Y = 4.99X - 1.98X^2 + u \\ \text{4A} & Y = -1.86 + 4.81X + v \end{array}}$$

Bob:

$$\boxed{\begin{array}{ll} \text{3B} & Y = 5X - 2.02X^2 + u \\ \text{4B} & Y = X + v. \end{array}}$$

how could differences have occured?

- for Alice we can assume average X was approx. 1

$$\mathbb{E}[Y] = 4.99 \cdot \mathbb{E}[X] - 1.98\,\mathbb{E}[X^2] + \underbrace{\mathbb{E}[u]}_{=\mathbb{E}[\mathbb{E}[u|X]] = 0}$$

$$\mathbb{E}[Y] = -1.86 + 4.81\,\mathbb{E}[X] + 0$$

$\therefore$ if $\mathbb{E}[X] \simeq 1$ then $\mathbb{E}[Y] = \begin{cases} 3.01 & \text{in 3A} \\ 2.95 & \text{in 4A} \end{cases}$

$\mathbb{E}[Y] \simeq 3$

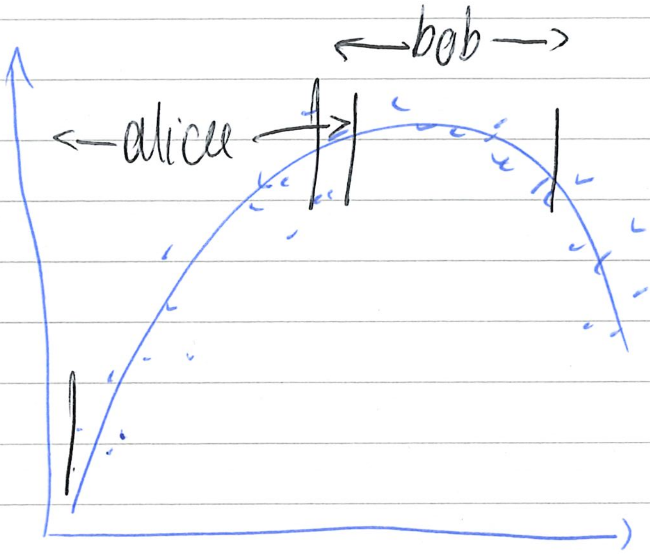
- for Bob we can assume $\mathbb{E}[X] \simeq 2$

then $\mathbb{E}[Y] = \begin{cases} 10 - 2.02 \cdot 4 \simeq 2 & \text{in 3B} \\ \simeq 2 & \text{in 4B} \end{cases}$

$\therefore$ depended on size of $\mathbb{E}[X]$.

Kevin Notes

(b)



using diff subsets of the data!!

for $\hat{\gamma}_0$

- standard errors large for B since he has data further from zero!

OR

(b)

$$\hat{\gamma}_1 = \underbrace{\hat{\beta}_0}_{=0} + \hat{\beta}_2 \frac{cov(X, X^2)}{var(X)}$$

$$\cancel{\hat{\gamma}_1 = \hat{\beta}_2 \frac{cov(X,X^2)}{var(X)}} \quad \hat{\gamma}_1 =$$

∴ ~~Alice must have had~~

$\cancel{4.81 = -1.98}$

Alice:

$$4.81 = 4.99 - 1.98 \frac{cov(X, X^2)}{var(X)}$$

$\therefore \frac{cov(X,X^2)}{var(X)}$ must have been small

$\Rightarrow$ high $var(X)$ & low $cov(X, X^2)$

Bob:

$$\cancel{5a}$$

$$1 = 5 - 2.02 \frac{cov(X,X^2)}{var(X)}$$

∴ must have been larger than A

$\Rightarrow$ lower $var(X)$ and higher $cov(X,X^2)$.

(c)

③ $$\left.\frac{\partial Y}{\partial X}\right|_{X=x} = \beta_1 + 2\beta_2 x$$

$$\left.\frac{\partial Y}{\partial X}\right|_{X=x} = \gamma_1$$

$$\gamma_1 = \frac{cov(Y, X)}{var(X)} = \frac{cov(\beta_0 + \beta_1 X + \beta_2 X^2 + u\ ,\ X)}{var(X)}$$

$$= \frac{\underbrace{cov(\beta_0, X)}_{\substack{=0 \\ (\text{constant}}} + \underbrace{\beta_1 cov(X, X)}_{=\beta_1 var(X)} + \beta_2 cov(X^2, X) + \underbrace{cov(u, X)}_{\substack{=0 \\ (\text{OR})}}}{var(X)}$$

$$= \beta_1 \frac{var(X)}{var(X)} + \beta_2 \frac{cov(X^2, X)}{var(X)}$$

hence

$$\gamma_1 = \cancel{\beta_1} + 2\cancel{\beta_2} x = \cancel{\beta_1} + \cancel{\beta_2} \frac{cov(X^2, X)}{var(X)}$$

$$\boxed{x = \frac{cov(X^2, X)}{2\, var(X)}}$$

(d)

(i) Problematic ~~st~~ as showed by Alice & Bob, model will only fit data well for similar samples of Y and X and cannot be used causally

(ii)

danger ~~over~~ of over fitting...

Best to fit ~~fith~~ fifth order - polynomial and then sequential t test that $H_0: \beta_i = 0$ ~~until~~ for $i=5$, then $i=4$, etc until insufficient evidence to reject H_0.

(e)

error = linear in errors.
hence ε^2 goes into constant!

$\cancel{Y = \beta_0 + \beta_1 Z + \beta_2 Z^2}$

$$Y = \beta_0 + \beta_1(\pi_0 + \pi_1 Z + \varepsilon) + \beta_2(\pi_0 + \pi_1 Z + \varepsilon)^2 + u$$

$$Y = (\beta_0 + \beta_1\pi_0) + \beta_1\pi_1 Z + \beta_2(\pi_0^2 + \pi_0\pi_1 Z + \pi_0\varepsilon + \pi_1^2 Z^2 + \pi_0\pi_1 Z + \pi_1 Z\varepsilon + \varepsilon^2 + \pi_0\varepsilon + \pi_1\varepsilon Z) + u + \beta_1\varepsilon$$

error! → ← constant!

$$Y = (\beta_0 + \beta_1\pi_0 + \pi_0^2) + (\beta_1\pi_1 + 2\pi_0\pi_1 + 2\pi_1\varepsilon)Z + \pi_1^2 Z^2 + u + \beta_1\varepsilon + \varepsilon^2 + \pi_0\varepsilon$$

$$\mathbb{E}[Y|Z] = \gamma_0 + \gamma_1 Z + \gamma_0 Z^2 + \mathbb{E}[u + \beta_1\varepsilon + \varepsilon^2 + \pi_0\varepsilon \,|\, Z]$$

$$= \underbrace{\mathbb{E}[u|Z]}_{=0} + \beta_1\underbrace{\mathbb{E}[\varepsilon|Z]}_{=0} + \pi_0\underbrace{\mathbb{E}[\varepsilon|Z]}_{=0} + \mathbb{E}[\varepsilon^2|Z]$$

$$\mathbb{E}[\varepsilon^2|Z]$$

$$\mathbb{E}[X|Z] = \pi_0 + \pi_1 Z$$

?

Q3, Q4, Q5

Quantitative Economics 2019

(1)

(a)

$$Var(X) = \cancel{\mathbb{E}[X - X^2]}\ \mathbb{E}[(X-\mathbb{E}(X))^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$Var(a+bX+cY) = \mathbb{E}[(a+bX+cY)^2] - (\mathbb{E}[a+bX+cY])^2$$

$$= \mathbb{E}[a^2+b^2X^2+c^2Y^2 + 2abX + 2acY + 2bcYX] - (a+b\mathbb{E}X + c\mathbb{E}Y)^2$$

$(a+bX+cY)(a+bX+cY)$

$= a^2 + abX + acY$

$+ abX + b^2X^2 + bcYX$

$+ acY + bcYX + c^2Y^2$

$= a^2 + 2abX + 2acY + 2bcYX + b^2X^2 + c^2Y^2$

$$= \cancel{a^2} + b^2\mathbb{E}[X^2] + c^2\mathbb{E}[Y^2] + \cancel{2ab\mathbb{E}[X]} + \cancel{2ac\mathbb{E}[Y]} + 2bc\,\mathbb{E}[XY]$$
$$- \cancel{a^2} - \cancel{2ab\mathbb{E}X} - \cancel{2ac\mathbb{E}Y} - 2bc\mathbb{E}X\mathbb{E}Y - b^2(\mathbb{E}[X])^2 - c^2(\mathbb{E}[Y])^2$$

$$= b^2[\mathbb{E}[X^2] - (\mathbb{E}[X])^2] + c^2[\mathbb{E}[Y^2] - (\mathbb{E}[Y])^2]$$
$$+ 2bc[\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]]$$

$$= b^2 \, var(X) + c^2 \, var(Y) + 2bc \, cov(X,Y)$$

$$cov(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

~~(b)~~

~~$X \sim N(1,7)$ $Y \sim N(4,8)$~~

~~$Z = 2X + 3Y$ (sum of normals is normal hence $Z \sim N$)~~

~~$2X \sim N(1,14)$ $3Y \sim N(4,$~~

~~$2X \sim N(1, 7\cdot 2^2)$ $3Y \sim N(4, 8\cdot 3^2)$~~

~~$2X \sim N(1, 28)$ $3Y \sim N(4, 72)$~~

(b)

$E[X] = 1 \quad var(X) = 7 \quad \therefore \quad E[2X] = 2 \quad var(2X) = 2^2 \cdot 7 = 28$

$E[Y] = 4 \quad var(Y) = 8 \quad \therefore \quad E[3Y] = 12 \quad var(3Y) = \cancel{8} \, 72$

Sum of normal rvs is normal hence

$Z \sim N(2 + 12, 28 + 72)$ given X and Y are independent

$Z \sim N(14, 100)$

[If they were dependent then $var(Z) = var(X) + var(Y) + 2cov(X, Y)$]

(c)

$P(Z \geq 25)? = \quad Q = \frac{Z - 14}{100} \sim N(0,1)$

$\frac{25 - 14}{100} = 0.11$

$P(Q \geq 0.11) = 1 - \phi(0.11) = \underline{0.456}$

(2)

$\{X_t\}$ granger causes $\{Y_t\}$ if

$$\mathbb{E}[(Y_{T+1} - \mathbb{E}[Y_{T+1} | Y_T, X_T])^2] < \mathbb{E}[(Y_{T+1} - \mathbb{E}[Y_{T+1} | Y_T])^2]$$

MSFE of Y_{T+1} including X_T < MSFE of Y_{T+1} without X_T

Test : F-test

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{i=1}^{r} \gamma_i X_{t-i} + u_i$$

- $p = r$ other-wise extra predicatibility could just come from different no. of lags

- Test :

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_p = 0$$

Estimate :

UN $\quad Y_t = \hat{\alpha}_{0,un} + \sum_{i=1}^{p} \hat{\beta}_{i,un} Y_{t-i} + \sum_{i=1}^{p} \hat{\gamma}_{i,un} X_{t-1} + \hat{u}_{i,un}$

RS $\quad Y_t = \hat{\alpha}_{0,RS} + \sum_{i=1}^{p} \hat{\beta}_{i,RS} Y_{t-i} + \hat{u}_{i,RS}$

$$F = \frac{SSR_{RS} - SSR_{UN}}{SSR_{UN}} \; \frac{n - 2p + 1}{p} \xrightarrow{d} F_{p,\infty}$$

reject if $F_{p,\infty} > CV_{F,\alpha}$

(3)

(a)

On average log(wages) are 0.15 lower in wales

$$\text{coefficient} = \mathbb{E}[\log(\text{wages}) \mid \text{wales}=1] - \mathbb{E}[\log(\text{wages}) \mid \text{wales}=0]$$

On average wages are ~~$(e^{-0.15} - 1) = -0.139) \times 100 = 13$~~
13.9% lower in wales

$$(e^{-0.15} - 1) \times 100 = -13.9\%$$

(b)

2 additional yrs. increases log(wages) by 0.1
or
increases wages by 10.5%

(c)

~~90% CI : 0.1~~

~~recall $\hat{\beta}$ is an estimator (rv) hence $2\hat{\beta}$ has $\mathbb{E}[2\hat{\beta}] = 2 \cdot \mathbb{E}[\hat{\beta}]$~~
~~and $var(2\hat{\beta}) = 4\, var(\hat{\beta})$~~
~~hence $se(\hat{\beta}) = 0.02$ $\quad s.e.(2\hat{\beta}) = 2^2 \cdot 0.02 = 0.08$~~

~~90% CI : $0.1 \pm$~~

recall $\hat{\beta}$ is an estimator (hence a RV) hence
for $2\hat{\beta}$ $\quad \mathbb{E}[2\hat{\beta}] = 2\mathbb{E}[\hat{\beta}]$
$var(2\hat{\beta}) = 4\, var(\hat{\beta}) \implies s.e(2\hat{\beta}) = \sqrt{4\, v\hat{a}r(\hat{\beta})}$
$= 2\, s.e.(\hat{\beta})$

90% CI : $0.1 \pm 1.645 \cdot 2 \cdot 0.02$
~~0.1 ±~~
$= [0.0342, 0.1658]$

(d)

$H_0: \beta_{Gender} = 0$

$$t = \frac{0.08 - 0}{0.03} = 2.667 \overset{a}{\sim} N(0,1) \quad \text{under } H_0$$

$P_1 = 2\phi(2.667) = 0.0077$

'what is the probability under H_0 of finding evidence against the null beyond the observed t-stat'

(e)

?

- No change to Gender estimator from removal of region dummies since likely that $cov(Gender, Region) = 0$ (gender is split approx 50:50 across all regions)
- experience and gender may be correlated as women leave work force in child birth hence more senior workers may be men.
- Result is $\hat{\beta}_{Gender}$ will be larger since it is picking up some of the effect of experience on log(wages)
- $s.e(\hat{\beta}_{Gender})$ will increase since model will likely fit data less well

$\therefore$ higher variance $\Rightarrow$ higher s.e.

(?) This seems ridiculous...

(4)

(a)

By iid CLT (Lindeberg - Levy):

$$\frac{\sqrt{n}\,(\bar{Y}_n - \mu_Y)}{\sigma_Y} \xrightarrow{d} N(0,1)$$

for n sufficiently large

- Rule of thumb is $n \geq 30$ hence $1000 = n$ is sufficient

$$\bar{Y}_n \sim N\left(\mu_Y, \frac{\sigma_Y^2}{n}\right)$$

(b)

95% CI : $55 \pm 1.96 \cdot \frac{10}{\sqrt{400}}$

$$= [54.02, 55.98]$$

(c)

$H_0: \mu_Y^{ox} = 50$

$H_1: \mu_Y^{ox} > 50$

$$t = \frac{55 - 50}{10/\sqrt{400}} = 10 \sim N(0,1) \text{ under } H_0$$

(one-tailed)

$$P(t) = 1 - \phi(10) \approx 1 - 1 = 0$$

(d)

$H_0: \quad \mu^{ox} - \mu_T^{ox} = 0$

$H_1: \quad \mu^{ox} \neq \mu_T^{ox} \iff \mu^{ox} - \mu_T^{ox} \neq 0$

$$t = \frac{55 - 57 - 0}{\sqrt{\frac{20^2}{900} + \frac{10^2}{400}}} = -2.4$$

reject H_0 if $\quad t < CV_{0.01}$

$CV_{0.01} = -2.58$

$\therefore$ not sufficient evidence to show treated mean was substantially different from untreated.

2 sided test

$$\mathbb{E}[\beta_{1i} | X_i] = \beta_1$$

(by mean independence).

(5)

(a)

$$\arg\min \{ \mathbb{E}[(Y_i - \beta_{1i} X_i - \beta_0 \cancel{u_i})^2]$$

foc:

$$\mathbb{E}[Y_i - \beta_{1i} X_i - \beta_0] \cancel{u_i} = 0$$

$$\beta_0 = \mathbb{E}[Y_i] - \mathbb{E}[\beta_{1i}]\mathbb{E}[X_i]$$

(by mean independence)

$$\mathbb{E}[(Y_i - \beta_{1i} X_i - \beta_0) X_i] = 0$$

$$\mathbb{E}[Y_i X_i - \beta_{1i} X_i^2 - \mathbb{E}[Y_i] X_i + \mathbb{E}[\beta_{1i}]\mathbb{E}[X_i] X_i] = 0$$

$$\mathbb{E}[\beta_{1i}]\left(\mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2\right) = \mathbb{E}[Y_i X_i] - \mathbb{E}[Y_i]\mathbb{E}[X_i]$$

$$\mathbb{E}[\beta_{1i}] = \frac{\operatorname{cov}(Y_i, X_i)}{\operatorname{var}(X_i)}$$

$\therefore$ recoverable by pop. lin. regression of X_i on Y_i

(b)

$\mathbb{E}[\beta_{1i}]$ = average causal effect

The ~~effect~~ expected causal effect of a randomly selected member of the population under study.

(c)

$$\cancel{\beta_{IV} = \frac{cov(Y_i, X_i)}{}}$$

$$\cancel{X = \pi_0 + \pi_{1i} Z_i + v_i \equiv X^* + v_i}$$

$$\cancel{Y_i = \beta_0 + \beta_{1i} X^* + \underbrace{(\beta_{1i} v_i + u_i)}_{=\varepsilon}}$$

$$\cancel{cov(\beta_{1i} v_i + u_i, X^*) = cov(\beta_{1i} v_i + u_i, \delta_0 + \delta_1 Z_i)}$$

$$\cancel{= \beta_{1i}\delta_1 cov(Z_i, \beta_{1i} v_i)} + \delta_1 cov(Z_i, u_i)$$

$=0$ $=0$

$\therefore$ OR holds.

$\boxed{\text{2SLS}}$

~~ILS:~~

$$\cancel{Y_i = \beta_0 + \beta_{1i}(}$$

$$X_i = \pi_0 + \pi_{1i} Z_i + v_i \equiv X^* + v_i$$

or

$$\beta_{IV} = \frac{cov(Y, X^*)}{var(X^*)}$$

$$= \frac{cov(Y, \pi_0 + \pi_1 Z_1)}{cov(X^*, X - v)}$$

$$= \frac{\pi_1 cov(Y, Z_1)}{cov(\pi_0 + \pi_1 Z_1, X - v)}$$

$$= (\pi_1)^2 var(Z_1)$$

$$\pi_1 = \frac{cov(X, Z)}{var(Z)}$$

etc.

$$\pi_{1i} = \frac{cov(X_i, Z_i)}{var(Z_i)}$$

$$Y_i = \beta_0 + \beta_{1i}(\pi_0 + \pi_{1i} Z_i + v_i) + u_i$$

$$= \beta_0 + \beta_{1i}\pi_0 + \beta_{1i}\pi_{1i} Z_i + \beta_{1i} v_i + u_i$$

$$\beta_{1i}\pi_{1i} = \frac{cov(Y_i, Z_i)}{var(Z_i)}$$

$$\beta_{1i} = \frac{cov(Y_i, Z_i) / \cancel{var(Z_i)}}{\pi_{1i}} = \frac{cov(Y_i, Z_i) / \cancel{var(Z_i)}}{cov(X_i, Z_i) / \cancel{var(Z_i)}}$$

$$= \frac{cov(Y_i, Z_i)}{cov(X_i, Z_i)}$$

as required.

go for expectations later on.

(d)

$$\beta_{IV} = \frac{cov(Y_i, Z_i)}{cov(X_i, Z_i)} = \frac{\mathbb{E}[(Z_i - \mu_Z) Y_i]}{\mathbb{E}[(Z_i - \mu_Z) X_i]} =$$

$$= \frac{\mathbb{E}[(Z_i - \mu_Z)(\beta_0 + \beta_{1i}\pi_0 + \beta_{1i}\pi_{1i} Z_i + \beta_{1i} v_i + u_i)]}{\mathbb{E}[(Z_i - \mu_Z)(\pi_0 + \pi_{1i} Z_i + v_i)]}$$

$$= \frac{\beta_0 \mathbb{E}[Z_i - \mu_Z] + \pi_0 \mathbb{E}[(Z_i - \mu_Z)\beta_{1i}] + \mathbb{E}[(Z_i - \mu_Z)\beta_{1i} Z_{1i} \pi_{1i}] + \mathbb{E}[(Z_i - \mu_Z)\beta_{1i} v_i] + \mathbb{E}[(Z_i - \mu_Z) u_i]}{\pi_0 \mathbb{E}[Z_i - \mu_Z] + \mathbb{E}[(Z_i - \mu_Z)\pi_{1i} Z_i] + \mathbb{E}[(Z_i - \mu_Z) v_i]}$$

$\mathbb{E}[Z_i - \mu_Z] = 0 \therefore$

$$\beta_{IV} = \frac{\beta_0 \times 0 + \pi_0 \mathbb{E}[(Z_i - \mu_Z)\beta_{1i}] + \mathbb{E}[(Z_i - \mu_Z)\beta_{1i} Z_{1i} \pi_{1i}] + \mathbb{E}[(Z_i - \mu_Z)\beta_{1i} v_i] + \mathbb{E}[(Z_i - \mu_Z) u_i]}{\pi_0 \times 0 + \mathbb{E}[(Z_i - \mu_Z)\pi_{1i} Z_i] + \mathbb{E}[(Z_i - \mu_Z) v_i]}$$

$$= \frac{\pi_0 \, cov(\underbrace{Z_i, \beta_{1i}}_{=0}) + \overset{\mathbb{E}[(Z_i - \mu_Z) Z_i \beta_{1i} \pi_{1i}]}{\cancel{cov(Z}} + cov(\underbrace{Z_i, \beta_{1i} v_i}_{=0}) + cov(\underbrace{Z_i, u_i}_{=0})}{\mathbb{E}[(Z_i - \mu_Z) Z_i \pi_{1i}] + \cancel{\mathbb{E}[Z} \; cov(\underbrace{Z_i, v_i}_{=0})}$$

(=0 by indep.)

$$= \frac{\mathbb{E}[(Z_i - \mu_Z) Z_i \beta_{1i} \pi_{1i}]}{\mathbb{E}[(Z_i - \mu_Z) Z_i \pi_{1i}]} = \frac{\cancel{var(Z_i)} \, \mathbb{E}[\beta_{1i}\pi_{1i}]}{\cancel{var(Z_i)} \, \mathbb{E}[\pi_{1i}]} \quad \text{(independent)}$$

$$\beta_{IV} = \frac{\mathbb{E}[\beta_{1i}\pi_{1i}]}{\mathbb{E}[\pi_{1i}]}$$

Late ($\mathbb{E}[\beta_{1i}]$ weighted by prob. of accepting treatment)

(e) ?

(f) ① if $\boxed{\cancel{\pi_{1i} = \mathbb{E}[\pi_{1i}]} \quad \pi_{1i} = \pi_1 \;\; \forall i}$

② $cov(\beta_{1i}, \pi_{1i}) = 0$

$\underbrace{cov(\beta_{1i}, \pi_{1i})}_{=0} = \mathbb{E}[\beta_{1i}, \pi_{1i}] - \mathbb{E}[\beta_{1i}]\,\mathbb{E}[\pi_{1i}]$

$\therefore \mathbb{E}[\beta_{1i}, \pi_{1i}] = \mathbb{E}[\beta_{1i}]\,\mathbb{E}[\pi_{1i}]$

$$\beta_{IV} = \frac{cov(Y_i, \check{X}_i)}{var(\check{X}_i)} = \frac{cov(Y_i, \delta_0 + \delta_1 Z_i)}{var(\check{X}_i)}$$

$$= \frac{\delta_1 cov(Y_i, Z_i)}{var(\delta_0 + \delta_1 Z_i)}$$

regressing Y on fitted value of $X = \check{X}$.

$$= \frac{cov(Y_i, Z_i)}{\delta_1 var(Z_i)}$$

$\check{X} = X^*$

...

$\underline{X = X^* + v_i}$

$$= \frac{cov(Y_i, Z_i)}{cov(X_i, Z_i)}$$

$$\boxed{\begin{aligned} X &= \pi_0 + \pi_1 Z_i + v_i \\ X^* &= \pi_0 + \pi_1 Z_i \end{aligned}}$$

(8)

(a) $y_t = \beta(\gamma y_{t-1} + v_t) + u_t$

$= \underbrace{\beta\gamma}_{=1} y_{t-1} + \underbrace{\beta v_t + u_t}_{= \varepsilon_t}$

$y_t = y_{t-1} + \varepsilon_t$ where $\varepsilon_t = \beta v_t + u_t$ and is iid since v_t and u_t are iid.

(b)

① $\mathbb{E}[y_t] = \mathbb{E}[y_{t-1}] + \underbrace{\mathbb{E}[\varepsilon_t]}_{=0}$

$y_t = (y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$

$= y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$

$y_t = y_{t-h} + \sum_{i=0}^{h-1} \varepsilon_{t-i}$

$\mathbb{E}[y_t] = \mathbb{E}[y_{t-h}] + \sum_{i=1}^{h-1} \underbrace{\mathbb{E}[\varepsilon_{t-i}]}_{=0}$

$\therefore \mathbb{E}[y_t] = \mathbb{E}[y_{t-h}] \quad \forall h$

② $var(y_t) = var(y_{t-h}) + \sum_{i=0}^{h-1} var(\varepsilon_{t-i})$

$cov(y_{t-h}, \sum_{i=1}^{h-1} \varepsilon_{t-i}) = 0$

since ε_{t-h+1} are after y_{t-h} $\therefore$ no correlation.

$var(y_t) = var(y_{t-h}) + h\sigma_\varepsilon^2$

let $h = t$

$var(y_t) = var(y_0) + t\sigma_\varepsilon^2$

depends on t.

$\therefore$ non-stationary

(C)

QE 2018

Part (A)

(1)

(a)

An estimator is unbiased iff its expected value is the thing it is an estimator of.
(estimator is random variable estimating a pop. parameter)

Eg: $\mathbb{E}[\hat{\theta}] = \theta \quad \therefore \hat{\theta}$ is unbiased for θ.

E.g. $y_i = \beta + u_i \qquad u_i \overset{iid}{\sim} (0, \sigma_u^2)$

$$\hat{\beta} = \arg\min \sum_{i=1}^{n} (y_i - \beta)^2$$

$$-2 \sum_{i=1}^{n} (y_i - \hat{\beta}) = 0$$

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} y_i = \beta + \frac{1}{n} \sum_{i=1}^{n} u_i$$

$$\mathbb{E}[\hat{\beta}] = \mathbb{E}\left[\beta + \frac{1}{n} \sum u_i\right]$$

$$= \underbrace{\mathbb{E}[\beta]}_{=\beta} + \frac{1}{n} \sum_{i=1}^{n} \underbrace{\mathbb{E}[u_i]}_{=0}$$

$$\underline{\underline{\mathbb{E}[\hat{\beta}] = \beta}}$$

(b)

Estimator is more efficient if it has a lower variance than another ~~varian~~ estimator.

E.g. Same example but ① : $u_i \sim iid(0, \sigma_u^2)$ ② : ~~$u_i \sim iid(0, \frac{n\sigma_u^2}{n})$~~ $u_i \sim iid(0, n\sigma_u^2)$

①
$$Var(\hat{\beta}_1) = \frac{1}{n^2} \sum_{i=1}^{n} var(u_i) = \frac{\sigma_u^2}{n}$$

($cov(u_i, u_j) = 0 \quad \forall i \neq j$
by iid)

②

$$var(\hat{\beta}_2) = \frac{1}{n^2}\sum_{i=1}^{n} var(\cancel{n} u_i) = \frac{1}{n^2} n \cdot n \sigma_u^2 = \sigma_u^2$$

$\sigma_u^2 > \frac{\sigma_u^2}{n}$ $\therefore$ $\hat{\beta}_1$ is more efficient than $\hat{\beta}_2$

(c)

Consistency: $\hat{\beta}$ is consistent iff $\hat{\beta} \xrightarrow{P} \beta$

Same example:

$$\hat{\beta} = \beta + \frac{1}{n}\sum_{i=1}^{n} u_i$$

$n \to \infty$ ~~by iid~~

$\cancel{\hat{\beta} \xrightarrow{P} \hat{\beta}_0}$

$\beta \to \beta$

$\frac{1}{n}\sum_{i=1}^{n} u_i \xrightarrow{P} 0$ (iid LLn)

hence $\hat{\beta} \xrightarrow{P} \beta$ (consistent.)

(2)

Show: $Var(Y) = var(\mathbb{E}[Y|X]) + Var(e)$

where: $Y = \mathbb{E}[Y|\cancel{X}] + e$

$\mathbb{E}[\cancel{X}e|X] = \mathbb{E}[e]$ (by MI)

① $\mathbb{E}[e] = \mathbb{E}[Y - \mathbb{E}[Y|X]] \underset{LIE}{=} \mathbb{E}[Y] - \mathbb{E}[Y] = 0$

hence $\mathbb{E}[e|X] = \mathbb{E}[e] = 0$

(by mean independence)

②

$$Cov(X, e) = \mathbb{E}[(X - \mathbb{E}X)(e - \mathbb{E}e)]$$

$$= \mathbb{E}[Xe + \mathbb{E}[X]\mathbb{E}[e] - e\mathbb{E}[X] - X\mathbb{E}[e]]$$

$$= \underbrace{\mathbb{E}[Xe]}_{=0} - \underbrace{\mathbb{E}[X]\underbrace{\mathbb{E}[e]}_{=0}}_{=0}$$

$$\mathbb{E}[Xe] = \mathbb{E}[\mathbb{E}[Xe|X]] = \mathbb{E}[X\underbrace{\mathbb{E}[e|X]}_{=0}] = 0$$

$\therefore$

$cov(X, e) = 0$

$$Var(Y) = var(\mathbb{E}[Y|X]) + Var(e) + 2Cov(\mathbb{E}[Y|X], e)$$

$\cancel{= 2Cov(Y - e, e)}$

$\cancel{= 2Cov(Y, e) - var(e)}$

$\mathbb{E}[Y|X]$ is a function of X $\therefore$

$cov(\mathbb{E}[Y|X], e) = 0.$

(3)

(a)

Spurious regression = the systemic tendency to find a statistically significant relationship between $\{X_t\}$ and $\{Y_t\}$

E.g. $\{X_t\}$ and $\{Y_t\}$ are independent random walks

$$X_t = \sum_{s=1}^{t} e_{x,s} \qquad Y_t = \sum_{s=1}^{t} e_{y,s}$$

$$\sim I(1) \qquad \sim I(1)$$

(b)

$I(0)$ functions
e.g. difference $\{X_t\}$ and $\{Y_t\}$

Test for cointegration (Engle - Granger test)
if do not reject null then spurious.

Part (B)

(4) Bernoulli $\sim X_i$ $\quad P(X_i=1)=p \ , \ P(X_i=0)=(1-p)$

(a)

density : $f(x) = p^x (1-p)^{(1-x)}$

distribution :
$$F(X) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ (1-p) & \text{if } 0 \le x < 1 \\ p & \text{if } 1 \le x < \infty \end{cases}$$

(b)

$$\mathbb{E}[X] = 0\cdot(1-p) + 1(p) = p$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \underbrace{\mathbb{E}[X]^2}_{=p^2}$$

$$\mathbb{E}[X^2] = 0^2\cdot(p^0)^2(1-p^{(1-0)})^2 + 1^2(p^1)^2(1-p^{(1-1)^2})$$
$$= p$$

$$\mathrm{Var}(X) = p(1-p)$$

$$Sk(X) = \mathbb{E}\left[\left\{\frac{X-\mathbb{E}[X]}{\sqrt{p(1-p)}}\right\}^3\right] = \frac{1}{\{p(1-p)\}^{3/2}}\mathbb{E}[(X-p)^3]$$

$(a-b)(a-b)(a-b)$

$(a^2-2ab+b^2)(a-b)$

$a^3-2a^2b+ab^2-ba^2+2ab^2-b^3$

$$= \frac{1}{\{p(1-p)\}^{3/2}}\mathbb{E}[X^3 - 2X^2p + \cancel{3}p^2X - pX^2 + 2Xp^2 - p^3]$$

$$= \frac{1}{\{p(1-p)\}^{3/2}}\left[\mathbb{E}[X^3] + \mathbb{E}[3p^2X] - \mathbb{E}[3pX^2] - \mathbb{E}[p^3]\right]$$

$$= \frac{1}{[p(1-p)]^{3/2}}\left[p + 3p^3 - 3p^2 - p^3\right]$$

$$Sk(x) = \frac{1}{\{p(1-p)\}^{3/2}}\left[p(2p^2-3p+1)\right]$$

$(p-1)(2p-1)$

$2p^2-3p+1.$

$$\cancel{\frac{1}{\{p(1-p)\}^{3/2}}\left[p(p-1)(2p-1)\right]}$$

$$= \cancel{\frac{(2p-1)}{\sqrt{p(1-p)}}}$$

$$= \frac{1}{p^{3/2}(1-p)^{3/2}}\, p(p-1)(2p-1)$$

$$= \frac{p(1-p)(1-2p)}{p^{3/2}(1-p)^{3/2}} = \frac{1-2p}{p^{1/2}(1-p)^{1/2}}$$

Now $\hat{p} = n^{-1}\sum_{i=1}^{n} X_i$

(c)

$$var(\hat{p}) = \frac{1}{n^2}\sum_{i=1}^{n} var(X_i) \qquad (cov(X_i, X_j) = 0 \text{ by iid assumption})$$

$$= \frac{1}{n^{\cancel{2}}}\cancel{n}\, p(1-p)$$

$$var(\hat{p}) = \frac{p(1-p)}{n} \qquad s.d. = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$$

$$\text{standard error} = \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

$$\boxed{\text{Standard error of estimator} = \text{estimate of standard deviation}}$$

(d)

$$t = \frac{\hat{p} - p}{se(\hat{p})} \xrightarrow{D} N(0,1)$$

$$se(\hat{p}) = \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

$$t = \frac{\sqrt{n}(\hat{p} - p)}{\sqrt{\hat{p}(1-\hat{p})}}$$

$$\sqrt{n}(\hat{p} - p) \xrightarrow{D} N(0, \cancel{\text{var}}\ s.d(\hat{p}))$$

$$\sqrt{\hat{p}(1-\hat{p})} \xrightarrow{p} \sqrt{p(1-p)}$$

$$t \xrightarrow{D} \frac{N(0, \hat{p}(1-\hat{p}))}{\sqrt{p(1-p)}} = N(0,1)$$

(e)

$$95\% \ CI = \hat{p} \pm 1.96 \cdot s.e.(\hat{p})$$

$$= 0.3 \pm 1.96 \frac{\sqrt{0.3 \cdot 0.7}}{\sqrt{100}}$$

$$= 0.3 \pm 1.96 \frac{\sqrt{21}}{100}$$

(6)

(a)

$\hat{\beta}_{woman} = -0.17$ $\qquad$ $\hat{\beta}_{woman}$ estimates $\mathbb{E}[\ln(wage) | woman = 1] - \mathbb{E}[\ln(wage) | woman = 0]$

- All else equal, on average, log(wage) ~~changes~~ is -0.17 lower for women than men.

$$\cancel{\Delta \log(wage) =}$$

$$wage = e^{-0.17} = 0.8436$$

- On average women earn 0.8436, or 15.6% less than men

99% CI : $\hat{\beta}_{woman} \pm 2.58 \cdot 0.03$

$$= [\hat{\beta}_{woman} - 0.0774 , \hat{\beta}_{woman} + 0.0774]$$

$$= [-0.2474 , -0.0926]$$

① CI gives a measure of uncertainty of the point estimate $\hat{\beta}_{woman}$
(Which values of β_{woman} are supported by the data at 99% level)

OR

② Set ~~cont~~ that would contain β_{woman} 99% of time if you repeated experiment 100 times.

(b)

- Age controls for labour market
- Education controls for human capital.

$H_0: \beta_{woman} = -0.17$
$H_1: \beta_{woman} \neq -0.17$

2-tailed at 10% : $C_{0.05} = 1.645$

$$t = \frac{\hat{\beta}_{woman} - (-0.17)}{se(\hat{\beta}_{woman})} \sim N(0,1) \text{ under } H_0$$

$$t = \frac{-0.13 + 0.17}{0.03} = 1.\dot{3}\dot{3}$$

reject H_0 if $|t| > C_{0.05} = 1.645$

$1.\dot{3}\dot{3} < 1.645$

∴ not sufficient evidence to reject H_0 at 10% level.

(c)

Disagree

- endogenous controls
- ~~Occupation dummies = a post - treatment effect ? X~~
-

(Don't really need to know)!

(d)

$H_0: \beta_{woman} = 0$

$H_1: \beta_{woman} = 1$

$$t = \frac{\hat{\beta}_{woman} - 0}{se(\hat{\beta}_{woman})} \sim N(0,1) \text{ under } H_0$$

Decision rule = reject H_0 if $p > \alpha$

$$t = \frac{-0.01}{0.04} = -\frac{1}{4}$$

$$2 \cdot \Phi\left(\frac{1}{4}\right) = 2 \cdot 0.40129 = \underline{\underline{0.8025}}$$

$0.8025 > \alpha$ for $\alpha = 0.05$ $\therefore$ do not reject H_0

p-value = probability of drawing a statistic at least as adverse to H_0 as $\hat{\beta}_{woman}$. under to assumption H_0 is true.

p-value = smallest significance level at which we can reject H_0

(8)

(a)

$$y_1 = y_{-1} + u_1 \quad ; \quad y_2 = y_0 + u_2 \; ; \; y_3 = y_1 + u_3$$

$$\mathbb{E}[y_1] = \mathbb{E}[y_{-1}] + \mathbb{E}[u_1] = 0$$
$$\mathbb{E}[y_2] = \mathbb{E}[y_0] + \mathbb{E}[u_2] = 0$$
$$\mathbb{E}[y_3] = \mathbb{E}[y_1] + \mathbb{E}[u_3] = 0$$

$$\text{var}(y_1) = \underbrace{\text{var}(y_{-1})}_{=0} + \underbrace{\text{var}(u_1)}_{=\sigma^2} + \underbrace{2\,\text{cov}(y_{-1}, u_1)}_{=0}$$
$=0$ since y_s and u_t independent for $s<t$

$$= \sigma^2$$

$$\text{var}(y_2) = \text{var}(y_0) + \text{var}(u_2) + 0$$
$$= \sigma^2$$

$$\text{var}(y_3) = \text{var}(y_1) + \text{var}(u_3) + 0$$
$$= 2\sigma^2$$

$\therefore$ not covariance-stationary as variance depends on t.

(b)

$|\phi| < 1 \qquad y_{-1}, y_0 \sim$ iid independent $(0, \sigma^2/(1-\phi^2))$

(i)

$$\mathbb{E}[y_t] = \phi\,\mathbb{E}[y_{t-2}] + \underbrace{\mathbb{E}[u_t]}_{=0}$$

covariance stationary hence $\mathbb{E}[y_{t-2}] = \mathbb{E}[y_t]$

$$\therefore \mathbb{E}[y_t] = 0$$

$$\text{Var}(y_t) = \phi^2\,\text{var}(y_{t-2}) + \underbrace{\text{var}(u_t)}_{=\sigma^2} + \underbrace{2\,\text{cov}(y_{t-2}, u_t)}_{=0}$$
$=0$ since independent for y_s, u_t $s<t$.

cov-stationary $\Rightarrow \text{var}(y_{t-2}) = \text{var}(y_t)$

$$\text{var}(y_t) = \frac{\sigma^2}{(1-\phi^2)}$$

(ii)

$$\rho_1 = \frac{cov(y_t, y_{t-1})}{var(y_t)}$$

$cov(y_t, y_{t-1})$

$= cov(\phi y_{t-2} + u_t, y_{t-1})$

$= \phi \, cov(y_{t-2}, y_{t-1}) + \underbrace{cov(u_t, y_{t-1})}_{=0 \text{ for } y_s, u_t, \; s<t}$

cov-stationary $\Rightarrow$ $cov(y_t, y_{t-h}) = \gamma_h$

$\therefore$ $cov(y_t, y_{t-1}) = 0$

$\therefore$ $\rho_1 = 0$

(iii)

$$\rho_h = \frac{cov(y_t, y_{t-h})}{var(y_t)}$$

$cov(y_t, y_{t-h})$ for h odd

$= \phi \, cov(y_{t-2}, y_{t-h}) + \underbrace{cov(y_{t-h} + u_t)}_{=0}$

prove by induction.

(iv)

$cov(y_t, y_{t-2}) = \phi \, cov(y_{t-2}, y_{t-2}) = \phi \, var(y_{t-2})$

backward substitution $\Rightarrow$ $cov(y_t, y_{t-h}) = \phi^{h/2} var(y_t)$

[Idea:

$y_t = \phi y_{t-2} + u_t$, $y_{t-2} = \phi y_{t-4} + u_{t-2}$, $y_{t-4} = \phi y_{t-6} + u_{t-4}$

$y_t = \phi^3 y_{t-6} + \phi^2 u_{t-4} + \phi u_{t-2} + u_t$

$$y_t = \phi^h y_{t-2h} + \sum_{s=0}^{h-1} \phi^s u_{t-2s}$$

$cov(y_t, y_{t-h}) = \phi^h cov(y_{t-2h}, y_{t-h})$?]

(C)

$\boxed{\mathbb{E}[y_{t+1} \mid y_t, y_{t-1} \cdots] \neq \phi y_{t-1}}$, $\boxed{\mathbb{E}[y_{t+2} \mid y_t, y_{t-1} \cdots] = \phi y_t}$,

$$\boxed{\begin{aligned} \mathbb{E}[y_{t+3} \mid y_t, \cdots] &= \phi \, \mathbb{E}[y_{t+1} \mid y_t \cdots] \\ &= \phi^2 y_{t-1} \end{aligned}}$$

$$\boxed{\mathbb{E}[y_{t+4} \mid y_t, \cdots] = \phi \, \mathbb{E}[y_{t+2} \mid y_t \cdots] = \phi^2 y_t}$$

(d)

$\mathbb{E}[y_{t+1} \mid y_t] = \phi \, \mathbb{E}[y_{t-1} \mid y_t]$, $\mathbb{E}[y_{t-1} \mid y_t] = \rho, y_t = 0$

$\therefore$ not useful forecast.