Practicum 2 Bookwork | 3.6, 3.11, 3.13, 3.18

Kieran Douglas

3.6 (i) MLR $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$: GM 1–4

I want to estimate $\hat{\theta}_1 = \hat{\beta}_1 + \hat{\beta}_2$ where it is an unbiased estimate of $\theta_1 = \beta_1 + \beta_2$

$\hat{\theta}_1$ is an unbiased estimator of θ_1 because it is entirely composed of estimators that satisfy MLR1 – MLR4, the necessary conditions for unbiasedness.

Proof: Since $E[\hat{\beta}_1] = \beta_1$ and $E[\hat{\beta}_2] = \beta_2$, $E[\hat{\theta}_1] = E[\hat{\beta}_1 + \hat{\beta}_2] = \beta_1 + \beta_2 = \theta_1$

(ii) $\mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}(\hat{\beta}_1) + \mathrm{Var}(\hat{\beta}_2) + 2\,\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)$

↳ Since $\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2) = \mathrm{Corr}(\hat{\beta}_1, \hat{\beta}_2)\sqrt{\mathrm{Var}(\hat{\beta}_1)\mathrm{Var}(\hat{\beta}_2)}$

↳ $\mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}(\hat{\beta}_1) + \mathrm{Var}(\hat{\beta}_2) + 2\,\mathrm{Corr}(\hat{\beta}_1, \hat{\beta}_2)\sqrt{\mathrm{Var}(\hat{\beta}_1)\mathrm{Var}(\hat{\beta}_2)}$

3.11 Population model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$; MLR1 → MLR4

estimated model: $\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 x_2 + u$

$$E(\tilde{\beta}_1) = \beta_1 + \beta_3 \frac{\sum \hat{r}_{i1} x_{i3}}{\sum r_{i1}^2} \longrightarrow \text{from } 3.22 \longrightarrow \hat{\beta}_1 = \frac{\sum \hat{r}_{i1} y_i}{\sum \hat{r}_{i1}^2}$$

(1) ↳ first, $x_{1i} = \pi_0 + \pi_1 x_{2i} + \hat{r}_{1i}$ $\longrightarrow$ $\tilde{\beta}_1 = \dfrac{\sum \hat{r}_1 (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i)}{\sum \hat{r}_{1i}^2}$

(2) ↳ $$\hat{\beta}_1 = \frac{\cancel{\beta_0 \sum \hat{r}_1} + \beta_1 \sum \hat{r}_{1i} x_{1i} + \cancel{\beta_2 \sum \hat{r}_{1i} x_{2i}} + \beta_3 \sum \hat{r}_{1i} x_{3i} + \sum \hat{r}_{1i} u_i}{\sum \hat{r}_{1i}^2}$$

(3) ↳ $$\hat{\beta}_1 = \frac{\beta_1 \sum \hat{r}_{1i} x_{1i} + \beta_3 \sum \hat{r}_{1i} x_{3i} + \sum \hat{r}_{1i} u_i}{\sum \hat{r}_{1i}^2} \longrightarrow \boxed{E(\tilde{\beta}_1) = \beta_1 + \beta_3 \left[\frac{\sum \hat{r}_{1i} x_{3i}}{\sum \hat{r}_{1i}^2}\right]}$$

3.13 (i) $y = \beta_0 + \beta_1 x + u \longrightarrow$ GM 1-4

For $g(x)$, define $z_i = g(x_i)$. Define a slope estimator as $\tilde{\beta}_1 = \dfrac{\Sigma(z_i - \bar{z})y_i}{\Sigma(z_i - \bar{z})x_i}$

Show that $\tilde{\beta}_1$ is linear and unbiased...

Proof: Linearity ~ Numerator $\Sigma_{i=1}^{n}(z_i - \bar{z})y_i$ is a linear combination of y_i values and the denominator is a nonrandom sample of x_i values, both satisfying GM 1-4 for β_1 and thus $\tilde{\beta}_1$

Unbiasedness ~

$$\tilde{\beta}_1 = \frac{\Sigma(z_i - \bar{z})(\beta_0 + \beta_1 x_i + u_i)u_i}{\Sigma(z_i - \bar{z})x_i}$$

$$\hookrightarrow \tilde{\beta}_1 = \cancel{\Sigma(z_i - \bar{z})}\beta_0 + \Sigma(z_i - \bar{z})\beta_1 x_i + \Sigma(z_i - \bar{z})u_i$$

$$\hookrightarrow \frac{\beta_1 \Sigma(z_i - \bar{z})x_i + \Sigma(z_i - \bar{z})u_i}{\Sigma(z_i - \bar{z})x_i}$$

$$\hookrightarrow \beta_1 + \frac{\Sigma(z_i - \bar{z})u_i}{\Sigma(z_i - \bar{z})u_i} = \tilde{\beta}_1$$

where z_i & x_i are under GM 1-4 including nonrandomness...

$\hookrightarrow$ so $E(\hat{\beta}_1 | X) = \beta_1 \longrightarrow \tilde{\beta}_1$ is unbiased

(ii) Adding MLR5 (Homoskedasticity), $Var(\tilde{\beta}_1) = \sigma^2 \left[\Sigma(z_i - \bar{z})^2\right] / \left[\Sigma(z_i - \bar{z})x_i\right]^2$

Proof: $\sigma^2\left[\Sigma(z_i - \bar{z})^2\right]\left[\Sigma(z_i - \bar{z})x_i\right]^{-2} \longrightarrow \sigma^2 = Var \Rightarrow Var\left[\Sigma(z_i - \bar{z})u_i\right]$

$\hookrightarrow = \Sigma(z_i - \bar{z})^2 \sigma^2 \Rightarrow \sigma^2 \Sigma(z_i - \bar{z})^2 \rightarrow$ so... $\dfrac{\sigma^2 \Sigma(z - \bar{z})^2}{\Sigma(z_i - \bar{z})x_i}$

(iii) Under GM, $Var(\hat{\beta}_1) \leq Var(\tilde{\beta}_1)$

I think the core intuition here is that the estimator $\hat{\beta}_1$ will have lower variance than the non OLS estimator $\tilde{\beta}_1$ because the OLS estimator accounts for weights and balances for sample info. It minimizes the distance between observed and predicted values, adjusting the model to better represent the data.

Proof

$$Var(\tilde{\beta}_1) = \frac{\sigma^2 \sum(z_i-\bar{z})^2}{(\sum(z_i-\bar{z}))^2} \quad \text{versus} \quad Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(x_i-\bar{x}_i)^2}$$

Cauchy-Schwarz says $\rightarrow$ $\left(n^{-1}\sum(z_i-\bar{z})(x_i-\bar{x})\right)^2 \leq \left(n^{-1}\sum(z_i-\bar{z})^2\right)\left(n^{-1}\sum(x_i-\bar{x})^2\right)$

$\hookrightarrow * n^2 \Rightarrow \left(\sum(z_i-\bar{z})(x_i-\bar{x})\right)^2 \leq \left(\sum(z_i-\bar{z})^2\right)\left(\sum(x_i-\bar{x})^2\right)$

$$\hookrightarrow \underbrace{\frac{1}{\left(\sum(z_i-\bar{z})(x_i-\bar{x})\right)^2}}_{\text{smaller Var}} \leq \frac{\left(\sum(z_i-\bar{z})^2\right)}{\left(\sum(z_i-\bar{z})(x_i-\bar{x})\right)^2}$$

3.18 (i) w = grant \$, $y(w)$ = College performance, assume $y(w) = \alpha + \beta w + v(0)$ where $y(w) = \alpha + v$

For each i, we can write $y_i = \alpha + \beta w_i + v_i$, $E(v_i|w_i) = 0$

* Since we can assume that "for all i's, w_i is independent of v_i", we would expect the average error v given some w should be 0. This is part of the MLR zero mean error assumption.

(ii) Given a random sample, I would estimate α and β with OLS. This is because the information provided allows us to assume the model is covered by GM assumptions 1-4, allowing for unbiased estimates.

(iii) we can write $y_i = \psi + \beta w_i + \gamma_1 x_{i1} + \ldots + \gamma_K x_{iK} + u_i$

$$E(u_i \mid w_i, x_{i1}, \ldots, x_{iK}) = \emptyset$$

we model V_i as a function of $E(X_i)$

$$\longrightarrow \quad y_i = \alpha + \beta w_i + (\eta + \gamma_1 x_{i1} + \ldots + \gamma_1 x_{iK}) + u_i$$

$$\hookrightarrow \quad y_i = \psi + \beta w_i + \gamma_1 x_{i1} + \ldots + \gamma_K x_{iK} + u_i$$

with $\psi = \alpha + \eta$ and $\mathbb{E}[u_i \mid w_i, X_{i1}, \ldots X_{iK}] = \emptyset$

Here, we cannot promise that the policy level (w_i) is independent of the error (V_i) meaning an endogeneity problem. We can however say that in the subset X_{ij} there is no correlation between W_i and V_i meaning the relationship flows only through their covariates. This is decomposition.

$$\text{Since} \quad \mathbb{E}[V_i \mid w_i, X_{ij} \ldots X_{iK}] = \mathbb{E}[V_i \mid X_{i1}, \ldots X_{iK}]$$

*If endogeneity is present in observable variables, we can condition on them to close that back door leaving only random error and making β unbiased.

(iv) To estimate β with ψ and γ_j in part (iii) I would run a regression with OLS, controlling for all included X_{ij} observable confounders so I can make unbiased estimates of some policy (w_i) effect on y_i after closing the back door.

the model is $\quad y_i = \psi + \beta w_i + \gamma_1 X_{ij} + \ldots + \gamma_K X_{iK} + u_i$

where β represents the average effect of a one unit change in policy level W controlling for all observed characteristics (holding all X_{ij} fixed). X_{ij} are covariates, γ_{ij} is their respective coefficients (avg effects CP), u_i is the error term with $E[u \mid w_i, X_{iK}] = \emptyset$, and ψ is the intercept (which may be sort of meaningless here?).