Advanced Econometric Theory (EC481)
Problem Set #2 Solutions

1. We have $A_i = W_i L_i$ is i.i.d. with

$$E(A_i^k) = E(W_i^k)E(L_i^k),$$

by independence. Furthermore,

$$E(W_i^k) = E[(W_0 + \eta_i)^k] = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} W_0^{k-j} E((\eta_i)^j).$$

In particular, we have

$$EA_i = W_0 L_0 \text{ and } E(A_i^2) = W_0^2 L_0^2 + W_0^2 \sigma_\varepsilon^2 + L_0^2 \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\eta^2.$$

If it is known that $\sigma_\varepsilon^2 = \sigma_\eta^2 = 1$, then we can estimate the parameters W_0 and L_0 very simply from the mean and variance, where

$$\text{var}(A_i) = W_0^2 + L_0^2 + 1$$

in this case. The centred moments seem to be more convenient here [they eliminate the common factor $(W_0 L_0)^k$]. Thus we take

$$\begin{aligned} m_1(A_i, \theta) &= A_i - WL \\ m_2(A_i, \theta) &= (A_i - \overline{A})^2 - W^2 - L^2 - 1 \end{aligned}$$

and find a zero of the vector $G_n(\theta) = (\sum_{i=1}^n m_1(A_i, \theta), \sum_{i=1}^n m_2(A_i, \theta))/n$. The uniform law of large numbers condition is easy to verify here because the parameters separate from the data [note that this does not quite fit into our framework because of $\overline{A}$, but it is easy to modify the proof in this case]. The identification condition is also easy here because the probability limit of $G_n(\theta)$ is

$$G(\theta) = \begin{pmatrix} W_0 L_0 - WL \\ W_0^2 + L_0^2 - (W^2 + L^2) \end{pmatrix},$$

which can easily be seen to be zero if and only if $W = W_0$ and $L = L_0$. Suppose that the variance parameter is unknown but common, i.e., $\sigma_\varepsilon^2 = \sigma_\eta^2 = \sigma^2$. We then have to compute the third centred moment. We use the following decomposition

$$A_i - EA_i = (W_i - W_0)L_0 + (L_i - L_0)W_0 + (W_i - W_0)(L_i - L_0),$$

which is of the form $(a + b + c)$ so that

$$E(A_i - EA_i)^3 = 6W_0L_0E(W_i - W_0)^2(L_i - L_0)^2 = 6W_0L_0\sigma^4,$$

because all other terms are zero (only abc terms are non zero in expectation) when ε and η are symmetrically distributed. Therefore, we can estimate the parameter directly by

$$\widehat{\sigma}^2 = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^n (A_i - \overline{A})^3}{6\frac{1}{n}\sum_{i=1}^n A_i}},$$

and then use the moments

$$\begin{aligned} m_1(A_i, \theta) &= A_i - WL \\ m_2(A_i, \theta) &= (A_i - \overline{A})^2 - \widehat{\sigma}^2W^2 - \widehat{\sigma}^2L^2 - \widehat{\sigma}^4. \end{aligned}$$

As we will see this will give the same results as if we knew σ^2. Alternatively, one can solve directly the three moments

$$\begin{aligned} m_1(A_i, \theta) &= A_i - WL \\ m_2(A_i, \theta) &= (A_i - \overline{A})^2 - \sigma^2W^2 - \sigma^2L^2 - \sigma^4 \\ m_3(A_i, \theta) &= (A_i - \overline{A})^3 - \sigma^4WL. \end{aligned}$$

If the error variances are different, we need an additional moment condition. When the errors are normally distributed,

$$\begin{aligned} E(\eta_iL_0 + \varepsilon_iW_0 + \eta_i\varepsilon_i)^4 &= E(\eta_i^4)L_0^4 + E(\varepsilon_i^4)W_0^4 + E(\eta_i^4)E(\varepsilon_i^4) + 6E(\eta_i^4)\sigma_\varepsilon^2L_0^2 + \\ &\quad 6E(\varepsilon_i^4)\sigma_\eta^2W_0^2 + 6\sigma_\eta^2\sigma_\varepsilon^2L_0^2W_0^2 \\ &= 3\sigma_\eta^4(L_0^4 + 6\sigma_\varepsilon^2L_0^2) + 3\sigma_\varepsilon^4(W_0^4 + 6\sigma_\eta^2W_0^2) + \sigma_\eta^2\sigma_\varepsilon^2(6L_0^2W_0^2 + 9) \end{aligned}$$

because $E(\eta_i^4) = 3\sigma_\eta^4$ and $E(\varepsilon_i^4) = 3\sigma_\varepsilon^4$.[1] Therefore, the fourth cumulant of A_i is this quantity minus three times the squared variance $[W_0^2\sigma_\varepsilon^2 + L_0^2\sigma_\eta^2 + \sigma_\varepsilon^2\sigma_\eta^2]$

$$\kappa_4(A_i) = 12\sigma_\eta^4\sigma_\varepsilon^2L_0^2 + 12\sigma_\varepsilon^4\sigma_\eta^2W_0^2 + 9\sigma_\varepsilon^2\sigma_\eta^2 - 3\sigma_\varepsilon^4\sigma_\eta^4.$$

[1] Note that

$$\begin{aligned} (a + b + c)^4 &= 12abc^2 + 12ab^2c + 12a^2bc \\ &\quad +4a^3b + 4a^3c + 6a^2b^2 \\ &\quad +6a^2c^2 + 4ab^3 + 4ac^3 \\ &\quad +4b^3c + 6b^2c^2 + 4bc^3 \\ &\quad +a^4 + b^4 + c^4. \end{aligned}$$

When the errors are not normally distributed, then maybe we need the fifth moment, but

$$E(\eta_i L_0 + \varepsilon_i W_0 + \eta_i \varepsilon_i)^5 = 20\sigma_\eta^2 L_0 E(\varepsilon_i^4) W_0^3 + 20E(\eta_i^4)E(\varepsilon_i^4) W_0 L_0 + 20E(\eta_i^4) L_0^3 \varepsilon_i^2 W_0$$

and this depends on the fourth moments of both error terms, which adds two extra unknown quantities. It seems like one can't solve this problem without normality.

Consider the related question where instead of the area one observes the perimeter $P_i = 2(W_i + L_i)$. In this case,

$$E(P_i^k) = 2^k E[(W_0 + L_0 + \eta_i + \varepsilon_i)^k] = 2^k \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} (W_0 + L_0)^{k-j} E((\eta_i + \varepsilon_i)^j).$$

For example,

$$EP_i = 2(W_0 + L_0), \quad EP_i^2 = 4(W_0 + L_0)^2 + \sigma_\eta^2 + \sigma_\varepsilon^2.$$

Unfortunately, every moment and hence cumulant is a function of $W_0 + L_0$ and one cannot get any new information from higher moments. In this case, the parameters are unidentifiable. In the normal case, the distribution of P_i is normal with mean $W_0 + L_0$ and variance $\sigma_\eta^2 + \sigma_\varepsilon^2$ that does not depend on the parameters of interest. ■

2. Let

$$\begin{aligned} G_n(\beta) &= \sum_{i=1}^{n} (y_i - \beta x_i)^3 h(x_i) \\ &= \sum_{i=1}^{n} (\varepsilon_i - (\beta - \beta_0) x_i)^3 h(x_i) \\ &= \sum_{i=1}^{n} \varepsilon_i^3 h(x_i) - 3(\beta - \beta_0) \sum_{i=1}^{n} \varepsilon_i^2 x_i h(x_i) + \\ &\quad + 3(\beta - \beta_0)^2 \sum_{i=1}^{n} \varepsilon_i x_i^2 h(x_i) - (\beta - \beta_0)^3 \sum_{i=1}^{n} x_i^3 h(x_i). \end{aligned}$$

We assume that $E(|\varepsilon_i|^6) < \infty$. In the usual case where x_i is some i.i.d. sequence, $\sum_{i=1}^n \varepsilon_i^3 h(x_i) = O_p(n^{1/2})$, $\sum_{i=1}^n \varepsilon_i^2 x_i h(x_i) = O_p(n)$, $\sum_{i=1}^n \varepsilon_i x_i^2 h(x_i) = O_p(n^{1/2})$, and $\sum_{i=1}^n x_i^3 h(x_i) = O_p(n)$, so that the dominant terms would be the second and third one and the correct normalization would be $1/n$. This question has chosen $x_i = i$. Suppose also that we take $h(i) = i^\theta$, $\theta > -1/2$, then $\sum_{i=1}^n \varepsilon_i^3 h(x_i) = O_p(\sqrt{\sum_{i=1}^n h^2(i)}) = O_p(n^{(2\theta+1)/2})$, $\sum_{i=1}^n \varepsilon_i^2 x_i h(x_i) = O_p(n^{\theta+2})$, $\sum_{i=1}^n \varepsilon_i x_i^2 h(x_i) = O_p(n^{(2\theta+5)/2})$, and $\sum_{i=1}^n x_i^3 h(x_i) = O_p(n^{4+\theta})$. The dominant term in this case is clearly the last one and we redefine the criterion function with this as the normalization. We can show that

$$G_n(\beta) = \frac{\sum_{i=1}^n (y_i - \beta x_i)^3 h(x_i)}{\sum_{i=1}^n x_i^3 h(x_i)} \to^p -(\beta - \beta_0)^3,$$

which has a unique zero at $\beta = \beta_0$.

■

3. There are several estimators we could consider. Maximum likelihood would solve the first order condition

$$\sum_{t=1}^{T} \frac{y_t - \Phi(\beta t)}{\Phi(\beta t)(1 - \Phi(\beta t))} t\phi(\beta t) = 0,$$

where ϕ is the standard normal density function, but an equally valid procedure can be to solve

$$\sum_{t=1}^{T} (y_t - \Phi(\beta t))h(t) = 0$$

for some function h, e.g., $h(t) = t$. In the second case it is easier to simulate the score function unbiasedly. We can replace $\Phi(\beta t)$ by

$$\Phi^{ns}(\beta t) = \frac{1}{ns} \sum_{j=1}^{ns} \{\eta_j \leq \beta t\},$$

where η_j are independent standard normal random variables. Clearly,

$$E_{sim} \Phi^{ns}(\beta t) = \Pr(\eta_j \leq \beta t) = \Phi(\beta t)$$

and so the simulator is unbiased. For the maximum likelihood case you must compute

$$\sum_{t=1}^{T} \frac{y_t - \Phi^{ns,1}(\beta t)}{\Phi^{ns,2}(\beta t)(1 - \Phi^{ns,2}(\beta t))} t\phi(\beta t),$$

where $\Phi^{ns,1}$ and $\Phi^{ns,2}$ are computed from different samples.

In the sequel we assume that we can compute the probability function with no error. We consider the first order condition

$$G_n(\beta) = \sum_{t=1}^{n} t(y_t - \Phi(\beta t)) = \sum_{t=1}^{n} t\varepsilon_t + \sum_{t=1}^{n} t(\Phi(\beta_0 t) - \Phi(\beta t)),$$

where $\varepsilon_t = y_t - \Phi(\beta_0 t)$. The error term has mean zero and variance $\Phi(\beta_0 t)(1 - \Phi(\beta_0 t))$ so that the first term on the right hand side has mean zero and variance

$$V_n = \sum_{t=1}^{n} t^2 \Phi(\beta_0 t)(1 - \Phi(\beta_0 t)).$$

Let $M_n(\beta) = \sum_{t=1}^{n} t(\Phi(\beta_0 t) - \Phi(\beta t))$. Clearly, $M_n(\beta), V_n$ are of smaller order than n^3 and n^2 respectively, but they can be much smaller depending on the value of β. Specifically, when $\beta_0 = 0$, $V_n = O(n^3)$ but for any other β_0, V_n is $O(1)$.

We look for an increasing sequence δ_n such that

$$\frac{1}{\delta_n^2} V_n \to 0 \text{ and } \frac{1}{\delta_n} M_n(\beta) \to M(\beta),$$

where the function $M(\beta) = 0$ if and only if $\beta = \beta_0$. When $\beta_0 = 0$, this can be true when we take $\delta_n = n^2$. In that case, the limit function is

$$M(\beta) = \begin{cases} 0 & \text{if } \beta = \beta_0 \\ \pm 1/4 & \text{any } \beta \neq \beta_0. \end{cases}$$

Roughly speaking, when $\beta = 0$, $\sum_{t=1}^n t\Phi(\beta t) \simeq \sum_{t=1}^n t/2$ but if $\beta > 0$, $\sum_{t=1}^n t\Phi(\beta t) \simeq \sum_{t=1}^n t$. while if $\beta < 0$, $\sum_{t=1}^n t\Phi(\beta t) \simeq -\sum_{t=1}^n t$. Note that the limiting function is discontinuous and any convergence argument that we make cannot be uniform in β, so we can't apply our first theorem. However, we can apply the second theorem that just bounds the behaviour of $M_n(\beta)$ when β is far from β_0.

Now consider what happens when $\beta_0 \neq 0$. We note that

$$\frac{\sum_{t=1}^n t\Phi(\beta t)}{\sum_{t=1}^n t\Phi(\beta_0 t)} \to \pm 1$$

as $n \to \infty$ for any β depending only on the signs of β_0, β. This is proved by showing that the derivative with respect to β, which is

$$\frac{\sum_{t=1}^n t^2\phi(\beta t)}{\sum_{t=1}^n t\Phi(\beta_0 t)},$$

tends to zero. But this is obvious - suppose that $\beta_0 > 0$, then $\sum_{t=1}^n t\Phi(\beta_0 t) \geq \sum_{t=1}^n t/2$ for large enough n, while $\sum_{t=1}^n t^2\phi(\beta t)$ is uniformly bounded for any $\beta > 0$. Therefore, suppose we normalize by $\sum_{t=1}^n t\Phi(\beta_0 t)$ we get

$$G_n(\beta) = \frac{\sum_{t=1}^n t(y_t - \Phi(\beta t))}{\sum_{t=1}^n t\Phi(\beta_0 t)} = \frac{\sum_{t=1}^n t\varepsilon_t}{\sum_{t=1}^n t\Phi(\beta_0 t)} + 1 - \frac{\sum_{t=1}^n t\Phi(\beta t)}{\sum_{t=1}^n t\Phi(\beta_0 t)} \to^p 1 \pm 1$$

for any β because in this case

$$V_n = \sum_{t=1}^n t^2\Phi(\beta_0 t)(1 - \Phi(\beta_0 t)) \simeq \sum_{t=1}^n t^2 e^{-\lambda t^2}, \quad \text{some } \lambda > 0,$$

is finite even in the limit. Therefore,

$$\text{var}\,[G_n(\beta)] = \frac{\sum_{t=1}^n t^2\Phi(\beta_0 t)(1 - \Phi(\beta_0 t))}{\left[\sum_{t=1}^n t\Phi(\beta_0 t)\right]^2} \to 0.$$

In conclusion, this normalization does not work because the limit of the mean function does not depend on β. In fact, no normalization works. For any increasing δ_n, we will have $\frac{1}{\delta_n}M_n(\beta) \to 0$ for all β or $\frac{1}{\delta_n}M_n(\beta) \to \infty$. This argument applies to other estimators, including the maximum likelihood estimator.

■

4. This is a pretty standard question. We suppose that the regressors are stationary. In the first part you have to get an unbiased variance estimator

$$\widehat{m}(x_i,\theta) = \frac{1}{n-1}\sum_{j=1}^{ns}(M(x_i,\theta,\eta_j) - \overline{M}(x_i,\theta))^2, \quad \overline{M}(x_i,\theta) = \frac{1}{n}\sum_{j=1}^{ns} M(x_i,\theta,\eta_j),$$

where η_j are i.i.d. P. The problem here is that

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{m}(x_i,\theta))^3 h(x_i) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta) - (\widehat{m}(x_i,\theta) - m(x_i,\theta)))^3 h(x_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta))^3 h(x_i) \\
&\quad -3\frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta))^2(\widehat{m}(x_i,\theta) - m(x_i,\theta))h(x_i) \\
&\quad +3\frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta))(\widehat{m}(x_i,\theta) - m(x_i,\theta))^2 h(x_i) \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}(\widehat{m}(x_i,\theta) - m(x_i,\theta))^3 h(x_i).
\end{aligned}$$

Because $\widehat{m}(x_i,\theta)$ is an unbiased simulator, the term

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta))^2(\widehat{m}(x_i,\theta) - m(x_i,\theta))h(x_i)$$

is of mean zero and can be ignored. Unfortunately, $E(\widehat{m}(x_i,\theta) - m(x_i,\theta))^j \neq 0$ for $j = 2,3$. We must estimate both these quantities unbiasedly by $B_{nij}(\theta)$ say and compute instead the first order condition

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{m}(x_i,\theta))^3 h(x_i) - 3\frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i,\theta))B_{ni2}(\theta)h(x_i) + \frac{1}{n}\sum_{i=1}^{n}B_{ni3}(\theta)h(x_i),$$

which is asymptotically unbiased and will give consistent estimates of the parameter estimates. The only remaining question is how to get unbiased estimates of $E(\widehat{m}(x_i,\theta) - m(x_i,\theta))^j$. This is just an extension of the usual idea.

■