LIMIT OF A SEQUENCE

Suppose that $\{a_1, a_2, \ldots, a_n\}$ constitutes a sequence of real numbers. If there exists a real number α such that for every real number ε however small there exists an integer $N(\varepsilon)$ such that for all $n > N(\varepsilon)$ $|a_n - \alpha| < \varepsilon$, then we say that α is the limit of the sequence $\{a_n\}$ and write $\lim_{n\to\infty} \{a_n\} = \alpha$.

LIMIT OF A FUNCTION

The function $f(x)$ has the limit A at the point x_0 if for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that $|f(x) - A| \, \varepsilon$ whenever $0 < |x - x_0| < \delta(\varepsilon)$.

CONVERGENCE IN DISTRIBUTION

Given a sequence of random variable X_n whose CDF is $F_n(x)$ and a cumulative distribution function $F_X(x)$ which corresponds to random variable X, we say that X_n converges in distribution to X and write

$X_n \xrightarrow{d} X$ if

$\lim_{n\to\infty} F_n(x) = F_X(X)$ at all points x at which $F_X(x)$ is continuous.

CONVERGENCE IN PROBABILITY

The sequence of random variables X_n converges in probability to the real number x if

$\lim_{n\to\infty} P\left[|X_n - x| > \varepsilon\right] = 0$ for each $\varepsilon > 0$.

Thus, it becomes less and less likely that the random variable $(X_n - x)$ lies outside the interval $(\pm \varepsilon)$.

The following statements are equivalent:

(a) $\lim_{n\to\infty} P\left(|X_n - x| < \varepsilon\right) = 1, \; \varepsilon > 0$.

(b) Given $\varepsilon > 0$ and $\delta > 0$ $\exists\, N(\varepsilon,\delta)$ such that

$$P\left(|X_n - x| > \varepsilon\right) < \delta \quad \text{for all } n > N(\varepsilon,\delta).$$

(c) $$P\left(|X_n - x| < \varepsilon\right) > 1-\delta \quad \text{for all } n > N.$$

That is, $P\left(|X_{N+1} - x| < \varepsilon\right) > 1-\delta$, $P\left(|X_{N-1} - x| < \varepsilon\right) > 1-\delta$ etc.

Convergence in probability is indicated by

$$X_n \xrightarrow{P} x \quad \text{or} \quad \boxed{\operatorname{plim} X_n = x}.$$

A familiar example of convergence in probability is:

$$\operatorname{plim} \bar{x}_n = \mu.$$

Remember that $E(\bar{x}_n) = \mu$ and $Var(\bar{x}_n) = \sigma^2/n$.

As the sample size (n) increases, the variance of the sampling distribution of the mean becomes smaller. Hence, the probability mass gets concentrated around the mean of the sampling distribution (μ). Hence, there is a large enough n such that $Pr\left[\bar{x}_n \in (\mu \pm \varepsilon)\right] > 1-\delta$.

MEAN SQUARED ERROR CONVERGENCE

A random variable X_n converges in the quadratic mean ~~if~~ to x if $E\left[(X_n - x)^2\right]$ exists and

$$\lim_{n\to\infty} E\left[(X_n - X)^2\right] = 0.$$

Again, the sample mean converges in the mean-squares to the population mean.

$$E\left[(\bar{X}_n - \mu)^2\right] = Var(\bar{X}_n) = \sigma^2/n$$

and $\lim_{n\to\infty} Var(\sigma^2/n) \to 0$.

ALMOST SURE CONVERGENCE:

The sequence of random variables X_n is said to converge almost surely to the real number x and is written as $X_n \xrightarrow{a.s.} x$ if

$$P\left[\lim X_n = x\right] = 1$$

This means that given ε and $\delta > 0$, there exist an integer N such that

$$Pr\left[\,|X_{N+1} - x| < \varepsilon,\ |X_{N+2} - x| < \varepsilon, \ldots\right] > 1 - \delta.$$

That is, the joint probability that these values can be made arbitrarily close to x stay arbitrarily close to 1.

CONSISTENCY

SIMPLE CONSISTENCY:

Let $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$ be a sequence of estimators of the parameter θ. Then $\hat{\theta}_n$ is a simple consistent estimator of θ if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\left[\,|\hat{\theta}_n - \theta| < \varepsilon\right] = 1.$$

This implies that $\text{plim}\,\hat{\theta}_n = \theta$.

SQUARED-ERROR CONSISTENCY:

The sequence $(\hat{\theta}_n)$ is a squared error consistent estimator of θ if $\lim_{n\to\infty} E\left[(\hat{\theta}_n - \theta)^2\right] = 0$.

STRONG CONSISTENCY:

$\hat{\theta}_n$ is a strongly consistent estimator of θ if

$$P\left[\lim_{n\to\infty} \hat{\theta}_n = \theta\right] = 1$$

Note that ~~___~~ strong consistency requires almost sure convergence

UNBIASEDNESS & CONSISTENCY

Consider a random sample $\{x_1, x_2, \ldots, x_n\}$ drawn from the population with $E(x) = \mu$. Let $x_1 = \tilde{\mu}$ be an estimator of μ. Clearly, $E(x_1) = \mu$ so that $\tilde{\mu}$ is an unbiased estimator. But x_1 does not converge in probability to μ no matter how large the sample size becomes. $\tilde{\mu}$ is not a consistent estimator.

ASYMPTOTIC UNBIASEDNESS

One often characterizes an estimator $\hat{\theta}_n$ as asymptotically unbiased if $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$. But this definition is not useful if $E(\hat{\theta}_n)$ does not exist.

Alternatively, $\hat{\theta}_n$ is an asymptotically unbiased estimator of θ if the mean of the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ is zero.

Asymptotic unbiasedness implies

$$AE(\hat{\theta}_n) = \theta.$$

Similarly, the asymptotic variance can be defined as either

$$\lim_{n\to\infty} E\left[\sqrt{n}\,(\hat{\theta}_n - \theta)\right]^2 \text{ or}$$

$$AE\left[\left\{\sqrt{n}\,(\hat{\theta}_n - \theta)\right\}^2\right].$$

The first is the limit of the variance. The other is the variance of the limiting distribution.

Note that in defining asymptotic variance we consider $\sqrt{n}(\hat{\theta}_n - \theta)$ so that variances do not tend to 0 and n increases.

ASYMPTOTIC UNBIASEDNESS & CONSISTENCY

Let $x = (x_1, x_2, \ldots, x_n)$ be a random sample drawn from $N(\theta, 1)$. Let

$$\hat{\theta}_n = \frac{1}{2} x_1 + \frac{1}{2n} \sum_{2}^{n} x_i$$

be an estimator of θ.

$$E\left[\hat{\theta}_n\right] = \frac{1}{2}\theta + \frac{n-1}{2n}\theta$$

$$= \frac{n\theta + (n-1)\theta}{2n} = \left(1 - \frac{1}{2n}\right)\theta$$

$\lim_{n \to \infty} E\left[\hat{\theta}_n\right] = \theta$. Thus $\hat{\theta}_n$ is an asymptotically unbiased estimator

But $\text{plim}\ \hat{\theta}_n = \text{plim}\left(\frac{1}{2}x_1\right) + \text{plim}\ \frac{(n-1)\bar{x}_{n-1}}{2n}$

$$= \frac{1}{2}x_1 + \frac{1}{2}\theta \neq \theta.$$

Hence, asymptotic unbiasedness does not ensure consistency.

STOCHASTIC REGRESSORS:

Consider the 2 variable regression

$$Y_i = \alpha + \beta X_i + u_i$$

and OLS estimator

$$\hat{\beta} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \beta + \frac{\sum (X_i - \bar{X})(u_i - \bar{u})}{\sum (X_i - \bar{X})^2}$$

$$= \beta + \frac{\sum x_i u_i}{\sum x_i^2} \quad \text{where } x_i \equiv X_i - \bar{X}.$$

Note that $E(\hat{\beta}) = 0$ when

$$E\left[\frac{x_1 u_1 + x_2 u_2 + \cdots + x_n u_n}{\sum x_i^2}\right] = 0.$$

When the Xs are non random, this expectation is

$$\frac{x_1}{\sum x_i^2} E(u_1) + \frac{x_2}{\sum x_2^2} E(u_2) + \cdots + \frac{x_n}{\sum x_n^2} E(u_n) = 0.$$

Even when the X_is are random so that $\frac{x_i}{\sum x_i^2}$ is also random, so long they are independent of the u_is, $E(\hat{\beta}) = \beta + \sum_i E\left(\frac{x_i}{\sum x_i^2}\right) \cdot E(u_i) = \beta$ no matter what $E\left[\frac{x_i}{\sum x_i^2}\right]$ is. because $E(u_i) = 0$

Even when X and u are not independent,

if $\text{plim} \frac{1}{n} \sum x_i u_i = 0$,

$$\text{plim}\, \hat{\beta} = \beta + \text{plim} \frac{\sum x_i u_i}{\sum x_i^2} = \beta + \frac{\text{plim} \frac{1}{n} \sum x_i u_i}{\text{plim} \frac{1}{n} \sum x_i^2}$$

$$= \beta.$$

Thus $\hat{\beta}$ would be a consistent estimator of β

Now suppose that the true model is

$$Y_i = \alpha + \beta x_i + u_i$$

but we can observe only $z_i = x_i + v_i$.

Thus
$$Y_i = \alpha + \beta (z_i - v_i) + u_i$$
$$Y_i = \alpha + \beta z_i + (u_i - \beta v_i)$$

Define $W_i = u_i - \beta v_i$

Then $Y_i = \alpha + \beta z_i + W_i$

$$\hat{\beta} = \frac{\sum (z_i - \bar{z})(Y_i - \bar{Y})}{\sum (z_i - \bar{z})^2} = \beta + \frac{\sum_i (z_i - \bar{z})(W_i - \bar{W})}{\sum (z_i - \bar{z})^2}$$

Note that
$$z_i - \bar{z} = (x_i - \bar{x}) + (v_i - \bar{v})$$
and
$$W_i - \bar{W} = (u_i - \bar{u}) - \beta (v_i - \bar{v})$$

Thus $\frac{1}{n} \sum (z_i - \bar{z})(W_i - \bar{W})$

$$= \text{Cov}(x, u) - \beta \text{Cov}(x, \beta v) + \text{Cov}(u, v) - \beta \text{Var}(v)$$

$$\frac{1}{n} \sum (z_i - \bar{z})^2 = \text{Var}(z) = \text{Var}(x) + \text{Var}(v) + 2 \text{Cov}(x, v)$$

Assume that
$$\text{plim Cov}(x, u) = \text{plim}(u, v) = \text{plim}(x, v) = 0$$

Thus
$$\text{plim } \hat{\beta} = \beta + \frac{\text{plim Cov}(z, W)}{\text{plim Var}(z)}$$

$$= \beta - \frac{\beta \sigma_v^2}{\sigma_x^2 + \sigma_v^2} = \beta \left(\frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2} \right) < \beta.$$

Assume that we have the "True" regression model

$$y = \tilde{X}\beta + u.$$

But the matrix of observed values of the explanator variables is

$$X = \tilde{X} + V. \Leftrightarrow \tilde{X} = X - V$$

Thus, $y = (X - V)\beta + u$

$$\Rightarrow y = X\beta + (u - V\beta).$$

If we look at the OLS estimator $\hat{\beta} = (X'X)^{-1}X'y$,

$$\hat{\beta} = \beta + (X'X)^{-1}X'(u - V\beta).$$

We may assume:

1) that the measurement errors V are uncorrelated in the limit with the true values $\tilde{X}$.

Thus. $\text{plim}\ \frac{\tilde{X}'V}{n} = 0$

Hence. $\text{plim}\left(\frac{1}{n}X'X\right) = \text{plim}\left(\frac{1}{n}\tilde{X}'\tilde{X}\right) + \text{plim}\left(\frac{1}{n}V'V\right)$

$$= \Sigma + \Omega$$

where $\Sigma = \text{plim}\left(\frac{1}{n}\tilde{X}'X\right)$ & $\Omega = \text{plim}\left(\frac{1}{n}V'V\right)$

(2) The measurement errors (V) are uncorrelated in the limit with the disturbance term u. Thus

$$\text{plim}\left(\frac{1}{n}V'u\right) = 0$$

But by specification $\tilde{X}$ & u are independent.

Hence $\text{plim}\left(\frac{1}{n}\tilde{X}'u\right) = 0$ also holds

Then $\text{plim}(\hat{\beta}) = \beta + \text{plim}\left(\frac{X'X}{n}\right)^{-1}\text{plim}\left(\frac{X'u}{n}\right)$

$$- \text{plim}\left(\frac{X'X}{n}\right)^{-1}\text{plim}\left(\frac{X'V}{n}\right)\beta$$

Now $\text{plim}\left(\frac{X'V}{n}\right)\beta = \text{plim}\left(\frac{\tilde{X}'V}{n} + \frac{V'V}{n}\right)\beta$

$$= \text{plim}\left(\frac{V'V}{n}\right)\beta = \Omega\beta$$

Therefore $\quad \text{plim}\ \hat{\beta} = \beta - (\Sigma + \Omega)^{-1} \Omega \beta$

Thus OLS estimators are not consistent.

The inconsistency arises out of interdependence between X and the composite error $(u - V\beta)$.

<u>Example</u>:

$$Y_i = \alpha + \beta \tilde{X}_i + u_i$$
$$X_i = \tilde{X}_i + v_i$$

Thus $\quad Y_i = \alpha + \beta X_i + (u_i - \beta v_i)$

$$\Sigma = \text{plim}\ \frac{1}{n}(\tilde{X}'\tilde{X}) = \text{plim} \begin{bmatrix} 1 & \frac{1}{n}\Sigma \tilde{X}_i \\ \frac{1}{n}\Sigma \tilde{X}_i & \frac{1}{n}\Sigma \tilde{X}_i^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \mu \\ \mu & \mu^2 + \sigma^2 \end{bmatrix}$$

Here μ & σ^2 are the mean and the variance of $\tilde{X}_i$.

$$\Omega = \text{plim}\left(\frac{1}{n} V'V\right) = \text{plim} \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{n}\Sigma v_i^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$$

$$\Sigma + \Omega = \begin{bmatrix} 1 & \mu \\ \mu & \mu^2 + \sigma^2 + \sigma_v^2 \end{bmatrix}$$

$$[\Sigma + \Omega]^{-1} = \frac{\begin{bmatrix} \mu^2 + \sigma^2 + \sigma_v^2 & -\mu \\ -\mu & 1 \end{bmatrix}}{\sigma^2 + \sigma_v^2}$$

$$[\Sigma + \Omega]^{-1} \Omega = \frac{1}{\sigma^2 + \sigma_v^2} \begin{bmatrix} 0 & -\mu \sigma_v^2 \\ 0 & +\sigma_v^2 \end{bmatrix}$$

$$\text{plim}\begin{bmatrix}\hat\alpha\\ \hat\beta\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - \frac{1}{\sigma^2+\sigma_v^2}\begin{bmatrix}0 & -\mu\sigma_v^2\\ 0 & -\sigma_v^2\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$$

$$= \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - \frac{1}{\sigma^2+\sigma_v^2}\begin{bmatrix}-\mu\sigma_v^2\beta\\ +\sigma_v^2\beta\end{bmatrix}$$

$$\text{plim } \hat\alpha = \alpha + \mu\frac{\sigma_v^2}{\sigma^2+\sigma_v^2}\beta$$

$$\text{plim } \hat\beta = \beta - \frac{\sigma_v^2}{\sigma^2+\sigma_v^2}\beta$$

$$= \beta\frac{\sigma^2}{\sigma^2+\sigma_v^2}$$

INSTRUMENTAL VARIABLE ESTIMATOR

The problem of inconsistency arises because the measured values X are correlated with the composite error $(u - V\beta)$.

Now assume that we can replace Xs by a matrix Z where the Zs are correlated with X but not correlated with u or V

~~Now~~ consider the estimator

$$\beta_{IV} = (Z'X)^{-1}Z'y$$

$$= (Z'X)^{-1}Z'(X\beta + u - V\beta)$$

$$= \beta + (Z'X)^{-1}Z'u - (Z'X)^{-1}Z'V\beta$$

$$\text{plim } \beta_{IV} = \beta + \text{plim}\left(\frac{Z'X}{n}\right)^{-1}\text{plim}\left(\frac{Z'u}{n}\right) - \text{plim}\left(\frac{Z'X}{n}\right)^{-1}\text{plim}\left(\frac{Z'V}{n}\right)\beta$$

$$= \beta + \Sigma_{xz}^{-1}0 - \Sigma_{zx}^{-1}0\,\beta = \beta$$

Consider again the regression

$$Y_i = \alpha + \beta X_i + (u_i - \beta v_i)$$

Suppose that there are n observations and n is even. Then define a variable z which takes the value 1 if X_i is above the median value and -1 if X_i is less than the median value of X. Now consider the matrix

$$Z = \begin{bmatrix} 1 & -1 \\ 1 & +1 \\ \cdot & \cdot \\ \cdot & \cdot \\ & +1 \\ 1 & -1 \end{bmatrix}$$

With this definition,

$$\begin{bmatrix} \alpha_{IV} \\ \beta_{IV} \end{bmatrix} = (Z'X)^{-1} Z'y$$

$$Z'X = \begin{bmatrix} 1 & 1 & \cdot & \cdot & 1 & 1 \\ -1 & +1 & \cdot & \cdot & +1 & -1 \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ \cdot & X_2 \\ \cdot & \\ & \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & n\bar{X} \\ 0 & \frac{n}{2}(\bar{X}_{II} - \bar{X}_I) \end{bmatrix}$$

where $\bar{X}_I$ = mean of values of X less than median

$\bar{X}_{II}$ = " " " X greater " " .

$$Z'y = \begin{bmatrix} 1 & 1 & \cdot & \cdot & 1 & 1 \\ -1 & +1 & & & +1 & -1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \\ Y_n \end{bmatrix} = \begin{bmatrix} n\bar{Y} \\ \frac{n}{2}(\bar{Y}_{II} - \bar{Y}_I) \end{bmatrix}$$

$$(Z'X)\begin{pmatrix}\alpha_{IV}\\ \beta_{IV}\end{pmatrix} = Z'y$$

$$\Rightarrow \begin{bmatrix} n & n\overline{X} \\ 0 & \frac{n}{2}\left(\overline{X}_{II} - \overline{X}_{I}\right)\end{bmatrix}\begin{bmatrix}\alpha_{IV}\\ \beta_{IV}\end{bmatrix} = \begin{bmatrix} n\overline{Y} \\ \frac{n}{2}\left(\overline{Y}_{II} - \overline{Y}_{I}\right)\end{bmatrix}$$

$$n\,\alpha_{IV} + n\,\beta_{IV} = n\overline{Y}$$

$$\frac{n}{2}\left(\overline{X}_{II} - \overline{X}_{I}\right)\beta_{IV} = \frac{n}{2}\left(\overline{Y}_{II} - \overline{Y}_{I}\right)$$

$$\Rightarrow \quad \beta_{IV} = \frac{\overline{Y}_{II} - \overline{Y}_{I}}{\overline{X}_{II} - \overline{X}_{I}}$$

and $\alpha_{IV} = \overline{Y} - \beta_{IV}\overline{X}$.

If n is odd, one should discard the central observation before applying this procedure. Under fairly general conditions this IV estimator (due to Wald) is consistent. But it usually has large standard errors.