Methods of Economic Investigation II (EC403)
Problem Set #3

Selected Solutions

1. Write

$$\begin{aligned}\hat{\alpha}_1 &= \bar{y}_1 - \hat{\beta}_1 \bar{x}_1 \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^{n_1}(y_i - \bar{y}_1)(x_i - \bar{x}_1)}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2}\end{aligned}$$

$$\begin{aligned}\hat{\alpha}^* &= \bar{y} - \hat{\beta}^* \bar{x}^* = \bar{y} - \hat{\beta}_1 \bar{x}_1 \\ \hat{\beta}^* &= \frac{\sum_{i=1}^{n}(x_i^* - \bar{x}_1)y_i}{\sum_{i=1}^{n}(x_i^* - \bar{x}_1)^2} = \frac{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)y_i}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2} = \hat{\beta}_1,\end{aligned}$$

because $\bar{x}_i^* = \bar{x}_1$ for all $i > n_1$.

(a) Standard assumptions of the linear regression model leads to both $\hat{\alpha}_1$ and $\hat{\beta}_1$unbiased estimators of α and β, i.e.,

$$E(\hat{\alpha}_1) = \alpha \text{ and } E(\hat{\beta}_1) = \beta$$

so these estimators are unbiased. Since $\hat{\beta}^* = \hat{\beta}_1$, this is obviously unbiased too. Note that

$$E(\overline{y}) = \alpha + \beta\overline{x},$$

where $\overline{x} = n^{-1}\sum_{i=1}^{n} x_i$, and so

$$E\left(\hat{\alpha}^*\right) = \alpha + \beta\left(\overline{x} - \overline{x}_1\right).$$

In general $\hat{\alpha}^*$ will be biased to the extent that $\overline{x} \neq \overline{x}_1$.

(b) We then calculate the variances. The variance of $\hat{\beta}_1$ is

$$\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2},$$

while

$$\begin{aligned}\text{var}(\hat{\alpha}_1) &= \text{var}(\bar{y}_1 - \hat{\beta}_1\bar{x}_1) = \text{var}(\bar{y}_1) + \bar{x}_1^2\text{var}(\hat{\beta}_1) - 2\bar{x}_1\text{cov}(\bar{y}_1, \hat{\beta}_1) \\ &= \frac{\sigma^2}{n_1} + \frac{\bar{x}_1^2\sigma^2}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2},\end{aligned}$$

because

$$\begin{aligned}\operatorname{cov}(\bar{y}_1, \hat{\beta}_1) &= \operatorname{cov}(\alpha + \beta\bar{x}_1 + \bar{\varepsilon}_1, \hat{\beta}_1) = \operatorname{cov}\left(\bar{\varepsilon}_1, \frac{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)\varepsilon_i}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2}\right) \\ &= \frac{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)E(\bar{\varepsilon}_1\varepsilon_i)}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2} = \frac{\frac{\sigma^2}{n}\sum_{i=1}^{n_1}(x_i - \bar{x})}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2} = 0.\end{aligned}$$

We do the same for $\hat{\alpha}^*$ and $\hat{\beta}^*$. Since $\hat{\beta}^* = \hat{\beta}_1$ we just calculate the variance of $\hat{\alpha}^*$, which is

$$\operatorname{var}(\hat{\alpha}^*) = \frac{\sigma^2}{n} + \operatorname{var}(\hat{\beta}_1)\bar{x}_1 - 2\operatorname{cov}(\bar{y}, \hat{\beta}_1) = \frac{\sigma^2}{n} + \frac{\bar{x}_1^2\sigma^2}{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2}$$

again because $\operatorname{cov}(\bar{y}, \hat{\beta}_1) = 0$, hence

$$\operatorname{var}(\hat{\alpha}^*) \leq \operatorname{var}(\hat{\alpha}_1).$$

2. Take

$$X_n \sim N(0, 1 + \frac{1}{n}) \text{ and } X \sim N(0,1) \text{ with } \operatorname{cov}(X_n, X) = 0 \;\; \forall\, n.$$

By definition $X_n \xrightarrow{d} X$ if $\lim_{n \longrightarrow \infty} F_{X_n}(x) = F_X(x)$ for every continuity point x. Denoting $\Phi(x) = F_X(x)$ the c.d.f. of a $N(0,1)$ then

$$F_{X_n}(x) = \Phi\left(\frac{x}{\sqrt{1 + \frac{1}{n}}}\right).$$

The point is: does $F_{X_n}(x) \longrightarrow F_X(x)$ as $n \longrightarrow \infty$? Yes because

$$\frac{x}{\sqrt{1 + \frac{1}{n}}} \longrightarrow x \text{ as } n \longrightarrow \infty$$

and hence, given that $\Phi(x)$ is continuous

$$\Phi\left(\frac{x}{\sqrt{1 + \frac{1}{n}}}\right) \longrightarrow \Phi(x) \text{ as } n \longrightarrow \infty.$$

For convergence in probability we need that $\forall \delta > 0$

$$\lim_{n \longrightarrow \infty} \Pr(|X_n - X| > \delta) = 0.$$

In our case let $Z_n = X_n - X$, which is $N(0, 2 + \frac{1}{n})$, i.e., $F_{Z_n}(x) = \Phi\left(\frac{x}{\sqrt{2+\frac{1}{n}}}\right)$. Then

$$\Pr(|Z_n| > \delta) = 2\Pr(Z_n > \delta) = 2(1 - \Pr(Z_n \leq \delta)) = 2\left(1 - \Phi\left(\frac{\delta}{\sqrt{2 + \frac{1}{n}}}\right)\right)$$

and so

$$\lim_{n\longrightarrow\infty}\left[\Pr(|Z_n|>\delta)\right]=2\left[1-\Phi\left(\frac{\delta}{\sqrt{2}}\right)\right]\neq 0,$$

at least for $\delta\neq\infty$. Therefore, $X_n\xrightarrow{d}X$ does not imply that $X_n\xrightarrow{p}X$.

Suppose that $T(\hat{\theta}-2\pi)\xrightarrow{d}N(0,1)$. Using a Taylor expansion we get

$$\begin{aligned}\sin\hat{\theta} &= \sin 2\pi+\cos(2\pi)(\hat{\theta}-2\pi)-\frac{1}{2}\sin(\tilde{\theta})(\hat{\theta}-2\pi)^2\\ &= (\hat{\theta}-2\pi)-\frac{1}{2}\sin(\tilde{\theta})(\hat{\theta}-2\pi)^2,\end{aligned}$$

where $\tilde{\theta}$ is an intermediate point between $\hat{\theta}$ and 2π. Then

$$T\sin(\hat{\theta})=T(\hat{\theta}-2\pi)-\frac{1}{2}\sin(\tilde{\theta})T(\hat{\theta}-2\pi)^2.$$

By assumption we know that $T(\hat{\theta}-2\pi)\xrightarrow{d}N(0,1)$, hence $T^2(\hat{\theta}-2\pi)^2\xrightarrow{d}\chi^2_{(1)}$, which means that $(\hat{\theta}-2\pi)^2$ is $O_p(T^{-2})$ and $T(\hat{\theta}-2\pi)^2$ is $O_p(T^{-1})=o_p(1)$. Given that $\sin(\tilde{\theta})$ is bounded between (-1,+1) it follows immediately that $T\sin(\hat{\theta})\xrightarrow{d}N(0,1)$.

3. Grouping the data into J categories we rewrite our regression model as

$$Y_j=\beta X_j+U_j \tag{1}$$

where all the variables are defined as in the problem. The OLS regression of Y_j on X_j gives the coefficient

$$\hat{\beta}=\frac{\sum_{j=1}^{J}X_jY_j}{\sum_{j=1}^{J}X_j^2} \tag{2}$$

the grouped regression model satisfies $E(U_j)=0$ and $E(U_jU_k)=0\ \forall j\neq k$, but we have heteroskedastic errors

$$E\left(U_j^2\right)=\frac{\sigma^2}{n_j}.$$

To derive the limiting distribution of

$$\hat{\beta}-\beta=\frac{\sum_{j=1}^{J}X_jU_j}{\sum_{j=1}^{J}X_j^2}$$

we will make use of the Lindberg-Feller CLT and Slutsky's theorem.

Assuming that:

$$\frac{1}{J}\sum_{j}^{J}X_j^2\xrightarrow{p}a>0,a<\infty$$

$$\frac{1}{J}\sum_{j}^{J}\frac{X_j^2}{n_j}\xrightarrow{p}q>0,q<\infty$$

$$\frac{1}{\sigma\sqrt{\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}}\sum_{j=1}^{J} X_j U_j \xrightarrow{d} N(0,1)$$

Then

$$\frac{1}{\sqrt{J}}\sum_{j=1}^{J} X_j U_j = \sigma\sqrt{\frac{1}{J}\sum_{j=1}^{J}\frac{X_j^2}{n_j}}\frac{1}{\sigma\sqrt{\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}}\sum_{j=1}^{J} X_j U_j \xrightarrow{d} N(0,\sigma^2 q).$$

Finally

$$\sqrt{J}(\hat{\beta}-\beta) = \frac{\frac{1}{\sqrt{J}}\sum_{j=1}^{J} X_j U_j}{\frac{1}{J}\sum_{j}^{J} X_j^2} \xrightarrow{d} N(0,\sigma^2\frac{q}{a^2}).$$

To prove this result we used the Lindberg-Feller CLT: let x_1, ..., x_n be i.n.i.d random variables with $E(x_i)=0$ for all i. If for all $\varepsilon>0$

$$\frac{1}{\sum_{i=1}^{n}\sigma_i^2}\sum_{i=1}^{n} E\left[x_i^2 1\left(x_i^2>\varepsilon\sum_{i=1}^{n}\sigma_i^2\right)\right] \longrightarrow 0$$

as $n\longrightarrow\infty$, then

$$\frac{1}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}}\sum_{i=1}^{n} x_i \xrightarrow{d} N(0,1)$$

Now we want to get the asymptotic distribution for

$$\sum_{j=1}^{J} X_j U_j \tag{3}$$

assuming $(X_j U_j)$ an independent sequence with $E\left(X_j U_j\right)=0$, $\text{var}(X_j U_j)=X_j^2\text{var}(U_j)=X_j^2\frac{\sigma^2}{n_j}$. If for all $\varepsilon>0$,

$$\frac{1}{\sigma^2\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}\sum_{j=1}^{J} X_j^2 E\left[U_j^2 1\left(X_j^2 U_j^2>\varepsilon\sum_{j=1}^{J}\frac{X_j^2}{n_j}\sigma^2\right)\right] \longrightarrow 0$$

as $n\longrightarrow\infty$, then

$$\frac{1}{\sigma\sqrt{\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}}\sum_{j=1}^{J} X_j U_j \xrightarrow{d} N(0,1)$$

as required.

To test $\beta=0$ we simply use the statistic perviously derived, but which requires to know σ, otherwise we can use

$$\frac{\hat{\beta}-\beta}{s\sqrt{\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}} = \frac{1}{s\sqrt{\sum_{j=1}^{J}\left(\frac{X_j^2}{n_j}\right)}}\sum_{j=1}^{J} X_j U_j \xrightarrow{d} N(0,1)$$

under H_0 if and only if s^2 is a consistent estimator for σ^2. The formula for s^2 is

$$s^2 = \frac{1}{J}\sum_{j=1}^{J} n_j (Y_j - \hat{\beta} X_j)^2.$$

To prove that it is a consistent estimator for σ^2 we proceed as follows

$$s^2 = \frac{1}{J}\sum_{j=1}^{J} n_j U_j^2 + \frac{1}{J}\sum_{j=1}^{J} n_j (\hat{\beta}-\beta)^2 X_j^2 - \frac{2}{J}\sum_{j=1}^{J} n_j U_j (\hat{\beta}-\beta) X_j$$

we then can finally say:

$$\frac{1}{J}\sum_{j=1}^{J} n_j U_j^2 \xrightarrow{p} \frac{1}{J}\sum_{j=1}^{J} E(n_j U_j^2) = \sigma^2$$

$$\frac{1}{J}\sum_{j=1}^{J} n_j (\hat{\beta}-\beta)^2 X_j^2 = O_p(\frac{1}{J})$$

$$\frac{2}{J}\sum_{j=1}^{J} n_j U_j (\hat{\beta}-\beta) X_j = O_p(\frac{1}{\sqrt{J}}).$$

4. To derive the asymptotic distribution of $\hat{\beta}_r$ we rewrite it as

$$\hat{\beta}_r = \frac{(\beta x + u)'(\beta x + u)}{x'(\beta x + u)} = \frac{\beta^2 x'x + \beta x'u + \beta u'x + u'u}{\beta x'x + x'u}$$

$$= \frac{\beta^2 x'x + 2\beta x'u + u'u}{\beta x'x + x'u}$$

and dividing and multiplying by $x'x$

$$\hat{\beta}_r = \frac{\beta^2 + 2\beta\frac{x'u}{x'x} + \frac{u'u}{x'x}}{\beta + \frac{x'u}{x'x}}.$$

At this point we must make some assumptions:

- x_i and u_i i.i.d.
- Let $E|x_i u_i| < \infty$ then $\frac{1}{n}\sum x_i u_i \xrightarrow{p} 0$ (call $w_i = x_i u_i$, then $E(w_i) = 0$).
- Let $E(x_i^2) = \sigma_x^2$ then $\frac{1}{n}\sum x_i^2 \xrightarrow{p} \sigma_x^2$.
- Let $E(u_i^2) = \sigma_u^2$ then $\frac{1}{n}\sum u_i^2 \xrightarrow{p} \sigma_u^2$.

We can therefore state that

$$\hat{\beta}_r \longrightarrow \frac{\beta^2 + \frac{\sigma_u^2}{\sigma_x^2}}{\beta} = \beta + \frac{\sigma_u^2}{\beta\sigma_x^2} = \beta_r.$$

Next, write

$$\hat{\beta}_r - \beta_r = \frac{(\beta^2 x'x + 2\beta x'u + u'u)\beta\sigma_x^2 - (\beta^2\sigma_x^2 + \sigma_u^2)(\beta x'x + x'u)}{(\beta x'x + x'u)\,\beta\sigma_x^2}$$

$$= \frac{\beta^3 x'x\sigma_x^2 + 22\beta x'u\sigma_x^2 + u'u\beta\sigma_x^2 - \beta^3 x'x\sigma_x^2 - \beta^2\sigma_x^2 x'u - \beta x'x\sigma_u^2 - \sigma_u^2 x'u}{(\beta x'x + x'u)\,\beta\sigma_x^2},$$

which after some simplifications and after adding and subtracting $\beta\sigma_x^2\sigma_u^2$ we get

$$\hat{\beta}_r - \beta_r = \frac{(\beta\sigma_x^2 - \sigma_u^2)\,x'u + \beta\sigma_x^2(u'u - \sigma_u^2) - \beta\sigma_u^2(x'x - \sigma_x^2)}{(\beta x'x + x'u)\,\beta\sigma_x^2}.$$

We can now write $\sqrt{n}\left(\hat{\beta}_r - \beta_r\right) = a_n' z_n$ (scalar), where

$$a_n = \left[\frac{\beta\sigma_x^2 - \sigma_u^2}{\frac{1}{n}(\beta x'x + x'u)\,\beta\sigma_x^2}\ ,\ \frac{1}{\frac{1}{n}(\beta x'x + x'u)}\ ,\ \frac{-\beta\sigma_u^2}{\frac{1}{n}(\beta x'x + x'u)\,\sigma_x^2}\right]'$$

$$z_n = \left[\frac{(x'u)}{\sqrt{n}}\ ,\ \frac{(u'u - \sigma_u^2)}{\sqrt{n}}\ ,\ \frac{(x'x - \sigma_x^2)}{\sqrt{n}}\right]'.$$

Under standard conditions

$$\frac{1}{n}(\beta x'x + x'u)\,\beta\sigma_x^2 \xrightarrow{p} \beta^2\sigma_x^4$$

etc and so $a_n \longrightarrow a$, where

$$a = \left[\frac{\beta\sigma_x^2 - \sigma_u^2}{\beta^2\sigma_x^4}\ ,\ \frac{1}{\beta\sigma_x^2}\ ,\ -\frac{\beta\sigma_u^2}{\beta\sigma_x^4}\right]'.$$

Furthermore,

$$z_n = \begin{pmatrix} \frac{(x'u)}{\sqrt{n}} \\ \frac{(u'u - \sigma_u^2)}{\sqrt{n}} \\ \frac{(x'x - \sigma_x^2)}{\sqrt{n}} \end{pmatrix} \xrightarrow{d} N\left[0\ ,\ \Sigma\right],$$

where

$$\Sigma = \begin{pmatrix} E(x_i^2 u_i^2) & \mathrm{cov}(x_i u_i, u_i^2) & \mathrm{cov}(x_i u_i, x_i^2) \\ \mathrm{cov}(x_i u_i, u_i^2) & \mathrm{var}(u_i^2) & \mathrm{cov}(u_i^2, x_i^2) \\ \mathrm{cov}(x_i u_i, x_i^2) & \mathrm{cov}(u_i^2, x_i^2) & \mathrm{var}(x_i^2) \end{pmatrix}.$$

Finally we can apply the Cramer's theorem: if $Z_n \xrightarrow{d} N(\mu, \Sigma)$ and $A_n \xrightarrow{p} A$ then $A_n Z_n \longrightarrow N(A\mu, A\Sigma A')$. This gives that

$$\sqrt{n}(\hat{\beta}_r - \beta_r) \longrightarrow N(0, a'\Sigma a).$$