Problem Set 7 – Solutions

ECON 837

Brian Krauth (adapted from problems by Simon Woodcock), Spring 2010

1. We know from the previous assignment that:

$$\begin{aligned}
\text{plim}\frac{1}{n}\sum_{i=1}^{n} x_i &= E[x_i] = \int_0^1 x_i dx_i = \left[\frac{1}{2}x_i^2\right]_0^1 = \frac{1}{2}. \\
\text{plim}\frac{1}{n}\sum_{i=1}^{n} x_i^2 &= E[x_i^2] = \int_0^1 x_i^2 dx_i = \left[\frac{1}{3}x_i^3\right]_0^1 = \frac{1}{3} \\
\text{plim}\frac{1}{n}\sum_{i=1}^{n} x_i^3 &= E[x_i^3] = \int_0^1 x_i^3 dx_i = \left[\frac{1}{4}x_i^4\right]_0^1 = \frac{1}{4}
\end{aligned}$$

all of which follows immediately from Khinchine's WLLN.

(a) We know that

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_i (x_i - \bar{x}_n) y_i}{\sum_i (x_i - \bar{x}_n)^2} = \frac{\sum_i (x_i - \bar{x}_n)(x_i^2 + \varepsilon_i)}{\sum_i (x_i - \bar{x}_n)^2} = \frac{\sum_i x_i^2 (x_i - \bar{x}_n) + \sum_i \varepsilon_i (x_i - \bar{x}_n)}{\sum_i (x_i - \bar{x}_n)^2} \\
&= \frac{\frac{1}{n}\sum_i x_i^3 - \frac{1}{n}\sum_i x_i^2 \bar{x}_n + \frac{1}{n}\sum_i x_i \varepsilon_i - \frac{1}{n}\sum_i \varepsilon_i \bar{x}_n}{\frac{1}{n}\sum_i (x_i - \bar{x}_n)^2}.
\end{aligned}$$

We can compute the probability limits of each term independently:

$$\begin{aligned}
\text{plim}\frac{1}{n}\sum_i (x_i - \bar{x}_n)^2 &= \text{plim}\frac{1}{n}\sum_i \left(x_i^2 - n\bar{x}_n^2\right) = \text{plim}\frac{1}{n}\sum_i x_i^2 - \text{plim}\bar{x}_n^2 = \frac{1}{3} - \text{plim}\left(\frac{1}{n}\sum_i x_i\right)^2 \\
&= \frac{1}{3} - \left(\text{plim}\frac{1}{n}\sum_i x_i\right)^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12} \\
\text{plim}\frac{1}{n}\sum_i x_i^3 &= \frac{1}{4} \\
\text{plim}\frac{1}{n}\sum_i x_i^2 \bar{x}_n &= \text{plim}\bar{x}_n \frac{1}{n}\sum_i x_i^2 = \text{plim}\bar{x}_n \text{plim}\frac{1}{n}\sum_i x_i^2 = \text{plim}\frac{1}{n}\sum_i x_i \text{plim}\frac{1}{n}\sum_i x_i^2 = \frac{1}{2}\frac{1}{3} = \frac{1}{6}.
\end{aligned}$$

Furthermore, $\frac{1}{n}\sum_i x_i\varepsilon_i$ is yet another sample mean of iid random variables $x_i\varepsilon_i$ that have finite first moment. Thus, by Khinchine's WLLN:

$$\text{plim}\frac{1}{n}\sum_i x_i \varepsilon_i = E[x_i \varepsilon_i] = E[\varepsilon_i] E[x_i] = 0$$

because x_i and ε_i are independent. Likewise, applying Khinchine's WLLN again:

$$\text{plim}\frac{1}{n}\sum_i \varepsilon_i \bar{x}_n = \text{plim}\bar{x}_n \frac{1}{n}\sum_i \varepsilon_i = \text{plim}\bar{x}_n \text{plim}\frac{1}{n}\sum_i \varepsilon_i = \text{plim}\frac{1}{n}\sum_i x_i \text{plim}\frac{1}{n}\sum_i \varepsilon_i = \frac{1}{2}E[\varepsilon_i] = 0.$$

Putting it all together,

$$\text{plim}\hat{\beta} = \frac{\frac{1}{4} - \frac{1}{6}}{\frac{1}{12}} = 1.$$

As for the intercept,

$$\hat{\alpha} = \bar{y}_n - \hat{\beta}\bar{x}_n = \frac{1}{n}\sum_i y_i - \hat{\beta}\frac{1}{n}\sum_i x_i = \frac{1}{n}\sum_i x_i^2 + \frac{1}{n}\sum_i \varepsilon_i - \hat{\beta}\frac{1}{n}\sum_i x_i$$

and hence

$$\begin{aligned}
\text{plim}\hat{\alpha} &= \text{plim}\frac{1}{n}\sum_i x_i^2 + \text{plim}\frac{1}{n}\sum_i \varepsilon_i - \text{plim}\hat{\beta}\text{plim}\frac{1}{n}\sum_i x_i \\
&= \frac{1}{3} + 0 - \frac{1}{2} = -\frac{1}{6}.
\end{aligned}$$

(b) The "average derivative" is

$$E\left[\frac{dy}{dx}\right]=E\left[\frac{d}{dx}\left(x_i^2+\varepsilon_i\right)\right]=E\left[2x_i\right]=2E\left[x_i\right]=1.$$

So we can say

$$\hat{\beta}\xrightarrow{p}E\left[\frac{dy}{dx}\right]$$

i.e., $\hat{\beta}$ is a consistent estimator of the average derivative.

(c) The least squares estimator of γ solves

$$\min_{\gamma}\sum_i\left(y_i-\gamma x_i\right)^2$$

which has FOC:

$$2\sum_i\left(y_i-\hat{\gamma}x_i\right)x_i=0$$

and hence the least squares estimator is

$$\hat{\gamma}=\frac{\sum_i x_i y_i}{\sum_i x_i^2}=\frac{\sum_i x_i\left(x_i^2+\varepsilon_i\right)}{\sum_i x_i^2}=\frac{\frac{1}{n}\sum_i x_i^3+\frac{1}{n}\sum_i x_i\varepsilon_i}{\frac{1}{n}\sum_i x_i^2}.$$

Taking probability limits,

$$\text{plim}\hat{\gamma}=\frac{\text{plim}\frac{1}{n}\sum_i x_i^3+\text{plim}\frac{1}{n}\sum_i x_i\varepsilon_i}{\text{plim}\frac{1}{n}\sum_i x_i^2}=\frac{\frac{1}{4}+0}{\frac{1}{3}}=\frac{3}{4}.$$

Therefore

$$\hat{\gamma}\overset{p}{\not\to}E\left[\frac{dy}{dx}\right]=1$$

i.e., $\hat{\gamma}$ is not a consistent estimator of the average derivative.

2. (a) By the invariance property, the MLE of γ is

$$\hat{\gamma}=\hat{\theta}_1+\frac{1}{2}\hat{\theta}_2^2=6+2=8.$$

See part b for the variance.

(b) Let's use the delta-method to derive the asymptotic distribution of $\hat{\gamma}$. It is,

$$\hat{\gamma}\overset{a}{\sim}N\left(\gamma,\frac{\partial\gamma}{\partial\theta}\left[Asy.Var\left[\theta\right]\right]\frac{\partial\gamma}{\partial\theta'}\right).$$

We estimate the variance by

$$\frac{\partial\gamma}{\partial\hat{\theta}}\left[Asy.Var\left[\hat{\theta}\right]\right]\frac{\partial\gamma}{\partial\hat{\theta}}=\left[\begin{array}{cc}1&\hat{\theta}_2\end{array}\right]\left[\begin{array}{cc}2&1\\1&2\end{array}\right]\left[\begin{array}{c}1\\\hat{\theta}_2\end{array}\right]=\left[\begin{array}{cc}1&2\end{array}\right]\left[\begin{array}{cc}2&1\\1&2\end{array}\right]\left[\begin{array}{c}1\\2\end{array}\right]=14.$$

And hence inference can be based on

$$\hat{\gamma}\overset{a}{\sim}N\left(\gamma,14\right)$$

or

$$z=\frac{\hat{\gamma}-\gamma}{\sqrt{14}}\overset{a}{\sim}N\left(0,1\right).$$

To test the hypothesis $H_0:\gamma=6$ vs. the alternative $H_1:\hat{\gamma}\neq6$, we construct the test statistic

$$z=\frac{\hat{\gamma}-6}{\sqrt{14}}=\frac{2}{\sqrt{14}}\approx0.5345$$

and compare it to critical values of the $N\left(0,1\right)$ distribution. We know the critical value for a test at the 5% level is 1.96. Hence we conclude that z is "small." That is, we do not reject the null hypothesis at the 5% level.

3. The log likelihood is

$$\ln L(\alpha,\beta)=\sum_{i=1}^{n}\ln f(y_i|x_i,\alpha,\beta)=\sum_{i=1}^{n}[\ln\gamma_i-\gamma_i y_i]=\sum_{i=1}^{n}[\alpha+\beta x_i-y_i\exp(\alpha+\beta x_i)].$$

The MLE solves the first order conditions:

$$\begin{aligned}\frac{\partial\ln L(\alpha,\beta)}{\partial\hat{\alpha}} &= \sum_{i=1}^{n}[1-y_i\exp(\alpha_{ML}+\beta_{ML}x_i)]=0\\ \frac{\partial\ln L(\alpha,\beta)}{\partial\hat{\beta}} &= \sum_{i=1}^{n}[x_i-y_i x_i\exp(\alpha_{ML}+\beta_{ML}x_i)]=0\end{aligned}$$

which have no closed form solution. The Hessian has elements

$$\begin{aligned}\frac{\partial^2\ln L(\alpha,\beta)}{\partial\alpha^2} &= -\sum_{i=1}^{n}y_i\exp(\alpha+\beta x_i)\\ \frac{\partial^2\ln L(\alpha,\beta)}{\partial\alpha\partial\beta} &= -\sum_{i=1}^{n}y_i x_i\exp(\alpha+\beta x_i)\\ \frac{\partial^2\ln L(\alpha,\beta)}{\partial\beta^2} &= -\sum_{i=1}^{n}y_i x_i^2\exp(\alpha+\beta x_i).\end{aligned}$$

You might recognize the density of y_i as being that of an exponential random variable with parameter γ_i. Hence

$$E[y_i]=\frac{1}{\gamma_i}=\frac{1}{\exp(\alpha+\beta x_i)}$$

(you can also show this by direct integration). Therefore, the elements of the information matrix are

$$\begin{aligned}-E\left[\frac{\partial^2\ln L(\alpha,\beta)}{\partial\alpha^2}\right] &= E\left[\sum_{i=1}^{n}y_i\exp(\alpha+\beta x_i)\right]=\sum_{i=1}^{n}E[y_i]\exp(\alpha+\beta x_i)=n\\ -E\left[\frac{\partial^2\ln L(\alpha,\beta)}{\partial\alpha\partial\beta}\right] &= E\left[\sum_{i=1}^{n}y_i x_i\exp(\alpha+\beta x_i)\right]=\sum_{i=1}^{n}E[y_i]x_i\exp(\alpha+\beta x_i)=\sum_{i=1}^{n}x_i\\ -E\left[\frac{\partial^2\ln L(\alpha,\beta)}{\partial\beta^2}\right] &= E\left[\sum_{i=1}^{n}y_i x_i^2\exp(\alpha+\beta x_i)\right]=\sum_{i=1}^{n}E[y_i]x_i^2\exp(\alpha+\beta x_i)=\sum_{i=1}^{n}x_i^2\end{aligned}$$

and we have the asymptotic distribution

$$\begin{aligned}&\begin{bmatrix}\sqrt{n}(\alpha_{ML}-\alpha)\\ \sqrt{n}(\beta_{ML}-\alpha)\end{bmatrix}\xrightarrow{d}N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1 & \frac{1}{n}\sum_{i=1}^{n}x_i\\ \frac{1}{n}\sum_{i=1}^{n}x_i & \frac{1}{n}\sum_{i=1}^{n}x_i^2\end{bmatrix}^{-1}\right),\quad\text{or}\\ &\begin{bmatrix}\alpha_{ML}\\ \beta_{ML}\end{bmatrix}\overset{a}{\sim}N\left(\begin{bmatrix}\alpha\\ \beta\end{bmatrix},\begin{bmatrix}n & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2\end{bmatrix}^{-1}\right).\end{aligned}$$

We know the MLE of $\delta=\alpha\exp(\beta)$ is

$$\delta_{ML}=\alpha_{ML}\exp(\beta_{ML})$$

by the invariance property. Its asymptotic distribution is given by the delta method:

$$\delta_{ML}\overset{a}{\sim}N\left(\delta,\begin{bmatrix}\exp(\beta) & \alpha\exp(\beta)\end{bmatrix}\begin{bmatrix}n & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2\end{bmatrix}^{-1}\begin{bmatrix}\exp(\beta)\\ \alpha\exp(\beta)\end{bmatrix}\right).$$

4. We know that

$$\begin{aligned}\hat{\beta} &= \frac{\sum(x_i^*-\bar{x}^*)y_i^*}{\sum(x_i^*-\bar{x}^*)^2}=\frac{\sum(x_i^*-\bar{x}^*)(y_i+\nu_i)}{\sum(x_i^*-\bar{x}^*)^2}=\frac{\sum(x_i^*-\bar{x}^*)(\alpha+\beta x_i+\varepsilon_i)}{\sum(x_i^*-\bar{x}^*)^2}+\frac{\sum(x_i^*-\bar{x}^*)\nu_i}{\sum(x_i^*-\bar{x}^*)^2}\\ &= \beta\frac{\frac{1}{n}\sum(x_i^*-\bar{x}^*)x_i}{\frac{1}{n}\sum(x_i^*-\bar{x}^*)^2}+\frac{\frac{1}{n}\sum(x_i^*-\bar{x}^*)\varepsilon_i}{\frac{1}{n}\sum(x_i^*-\bar{x}^*)^2}+\frac{\frac{1}{n}\sum(x_i^*-\bar{x}^*)\nu_i}{\frac{1}{n}\sum(x_i^*-\bar{x}^*)^2}.\end{aligned}$$

Applying a WLLN to each term, we have:

$$\begin{aligned}
\text{plim}\frac{1}{n}\sum(x_i^*-\bar{x}^*)^2 &= E\left[(x_i^*-\bar{x}^*)^2\right] = Var\left[x_i^*\right] = \sigma_x^2+\sigma_\eta^2 \\
\text{plim}\frac{1}{n}\sum(x_i^*-\bar{x}^*)\,x_i &= E\left[(x_i^*-\bar{x}^*)\,x_i\right] = E\left[(x_i+\eta_i-\bar{x})\,x_i\right] = Var\left[x_i\right]+Cov\left[x_i,\eta_i\right] = \sigma_x^2 \\
\text{plim}\frac{1}{n}\sum(x_i^*-\bar{x}^*)\,\varepsilon_i &= E\left[(x_i^*-\bar{x}^*)\,\varepsilon_i\right] = E\left[(x_i+\eta_i-\bar{x})\,\varepsilon_i\right] = Cov\left[x_i,\varepsilon_i\right]+Cov\left[\eta_i,\varepsilon_i\right] = 0 \\
\text{plim}\frac{1}{n}\sum(x_i^*-\bar{x}^*)\,\nu_i &= E\left[(x_i^*-\bar{x}^*)\,\nu_i\right] = E\left[(x_i+\eta_i-\bar{x})\,\nu_i\right] = Cov\left[x_i,\nu_i\right]+Cov\left[\eta_i,\nu_i\right] = 0
\end{aligned}$$

so that $\text{plim}\hat{\beta} = \beta\sigma_x^2/\left(\sigma_x^2+\sigma_\eta^2\right)$

(a) If $\sigma_\eta^2 = 0$, then $\text{plim}\hat{\beta} = \beta\sigma_x^2/\sigma_x^2 = \beta$ (the value of σ_ν^2 is irrelevant here). Hence the least squares estimator is consistent.

(b) If $\sigma_\eta^2 \neq 0$, then $\text{plim}\hat{\beta} = \beta\sigma_x^2/\left(\sigma_x^2+\sigma_\eta^2\right) \neq \beta$ (again, the value of σ_ν^2 is irrelevant). Hence the least squares estimator is inconsistent.

(c) Clearly, measurement error in x_i is the problem. Measurement error in y_i is irrelevant – it just becomes part of the error term.

(d) Suppose we observe $w_i = x_i + \xi_i$. Notice that:

$$\begin{aligned}
Cov\left[w_i,x_i^*\right] &= E\left[(x_i+\xi_i-\mu)\,(x_i+\eta_i-\mu)\right] = E\left[(x_i-\mu)^2\right]+E\left[(x_i-\mu)\,\eta_i\right]+E\left[(x_i-\mu)\,\xi_i\right]+E\left[\eta_i\xi_i\right] \\
&= Var\left[x_i\right]+Cov\left[x_i,\eta_i\right]+Cov\left[x_i,\xi_i\right]+Cov\left[\eta_i,\xi_i\right] \\
&= \sigma_x^2
\end{aligned}$$

because the covariances are all zero by assumption. So we can easily estimate σ_x^2 from this covariance, and we can estimate $\sigma_x^2+\sigma_\eta^2$ directly from the observed x_i^*. Using these, we can correct our estimate $\hat{\beta}$ and obtain a consistent estimate of β (note: this is an IV estimator, where w_i is an instrument for x_i).