Problem Set IV

Econometrics I - FGV EPGE
Instructor: Raul Guarini Riva
TA: Taric Latif Padovani

Problem 1 – (points: 2)

In the next questions, you will be asked to derive CLTs for the sample mean under different assumptions. You should decide what type of theorem to use in each case and *impose the assumptions you think you need*. You have less assumptions than you need on purpose. Justify your choices.

a) Suppose X_t is an α-mixing weakly stationary process with $\mathbb{E}[X_t] = 0$ and covariance function $\gamma(h)$. Let $Y_t = X_t + 0.5_{t-1}$. Derive a CLT for $\frac{1}{T}\sum_{t=2}^{T} Y_t$.

b) Suppose that X_t is an i.i.d. scalar sequence with zero mean and finite fourth moment. If $Y_t \equiv X_t X_{t-11}$, derive a CLT for $\frac{1}{T}\sum_{t=2}^{T} Y_t$

Problem 2 – (points: 1)

During the derivation of the Augmented Dickey-Fuller test, we used a trick and invoked a result without proving it. You will prove it now. Consider p coefficients from an AR(p) process given by $\phi_1, \phi_2, \ldots, \phi_p$. Consider the following $p \times p$ matrix $\boldsymbol{B}$:

$$\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix}$$

Let $\boldsymbol{y}_{t-1} = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$ be a $p \times 1$ vector. Show that:

$$\boldsymbol{B}\boldsymbol{y}_{t-1} = \begin{bmatrix} y_{t-1} \\ \Delta y_{t-1} \\ \Delta y_{t-2} \\ \vdots \\ \Delta y_{t-p+1} \end{bmatrix}$$

Furthermore, let $\phi = (\phi_1, \phi_2, \ldots, \phi_p)'$ and assume that there exist scalars $\rho \in \mathbb{R}$ and a vector $\beta = (\beta_1, \beta_2, \ldots, \beta_{p-1})' \in \mathbb{R}^{p-1}$ such that:

$$\phi = \boldsymbol{B}' \begin{bmatrix} \rho \\ \beta \end{bmatrix}$$

Show that this transformation implies:

(a) $\rho = \phi_1 + \phi_2 + \cdots + \phi_p = \sum_{j=1}^{p} \phi_j$;

(b) An AR(p) process can be equivalently written as:

$$y_t = \rho y_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \cdots + \beta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

where ε_t is the innovation term.

Problem 3 – (points: 1)

Consider an AR-DL(1,2) model where y_t depends on its own lag and on current and lagged values of an exogenous variable x_t:

$$y_t = \alpha + \rho y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t$$

where $|\rho| < 1$ ensures stationarity, and $\mathbb{E}[\varepsilon_t | y_{t-1}, y_{t-2}, \ldots, x_t, x_{t-1}, \ldots] = 0$.

(a) **Impact Multiplier**: Suppose that at time $t = 0$, the system is in equilibrium with $y_t = \bar{y}$ for all $t \leq 0$ and $x_t = \bar{x}$ for all $t \leq 0$. At time $t = 1$, there is a one-time shock such that $x_1 = \bar{x} + 1$ (i.e., x increases by one unit), and then $x_t = \bar{x}$ for all $t \geq 2$.

Compute the *impact multiplier*, defined as the immediate effect on y_1:

$$\frac{\partial y_1}{\partial x_1} = ?$$

(b) **Dynamic Multipliers**: Continuing from part (a), compute the effect of this shock on y_t for $t = 2, 3, 4$. Show that:

$$\begin{aligned} y_2 - \bar{y} &= \rho\beta_0 + \beta_1 \\ y_3 - \bar{y} &= \rho^2\beta_0 + \rho\beta_1 + \beta_2 \\ y_4 - \bar{y} &= \rho^3\beta_0 + \rho^2\beta_1 + \rho\beta_2 \end{aligned}$$

What is the general pattern here?

(c) **Long-Run Multiplier**: The *long-run multiplier* is defined as the cumulative effect of a permanent one-unit increase in x on the long-run equilibrium value of y.

Suppose instead that starting at $t = 1$, x increases permanently from $\bar{x}$ to $\bar{x} + 1$ (i.e., $x_t = \bar{x} + 1$ for all $t \geq 1$). Let y^* denote the new long-run equilibrium value of y. Show that:

$$y^* = \frac{\alpha + (\beta_0 + \beta_1 + \beta_2)(\bar{x} + 1)}{1 - \rho}$$

and therefore the long-run multiplier is:

$$\frac{dy^*}{dx} = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \rho}$$

Hint: In the new steady state, $y_t = y_{t-1} = y^*$ and $x_t = x_{t-1} = x_{t-2} = \bar{x} + 1$.

(d) **General AR-DL(1,q) Case**: Consider the more general AR-DL(1,q) model:

$$y_t = \alpha + \rho y_{t-1} + \sum_{j=0}^{q} \beta_j x_{t-j} + \varepsilon_t$$

Using the same reasoning as in part (c), show that the long-run multiplier is:

$$\text{LRM} = \frac{\sum_{j=0}^{q} \beta_j}{1 - \rho}$$

Interpret this result: why does the autoregressive coefficient ρ amplify (if $\rho > 0$) or dampen (if $\rho < 0$) the long-run effect relative to the sum of the contemporaneous and lagged coefficients on x?

Problem 4 – (points: 2)

Take the linear model $Y_t = X_t'\beta + e_t$ with $\mathbb{E}[Z_t e_t] = 0$ for some variable $Z_t \in \mathbb{R}^\ell$, where $Y_t \in \mathbb{R}$ and $X_t \in \mathbb{R}^k$. Consider the GMM estimator $\hat{\beta}$ of β and assume $\ell \geq k$. Let $J = T\hat{m}_T(\hat{\beta})'\hat{\boldsymbol{S}}^{-1}\hat{m}_T(\hat{\beta})$ be the J-statistic for the test regarding overidentifying restrictions. $\boldsymbol{S}$ is the asymptotic variance of the moment conditions and $\hat{\boldsymbol{S}}$ is a consistent estimator.

We also let $\boldsymbol{X}_{T\times k} \equiv (X_1', \ldots, X_T')'$ and $\boldsymbol{Z}_{T\times \ell} \equiv (Z_1', \ldots, Z_T')'$. Finally, $\boldsymbol{I}_\ell$ is the $\ell \times \ell$ identity matrix. We will now show that $J \xrightarrow{d} \chi^2_{\ell-k}$ as $T \to \infty$ by demonstrating the following items:

(a) Argue that we can write $\boldsymbol{S}^{-1} = \boldsymbol{C}\boldsymbol{C}'$ and $\boldsymbol{S} = \boldsymbol{C}'^{-1}\boldsymbol{C}^{-1}$ for some matrix $\boldsymbol{C}$.

(b) Show that we can write $J = T\left(\boldsymbol{C}'\hat{m}_T(\hat{\beta})\right)'\left(\boldsymbol{C}'\hat{\boldsymbol{S}}\boldsymbol{C}\right)^{-1}\boldsymbol{C}'\hat{m}_T(\hat{\beta})$.

(c) Show that $\boldsymbol{C}'\hat{m}_T(\hat{\beta}) = \boldsymbol{A}_T\boldsymbol{C}'\hat{m}_T(\beta)$ where $\hat{m}_T(\beta) = \frac{1}{T}\sum_{t=1}^{T} Z_t e_t$ and

$$\boldsymbol{A}_T = \boldsymbol{I}_\ell - \boldsymbol{C}'\left(\frac{1}{T}\boldsymbol{Z}'\boldsymbol{X}\right)\left[\left(\frac{1}{T}\boldsymbol{X}'\boldsymbol{Z}\right)\hat{\boldsymbol{S}}^{-1}\left(\frac{1}{T}\boldsymbol{Z}'\boldsymbol{X}\right)\right]^{-1}\left(\frac{1}{T}\boldsymbol{X}'\boldsymbol{Z}\right)\hat{\boldsymbol{S}}^{-1}\boldsymbol{C}'^{-1}.$$

(d) Show that $\boldsymbol{A}_T \xrightarrow{p} \boldsymbol{I}_\ell - \boldsymbol{R}(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'$ where $\boldsymbol{R} = \boldsymbol{C}'\mathbb{E}[Z_t X_t']$.

(e) Show that $T^{1/2}\boldsymbol{C}'\hat{m}_T(\beta) \xrightarrow{d} u \sim N(0, \boldsymbol{I}_\ell)$.

(f) Show that $J \xrightarrow{d} u'\left(\boldsymbol{I}_\ell - \boldsymbol{R}(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'\right)u$.

(g) Show that $u'\left(\boldsymbol{I}_\ell - \boldsymbol{R}(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'\right)u \sim \chi^2_{\ell-k}$.

Hint: $\boldsymbol{I}_\ell - \boldsymbol{R}(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'$ is a projection matrix. What do we know about these matrices?

Problem 5 – (points: 2)

You want to estimate $\mu = \mathbb{E}[Y_i]$ under the assumption that $\mathbb{E}[X_i] = 0$, where Y_i and X_i are scalars and observed from a random sample $\{(Y_i, X_i)\}_{i=1}^n$. Find an efficient GMM estimator for μ. Why the information about X_i might help to improve the estimation of μ?

Problem 6 – (points: 2)

In this question, we will explore the asymptotic distribution of the GMM estimator under model misspecification.

The observed data is $\{Y_i, X_i, Z_i\} \in \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^\ell$, $k > 1$ and $\ell > k > 1$, $i = 1, \ldots, n$. The model assumed by the econometrician is $Y = X'\beta + e$ with $\mathbb{E}[Ze] = 0$.

(a) Given a weight matrix $W > 0$ write down the GMM estimator $\hat{\beta}$ for β.

(b) Suppose the model is misspecified. Specifically, assume that for some $\delta \neq 0$,

$$e = \frac{\delta}{\sqrt{n}} + u$$
$$\mathbb{E}[u \mid Z] = 0$$

with $\mu_Z = \mathbb{E}[Z] \neq 0$. Show that (13.32) implies that $\mathbb{E}[Ze] \neq 0$.

(c) Express $\sqrt{n}(\hat{\beta} - \beta)$ as a function of W, n, δ, and the variables (X_i, Z_i, u_i).

(d) Find the asymptotic distribution of $\sqrt{n}(\hat{\beta} - \beta)$ in this case and show that it has an asymptotic bias.

(e) Is this misspecification bias eliminated if we use the optimal weight matrix? Justify your answer.

(f) Will this misspecifcation affect consistency?