# Econometric Problem Set 1

*Gigi*

*September 19, 2018*

## Question 1

**Prove the Chebyshev Inequality, and then the WLLN.**

The Chebyshev Inequality is

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

**Proof**: Let $X$ be a random variable with mean $\mu$ and variance $\sigma^2$. Let $\epsilon > 0$ be any postive real number. I will show for the **discrete case**, but the continuous case is identical to this proof. Let $f(x)$ be the distribution function of $X$. Then the probability that $X$ deviates from $\mu$ by at least $\epsilon$ is

$$P(|X - \mu| \geq \epsilon) = \sum_{|x-\mu| \leq \epsilon} f(x) \tag{1}$$

Let us explicitly define the variance of $X$:

$$\begin{aligned}
\sigma^2 &= \sum_x (x-\mu)^2 f(x) && (2)\\
&\geq \sum_{|x-\mu| \leq \epsilon} (x-\mu)^2 f(x) && (3)\\
&\geq \sum_{|x-\mu| \geq \epsilon} \epsilon^2 f(x) = \epsilon^2 \sum_{|x-\mu| \geq \epsilon} f(x) && (4)\\
&= \epsilon^2 \, P(|X-\mu| \geq \epsilon). && (5)
\end{aligned}$$

From (2) and (5), we now have

$$\sigma^2 \geq \epsilon^2 \, P(|X - \mu \geq \epsilon),$$

which by simplifying is the Chebyshev Inequality

$$\frac{\sigma^2}{\epsilon^2} \geq P(|X - \mu \geq \epsilon).$$

An assumption to this inequality is that $\sigma^2 \neq \infty$.

The **Weak Law of Large Numbers** says that $X_1, X_2, \ldots, X_n$ are an independent trials process, with finite expected value $\mu = E(X_i)$ and finite variance $\sigma^2 = V(X_i)$. Let $S_n = X_1, X_2, \ldots, X_n$. Then for any $\epsilon > 0$,

$$P(|\frac{S_n}{n} - \mu) \leq \epsilon) \to 0$$

as n $\to \infty$, or,

$$P(|\frac{S_n}{n} - \mu) < \epsilon) \to 1$$

as n $\to \infty$

**Proof**: We will use Chebyshev's Inequality (which was proved above) as this assumes mutual independence. Let $S_n = X_1, X_2, \ldots, X_n$ be a sequence of mutually independent random variables, with $\mu = E(X_i)$ and variance $Var(S_i) = \sigma^2$. Since $X_1, X_2, \ldots, X_n$ are independently distributed: for this sequence,

$$Var(S_i) = n\sigma^2,$$

and

$$Var(\frac{S_n}{n}) = \frac{\sigma^2}{n}.$$

Which gives

$$E(\frac{S_n}{n}) = \mu.$$

By Chebyshev's Inequality, for any $\epsilon > 0$,

$$P(|\frac{S_n}{n} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}.$$

So, for any constant $\epsilon$,

$$P(|\frac{S_n}{n} - \mu) \leq \epsilon) \to 0$$

as n $\to \infty$ or

$$P(|\frac{S_n}{n} - \mu) < \epsilon) \to 1$$

as n $\to \infty$

□

---

## Question 2

**Show that if X and Y are continuous independent random variables, $E[XY] = E[X]E[Y]$.**

**Proof**: Let $X$ and $Y$ be two independent continuous random variables. Independence implies the **covariance** between $X$ and $Y$ is 0: $\sigma_{X,Y} = 0$.

$$\begin{aligned}
\sigma_{X,Y} &= E[(X - E(X))(Y - E(Y))] && (1)\\
&= E[XY - XE(Y) - YE(X) + E(X)E(Y)] && (2)\\
&= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) && (3)\\
&= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) && (4)\\
&= E(XY) - E(X)E(Y) && (5)
\end{aligned}$$

Since $\sigma_{X,Y} = 0$,

$$\begin{aligned}
\sigma_{X,Y} &= E(XY) - E(X)E(Y) && (6)\\
0 &= E(XY) - E(X)E(Y) && (7)\\
E(XY) = E(X)E(Y), & && (8)
\end{aligned}$$

□

where (1) is the definition of covariance; (2) follows from distributing the terms inside the brackets; (3) is the communitivity property of expectation; (4) is the distributive law; (5) cancels out like terms, and (8) is our result.

---

## Question 2_a

Let $X_1, \ldots, X_n$ be a sequence of N independent and identically distributed (iid) random variables, with $E[X_i] = \mu < \infty$. Then,

$$\bar{X_N} \xrightarrow{\text{p}} \mu$$

.

## Question 3

### Linear regression model

Obtain the covariance matrix of $\mathbf{u} = (u_1, \ldots, u_T)'$ as a function of $\rho$ and $\sigma_u^2$.

**Work**: Our linear regression model is $y_t = x_t\beta + u_t$ in the presence of residual autocorrelation of the form $u_t = \rho u_{t-1} + \epsilon_t$. $E(\epsilon_t) = 0$, $Var(\epsilon_t) = \sigma_\epsilon^2$, $Var(u_t) = \sigma_u^2$ for $t = 1, \ldots, T$, and $Cov(\epsilon_s, \epsilon_t) = 0$ for $s, t = 1, \ldots, T,$, where $s \neq t$.

We have $E(\epsilon_t, \epsilon_{t+s}) = \sigma_\epsilon^2$ if $(s = 0)$, and $E(\epsilon_t, \epsilon_{t+s}) = 0$ if $(s \neq 0)$

Now,

$$\begin{aligned} u_t =& \rho u_{t-1} + \epsilon_t && (1) \\ =& \rho(u_{t-2} + \epsilon_{t-1} + \epsilon_t) && (2) \\ =& \vdots \\ =& \epsilon_t + \rho\epsilon_{t-1} + \rho^2\epsilon_{t-2} && (3) \\ =& \sum_{r=0}^{\infty} \rho^2 \epsilon_{t-r} && (4) \end{aligned}$$

Since $E(u_t) = 0$, taking the expected value of (3) $E(u_t)$, we get

$$\begin{aligned} E(u_t^2) =& E(\epsilon_t^2) + \rho^2 E(\epsilon_{t-1}^2) + \rho^4 E(\epsilon_{t-2}^2) + \ldots \\ =& \sigma_{\epsilon^2}(1 + \rho^2 + \rho^4 + \ldots) \\ E(u_t^2) =& \sigma_u^2 = \frac{\sigma_\epsilon^2}{1 - \rho^2} \end{aligned}$$

for all i.

$$\begin{aligned} E(u_t u_{t-1}) =& E[(\epsilon_t + \rho\epsilon_{t-1} + \rho^2\epsilon_{t-2} + \ldots) \times (\epsilon_{t-1} + \rho\epsilon_{t-2} + \rho^2\epsilon_{t-3} + \ldots)] \\ =& E[[\epsilon_t + \rho(\epsilon_{t-1} + \epsilon_{t-2} + \ldots)](\epsilon_{t-1} + \rho\epsilon_{t-2} + \rho^2\epsilon_{t-3} + \ldots)]] \\ =& \rho E[(\epsilon_{t-1} + \rho\epsilon_{t-2} + \ldots)^2] \\ =& \rho\sigma_u^2 \end{aligned}$$

as $E(\epsilon_t) = 0$.

Also $E(u_t u_{t-2}) = \rho^2\sigma_u^2$.

In general,

$E(u_t u_{t-s}) = \rho^s \sigma_u^2$

$$E(uu') = \Omega = \sigma_u^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}$$

Which is the covariance matrix of this linear regression.