PROBLEM SET # 1

ECONOMICS 240B -SPRING 2007

Almost all of the answers are the work of Tomas Rau, who was the GSI in 2005. Tomas, I can't thank you enough!

13.5 a) We want to show $E_N[U] = N^{-1}\sum_{n=1}^N U_n \xrightarrow{p} E[U] = \mu$, where U_n are independent but not identically distributed with $E[U_n] = \mu_n$ and $Var[U_n] = \sigma_n^2$. Let,

$$\bar{\mu}_N = N^{-1}\sum_{n=1}^{N}\mu_n \qquad \text{and} \qquad \bar{\sigma}_N^2 = N^{-1}\sum_{n=1}^{N}\sigma_n^2$$

and suppose that, $\lim_{N\to\infty}\bar{\mu}_N = \mu$ and $\lim_{N\to\infty} N^{-1}\bar{\sigma}_N^2 = 0$

First, Chebychev's inequality implies that

$$\begin{aligned}
&P(|E_N(U) - \mu| > \epsilon) = \\
&= P(|(E_N(U) - \bar{\mu}_N) + (\bar{\mu}_N - \mu)|) \\
&< E[(E_N(U) - \bar{\mu}_N) + (\bar{\mu}_N - \mu)]^2/\epsilon^2 \\
&= E[(E_N(U) - \bar{\mu}_N)^2] + 2(E_N(U) - \bar{\mu}_N)(\bar{\mu}_N - \mu) + (\bar{\mu}_N - \mu)^2]/\epsilon^2
\end{aligned}$$

Since $E[E_N(U)] = E[N^{-1}\sum_{n=1}^N U_n] = N^{-1}\sum_{n=1}^N \mu_n = \bar{\mu}_N$ and $Var[E_N(U)] = N^{-2}\sum_{n=1}^N \bar{\sigma}_N^2 = N^{-1}\bar{\sigma}_N^2$, we have that

$$P(|E_N(U) - \mu| > \epsilon) < \bar{\sigma}_N^2/\epsilon^2 N + (\bar{\mu}_N - \mu)^2/\epsilon^2 \longrightarrow 0 \quad \text{as} \quad N \longrightarrow \infty$$

Then, $E_N[U] \xrightarrow{p} \mu$ by assumptions above stated

b) Now we have $\mathbb{E}(U_n) = \mu$ and $Var(U_n) = \sigma^2$ but instead of independence we have that covariances equal to zero. It is easy to check that

$$P(|E_N(U) - \mu| > \epsilon) < E[(E_N(U) - \mu)^2]/\epsilon^2 = Var[E_N(U)]/\epsilon^2$$

since $E[E_N(U)] = E[N^{-1}\sum_{n=1}^N U_n] = N^{-1}\sum_{n=1}^N \mu = \mu$. Now, it is easy to check that $Var[E_N(U)] = \sigma^2/N$. So,

$$P(|E_N(U) - \mu| > \epsilon) \longrightarrow 0 \quad \text{as} \quad N \longrightarrow \infty$$

Then, $E_N[U] \xrightarrow{p} \mu$

Date: Feb 09, 2007.

c) In this case, $\mathbb{C}(U_n, U_m) = \sigma^2\rho^{|n-m|}$ we have that,

$$\begin{aligned}
\mathbb{V}(E_N[U]) &= \tfrac{\sigma^2}{N} + \tfrac{1}{N^2}\sum_{n=1}^{N}\sum_{\substack{m=1\\ m\neq n}}^{N} \mathbb{C}(U_n, U_m)\\
&= \tfrac{\sigma^2}{N} + \tfrac{2}{N^2}\sum_{n=1}^{N}\sum_{m=n+1}^{N} \mathbb{C}(U_n, U_m)\\
&= \tfrac{\sigma^2}{N} + \tfrac{2\sigma^2}{N^2}\sum_{n=1}^{N}\sum_{m=n+1}^{N} \rho^{m-n} = \tfrac{\sigma^2}{N} + \tfrac{2\sigma^2}{N^2}\sum_{n=1}^{N}\tfrac{\rho(1-\rho^{N-n})}{1-\rho}\\
&= \tfrac{\sigma^2}{N} + \tfrac{2\sigma^2}{N^2}\left[\tfrac{(N-1)\rho}{1-\rho} - \left(\tfrac{\rho}{1-\rho}\right)^2(1-\rho^{N-1})\right]
\end{aligned}$$

$$P(|E_N(U) - \mu| > \epsilon) < Var[E_N(U)]/\epsilon^2 = \frac{\sigma^2}{N\epsilon^2} + \frac{2\sigma^2}{N^2\epsilon^2}\left[\frac{(N-1)\rho}{1-\rho} - \left(\frac{\rho}{1-\rho}\right)^2(1-\rho^{N-1})\right]$$

and $Var[E_N(U)]$ goes to zero as N goes to ∞, hence $E_N[U] \xrightarrow{p} \mu$.

13.6 Let $Z_N \xrightarrow{d} Z$ where the distribution for Z_N is a mixture of two distributions with weights that depend on N: let $F_{Z_N}(z) = \lambda(N)F_Z(z) + (1-\lambda(N))F_Y(z)$ where $0 < \lambda(N) < 1$, $F_Z(z)$ is the c.d.f. of Z, and $FY(z)$ is the c.d.f. of another random variable Y. For simplicity suppose that $F_Z(z)$ and $F_W(z)$ are continuous. If $\lambda(N) \longrightarrow 1$ as $N \longrightarrow \infty$ then $Z_N \xrightarrow{d} Z$ according to definition of convergence in distribution (Ruud's book pg. 259).

Now, is easy to find one example looking at the mean of Z_N:

$$\begin{aligned}
E[Z_N] &= \int_{-\infty}^{\infty} z dF_{Z_N}(z)\\
&= \lambda(N)\int_{-\infty}^{\infty} z dF_Z(z) + (1-\lambda(N))\int_{-\infty}^{\infty} z dF_Y(z)\\
&= \lambda(N)E[Z] + (1-\lambda(N))E[Y]
\end{aligned}$$

Now, if the distribution of Y doesn't have a mean (eg. cauchy distribution), then for every N, we have that $E[Z_N]$ does not exist either and $E[Z_N]$ does not converge to $E[Z]$.

13.9 The paradox is that we treat s^2 as if it were constant (σ_0^2) in inferences about $\hat{\beta}$ based upon asymptotic approximations ($s^2 \xrightarrow{p} \sigma_0^2$). On the other hand, we know that we can approximate the distribution of s^2 with a normal distribution (See exercise 13.8).

Note, however, that the approximate distribution of s^2 is normal with mean σ_0^2 and variance $(\mu_4 - \sigma_0^2)/N$ where μ_4 is the fourth moment. For large samples, the variance of s^2 will be small. The equivalence above reflects Lemma 13.3 (Slutsky, p. 261): multiplying a sequence that converges in distribution by a sequence that converges in probability to a constant is asymptotically equivalent to multiplying the first sequence by the limiting constant. If we were examining the distribution of a statistic depending only on $\sqrt{N}(\hat{\beta} - \beta_0)$ and $\sqrt{N}(s^2 - \sigma_0^2)$, matters would be different. $\sqrt{N}(s^2 - \sigma_0^2)$ does not converge in probability. Yet there is no motivation for studying such statistics. For inference about β_0 or σ_0^2, we require statistics that are pivotal and such statistics are not pivotal.

13.11 Definition 23 says that $\hat{\theta}_n \xrightarrow{p} \theta_0$ for all possible θ_0 then $\hat{\theta}_n$ is a consistent estimator of θ_0. Note that convergence in probability requires that the distribution of $\hat{\theta}_n$ collapses to that of a constant as $N \longrightarrow \infty$. Therefore, a consistent estimator can be biased. Indeed, a consistent estimator need not have a mean at all.

On the other hand, an unbiased estimator will be inconsistent if its distribution does not collapse to the distribution of a constant. An asymptotically unbiased estimator can be inconsistent for the same reason. A vanishing bias does not imply convergence in distribution, let alone convergence in probability.

Examples: Consider an i.i.d. sample $\{X_i\}_{i=1}^N$ with $\hat{\mu}_N = E_N[X]$ and finite variance. Hence, $\hat{\mu}_N$ is unbiased and consistent for the population mean (μ). Now, consider $\hat{\lambda}_N = \hat{\mu}_N + 1/N$, this is a consistent estimator of the population mean but is biased. Now, consider the following estimator $\hat{\alpha}_N = X_1$, this is an unbiased estimator of μ but is not consistent. Finally, consider $\hat{\theta}_N = X_1 + 1/N$, this is an asymptotically unbiased estimator but inconsistent.

13.12 Note,

$P(|U_n| > \epsilon) = P(U_n < -\epsilon) + P(U_n > \epsilon) = P(\frac{\sqrt{N}U_n}{\sigma} < \frac{-\sqrt{N}\epsilon}{\sigma}) + P(\frac{\sqrt{N}U_n}{\sigma} > \frac{\sqrt{N}\epsilon}{\sigma})$

Note that for a small arbitrarily $\delta > 0$ and N large enough,

$$\begin{aligned} P(\tfrac{\sqrt{N}U_n}{\sigma} < \tfrac{-\sqrt{N}\epsilon}{\sigma}) &\le \Phi(\tfrac{-\sqrt{N}\epsilon}{\sigma}) + \delta \quad \text{and} \\ P(\tfrac{\sqrt{N}U_n}{\sigma} > \tfrac{\sqrt{N}\epsilon}{\sigma}) &\le 1 - \Phi(\tfrac{\sqrt{N}\epsilon}{\sigma}) + \delta \end{aligned}$$

by $\sqrt{N}U_n \xrightarrow{d} (0, \sigma^2)$. Hence,

$P(|U_n| > \epsilon) \le 1 - \Phi(\frac{\sqrt{N}\epsilon}{\sigma}) + \delta + \Phi(\frac{-\sqrt{N}\epsilon}{\sigma}) + \delta \quad \text{as} \quad N \longrightarrow \infty$

Since,$\Phi(\frac{\sqrt{N}\epsilon}{\sigma}) \longrightarrow 1$ and $\Phi(\frac{-\sqrt{N}\epsilon}{\sigma}) \longrightarrow 0$, we can make $P(|U_n| > \epsilon)$ arbitrarily small as $N \longrightarrow \infty$, so $U_n \xrightarrow{p} 0$

25.11 consider the MA(1) process $\epsilon_t = u_t + \theta u_t$ where $\{u_t\}$ is i.i.d. zero mean and $\sigma_u^2 > 0$. We know that the correlation between ϵ_t and ϵ_{t-1} is given by,

$$\rho = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta\sigma_u^2}{\sigma_u^2(1+\theta^2)} = \frac{\theta}{1+\theta^2}$$

Now, we want to prove that $|\theta/(1+\theta^2)| \le 1/2$. Let's proceed by contradiction: suppose not, i.e. $\exists \theta \in \Re$ s.t. $\|\theta/(1+\theta^2)| > 1/2$. This is equivalent to $\theta/(1+\theta^2) > 1/2$ and $\theta/(1+\theta^2) < -1/2$.

First, $\theta/(1+\theta^2) > 1/2 \Longrightarrow (1-\theta^2)^2 < 0$ which is a contradiction since $\theta \in \Re$. Second, $\theta/(1+\theta^2) < -1/2 \Longrightarrow (1+\theta^2)^2 < 0$ which is a contradiction.

25.13 Consider two independent AR(1) processes: $y_t = \alpha y_{t-1} + \epsilon_t$ and $x_t = \beta x_{t-1} + u_t$ where $\{\epsilon_t\}$ and $\{u_t\}$ are *i.i.d.*, zero-mean with variances σ_ϵ^2 and σ_u^2 respectively. We want to show that $z_t = y_t + x_t$ follows an $ARMA(2,1)$ process. First, note that we can rewrite the AR processes using the *Lag* operator, so

$$\begin{array}{ll} (1-\alpha L)y_t & = \epsilon_t \\ (1-\beta L)x_t & = u_t \end{array}$$

multiplying the first equation by $(1-\beta L)$ and the second by $(1-\alpha L)$ we have,

$$\begin{array}{ll} (1-\alpha L)(1-\beta L)y_t & = (1-\beta L)\epsilon_t \\ (1-\alpha L)(1-\beta L)x_t & = (1-\alpha L)u_t \end{array}$$

note that $(1-\alpha L)(1-\beta L)$ is a 2nd order polynomial of lags. Now, adding both equaitons we get,

$$\begin{array}{ll} (1-\alpha L)(1-\beta L)(y_t + x_t) & = (1-\beta L)\epsilon_t + (1-\alpha L)u_t \\ (1-\alpha L)(1-\beta L)z_t & = c_t \end{array}$$

where $c_t = (1-\beta L)\epsilon_t + (1-\alpha L)u_t$. Hence we have the AR(2) part in the left-hand-side but we need to show that c_t follows a MA(1), i.e. we need to show that it is weak stationary and that $cov(c_t, c_{t-s}) = 0$ for $s > 1$. It can be easily seen that,

$$\mathbb{E}(c_t, c_{t-s}) = \begin{cases} (1+\beta^2)\sigma_\epsilon^2 + (1+\alpha^2)\sigma_u^2 & s = 0 \\ \beta\sigma_\epsilon^2 + \alpha\sigma_u^2 & s = 1 \\ 0 & s > 1 \end{cases}$$

hence c_t admits a MA(1) representation, therefore there exist an *i.i.d.* process v_t and $\phi \in \Re$ such that $c_t = (1-\phi L)v_t$. Finally,

$$(1-\alpha L)(1-\beta L)(x_t + y_t) = (1-\phi L)v_t$$

follows an ARMA(2,1) process.

Additional Problems

(1) We know that

$$\sqrt{n}(\widehat{\beta} - \beta) \longrightarrow_d N(0, \sigma^2 E(x_i x_i')^{-1})$$

We'll use the Delta Method:

$$\begin{aligned} g(y,z) &= \frac{y}{z} \\ &\Longrightarrow \frac{\partial g(y,z)}{\partial (y,z)'} = \left(\frac{1}{z}, -\frac{y}{z^2}\right) \\ &\Longrightarrow G_0 = \frac{\partial g(\beta_1,\beta_2)}{\partial (y,z)'} = \left(\frac{1}{\beta_2}, -\frac{\beta_1}{\beta_2^2}\right) = \frac{1}{\beta_2}(1,-\gamma) \end{aligned}$$

so

$$\sqrt{n}\left(\frac{\widehat{\beta}_1}{\widehat{\beta}_2} - \gamma\right) \longrightarrow_d N\left(0, \frac{\sigma^2}{\beta_2^2}(1,-\gamma)\,E(x_i x_i')^{-1}(1,-\gamma)'\right)$$

But we have 3 results:

$$\begin{aligned} 1)\sqrt{(1,-\widehat{\gamma})\,(\frac{X'X}{n})^{-1}\,(1,-\widehat{\gamma})'} &\longrightarrow {}_p\sqrt{(1,-\gamma)\,E(x_i x_i')^{-1}\,(1,-\gamma)'} \\ 2)\quad \widehat{\beta}_2 &\longrightarrow {}_p\beta_2 \\ 3)\quad s &\longrightarrow {}_p\sigma \end{aligned}$$

So we use Slustky to conclude that

$$\frac{\widehat{\beta}_2\sqrt{n}\left(\frac{\widehat{\beta}_1}{\widehat{\beta}_2} - \gamma\right)}{s\sqrt{(1,-\widehat{\gamma})\,(\frac{X'X}{n})^{-1}\,(1,-\widehat{\gamma})'}} \longrightarrow_d N(0,1)$$

So, an approximate confidence interval when $\widehat{\gamma} = \frac{\widehat{\beta}_1}{\widehat{\beta}_2} = -1$ is:

$$\left|\frac{-\frac{1}{4}(-1-\gamma)}{\frac{1}{\sqrt{10}}\sqrt{\begin{bmatrix}1 & -\widehat{\gamma}\end{bmatrix}\begin{bmatrix}\frac{1}{20} & -\frac{1}{10} \\ -\frac{1}{10} & \frac{2}{5}\end{bmatrix}\begin{bmatrix}1 \\ -\widehat{\gamma}\end{bmatrix}}}\right| \leq 1.96$$

$$\begin{aligned} &\Longleftrightarrow \left|\frac{\frac{1}{4}(1+\gamma)}{\frac{1}{\sqrt{10}}\sqrt{\frac{1}{20} + \frac{2}{5}\widehat{\gamma} + \frac{2}{5}\widehat{\gamma}^2}}\right| \leq 1.96 \\ &\Longleftrightarrow \left|\frac{\frac{1}{4}(1+\gamma)}{\frac{1}{\sqrt{10}}\sqrt{\frac{1}{4}}}\right| \leq 1.96 \\ &\Longleftrightarrow -2.24 \leq \gamma \leq 0.24 \end{aligned}$$

So, $\gamma_0 = 0$ is in this interval.

(2) a) when $\mu = 0$ we have that $\overline{X} \xrightarrow{p} 0$ by LLN, $\sqrt{n}\overline{X} \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ by CLT, and $s^2 \xrightarrow{p} \sigma^2$ by LLN, therefore we use Slutsky theorem having

$$\sqrt{n}T_n = \frac{\sqrt{n}\overline{X}}{s} \xrightarrow{d} \frac{1}{\sigma}\mathcal{N}(0, \sigma^2) = \mathcal{N}(0, 1)$$

b) Consider the vector

$$W_n = \begin{pmatrix} \overline{X} \\ s^2 \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n} \begin{pmatrix} X_i \\ (X_i - \mu)^2 \end{pmatrix} - \begin{pmatrix} 0 \\ (\overline{X} - \mu)^2 \end{pmatrix}$$

Due to the LLN, the last term goes in probability to the zero vector, and the first term, and thus the whole W_n, converges in probability to

$$\text{plim} W_n = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$$

We can apply multivariate CLT to W_n, hence

$\sqrt{n}\,(W_n - \text{plim}W_n) \xrightarrow{d} \mathcal{N}(0, V)$

where $V = \mathbb{V}\begin{pmatrix} X_i \\ (X_i - \mu)^2 \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \tau - \sigma^4 \end{pmatrix}$

and now we can use the so-called delta method with function:

$$g\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \equiv \frac{t_1}{\sqrt{t_2}} \Rightarrow G \equiv g'\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{1}{\sqrt{t_2}}\begin{pmatrix} 1 & -\frac{t_1}{2t_2} \end{pmatrix}$$

hence,

$$\sqrt{n}\,(T_n - T_0) \xrightarrow{d} \mathcal{N}(0, GVG')$$

$$\sqrt{n}\,(T_n - T_0) \xrightarrow{d} \mathcal{N}\left(0, 1 + \frac{\mu^2(\tau - \sigma^4)}{4\sigma^6}\right)$$

c) Similarly, consider the vector:

$$W_n = \begin{pmatrix} \overline{X} \\ \hat{\sigma}^2 \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n} \begin{pmatrix} X_i \\ X_i^2 \end{pmatrix}$$

Due to LLN, W_n converge in probability to

$$\text{plim} W_n = \begin{pmatrix} \mu \\ \mu^2 + \sigma^2 \end{pmatrix}$$

Next, $\sqrt{n}(W_n - \text{plim}W_n) \xrightarrow{d} \mathcal{N}(0, V)$ where $V = \mathbb{V}\begin{pmatrix} X_i \\ X_i^2 \end{pmatrix}$. Lets calculate V. First, $\mathbb{V}(X_i) = \sigma^2$ and $\mathbb{V}(X_i^2) = \mathbb{E}[(X_i^2 - \mu^2 - \sigma^2)^2] = \tau + 4\mu^2\sigma^2 - \sigma^4$. Second, $\mathbb{C}[X_i, X_i^2] = \mathbb{E}[(X_i - \mu)(X_i^2 - \mu^2 - \sigma^2)] = 2\mu\sigma$. Therefore,

$$\sqrt{n}\,(W_n - \text{plim}W_n) \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 2\mu\sigma^2 \\ 2\mu\sigma^2 & \tau + 4\mu^2\sigma^2 - \sigma^4 \end{pmatrix}\right)$$

Now we can use the so-called delta method with $g\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \equiv \frac{t_1}{\sqrt{t_2}}$ to get

$$\sqrt{n}\,(R_n - R_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{\mu^2\tau - \mu^2\sigma^4 + 4\sigma^6}{4(\mu^2 + \sigma^2)^3}\right)$$

The answer reduces to that of Part (b) iff $\mu = 0$. Under this condition, T_n and R_n are asymptotically equivalent.

E-mail address: carolina@econ.berkeley.edu

Answer Key