- LINEAR MODEL WITH **NON-RANDOM REGRESSOR X**

$$Y = X\beta + e \,, \quad e \sim (0, \sigma^2 I_T)$$

THE OLS ESTIMATOR $b$ OF $\beta$ IS ESTIMATED FROM A SAMPLE OF SIZE **T**

1) $b$ is BLUE (Gauss-Markov theorem)

- this result holds for $b$'s from samples of ANY SIZE T BIG OR/AND SMALL

2) IF ADDITIONALLY $e \sim \mathbf{N}(0, \sigma^2 I_T)$, then we know

- the exact distribution of $b \sim N(\beta, \sigma^2 (X'X)^{-1})$
- — " — — " — — " — of $t = b_j / \hat{se}(b_j) \sim t_{T-K}$
- — " — — " — — " — of $F = \dots \quad \sim F_{J, T-K}$

**IN ANY SAMPLE BIG OR SMALL**

- WHEN THE ASSUMPTION CHANGES TO: X IS RANDOM AND/OR $e$ is NOT NORMAL $\Rightarrow$

- **the BLUE PROPERTY MAY NOT HOLD IN SMALL SAMPLE** (i.e. for $b$ calculated from a sample of size $T < 30$)
- **the distribution of $b$, $t$, $F$ statistics MAY NOT BE** $N()$, $t_{T-K}$, $F_{J,T-K}$ **IN SMALL SAMPLE**

→ IN GENERAL, THEY WILL BE UNKNOWN IN SMALL SAMPLE

ALL WE CAN HOPE FOR is THAT THE "NICE" PROPERTY OF THE ESTIMATOR AND "NICE" DISTRIBUTIONS OF test statistics CAN BE FOUND IN LARGE SAMPLES, IN THE LIMIT AS $T \to \infty$

FOR THIS, WE CONSIDER A SEQUENCE OF ESTIMATORS BASED ON SAMPES OF INCREASING SIZES T:

$b_{10}$

$b_{100}$ . $b_T$

$b_{1000}$

$b_{10000}$

⋮

ect

this is a sequence $b_T$. If $b_T$ converges in probability $b_T \xrightarrow{P} \beta$. then **b is said to be CONSISTENT (IN LARGE SAMPLE)**

Later, we'll look for the limiting distributions of b, t, etc when $T \to \infty$

$$\lim_{T \to \infty} P(|b_T - \beta| < \varepsilon) = 1 \quad (\text{illustration})$$

# 2. LAW OF LARGE NUMBERS

## strong

If $\bar{X}_n$ is sample mean of a random sample of size $n$ from a distribution with mean $\mu$

then

$$P\left(\lim_{n\to\infty} \bar{X}_n = \mu\right) = 1$$

$$\bar{X}_n \xrightarrow{a.s.} \mu$$ almost sure convergence

## weak

$x_1, \dots, x_n$ random sample with mean $\mu$.

$\bar{X}_n$ is sample mean. Then

$$\operatorname{plim} \bar{X}_n = \mu$$ convergence in probability

i.e. $$\lim_{n\to\infty} P\left(|\bar{X}_n - \mu| < \varepsilon\right) = 1$$

$$\bar{X}_n \xrightarrow{P} \mu$$

# CONSISTENCY AND PLIM OPERATOR

## CONSISTENCY:

- intuition: property of estimator such that it very likely is very close to the true, unknown parameter value when $T \to \infty$
- definition:

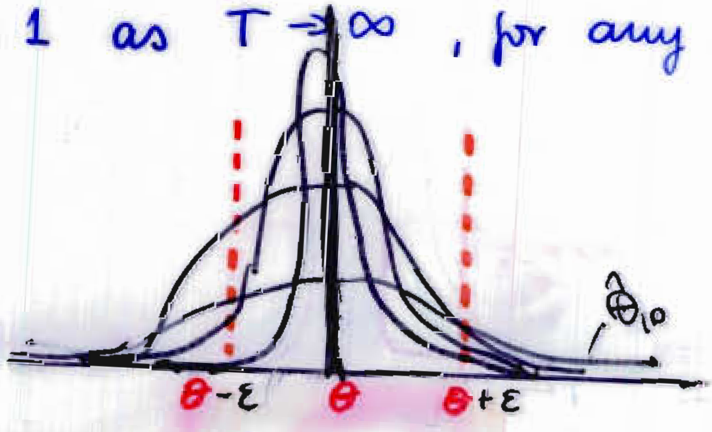
$\theta$: unknown parameter

$\hat{\theta}_T$: estimator of $\theta$ based on sample of T obs.

$$\lim_{T \to \infty} P\left(\left|\hat{\theta}_T - \theta\right| < \varepsilon\right) = 1$$

where $\varepsilon$ small, positive.

"probability of $\hat{\theta}_T$ falling in the interval $[\theta-\varepsilon, \theta+\varepsilon]$ approaches 1 as $T \to \infty$, for any $\varepsilon$, even very small".




T = 10
T = 100
T = 1000
T = 1000 $\hat{\theta}_{1000}$

consider varying T, increasing say: 100, 1000, 10 000 etc. $\hat{\theta}_{100}, \hat{\theta}_{1000}, \hat{\theta}_{10000}$

Consider estimators $\hat{\theta}_T$ based on samples of sizes T = 100, T = 1000, T = 10 000 etc. They form a sequence. This sequence CONVERGES IN PROBABILITY TO $\theta$.

$\theta$ IS THE PROBABILITY LIMIT OF THE SEQUENCE $\theta_T$.

EQUIVALENTLY

$$\text{plim } \hat{\theta}_T = \theta$$

We say, the estimator is consistent if this holds.

TO CHECK FOR CONSISTENCY IN PRACTICE WE CHECK IF:

- $$\lim_{T \to \infty} E(\hat{\theta}_T) = \theta$$

- $$\lim_{T \to \infty} \text{var}(\hat{\theta}_T) = 0$$

hold jointly. (!), i.e. the distribution of $\hat{\theta}_T$ collapses around $\theta$.

THE FOLLOWING RULES ARE IMPORTANT :

1. plim of a constant, say $\theta$, is equal to the constant

$$\text{plim}\ \beta = \beta$$

2. plim of the sum or difference of two r.v. say, $Y_1$ and $Y_2$ is equal to the sum of their plims respectively

$$\text{plim}\,(Y_1 \pm Y_2) = \text{plim}\,(Y_1) \pm \text{plim}\,(Y_2)$$

3. plim of a product (ratio) of two r.v. is equal to the product/ratio of their plims

$$\text{plim}\,(Y_1 \cdot Y_2) = \text{plim}\,Y_1 \cdot \text{plim}\,Y_2 \quad , \quad \text{plim}\,\frac{Y_1}{Y_2} = \frac{\text{plim}\,Y_1}{\text{plim}\,Y_2}$$

4. Slutsky's : plim of a cont. function $g$ of a r.v. $Y$ is equal to the fuction of plim

$$\text{plim}\left[g(Y)\right] = g(\text{plim}\,Y)$$

SAME HOLDS FOR THE EXPECTATION OPERATOR:

1. $E(\theta) = \theta$
2. $E(Y_1 + Y_2) = E(Y_1) + E(Y_2)$ ; $E(Y_1 - Y_2) = E(Y_1) - E(Y_2)$

BUT NOT FOR RULES 3 AND 4 !

ONLY IF $Y_1$ and $Y_2$ INDEPENDENT

$$E(Y_1 \cdot Y_2) = E(Y_1) \cdot E(Y_2)$$

ALSO: $E(Y_1/Y_2) \neq E(Y_1)/E(Y_2)$, $E[g(Y)] \neq g[E(Y)]$ NEVER. HENCE, plim better to check for consistency as E to check for unbiasedness.

Consider the model:

$$y = X\beta + e \quad , \quad e \sim (0, \sigma^2 I_T)$$

Where X is RANDOM

we check for the consistency of b, the OLS estimator of $\beta$:

$$b = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'e$$

$$= \beta + \left(\frac{X'X}{T}\right)^{-1}\left(\frac{X'e}{T}\right)$$

1° $\left(\frac{X'X}{T}\right)$: components are $\left(\sum_t x_{t2}\right)/T$ and $\left(\sum_t x_{t2}x_{t3}\right)/T$, $\left(\sum_t x_{t2}x_{t4}\right)/T$ etc....

We assume that those terms have finite plims.
(This assumptions is satisfied if random regressors X's have finite variances and covariances)

we denote: $$\text{plim}\left(\frac{X'X}{T}\right) = \Sigma_{XX}$$

and assume $\Sigma_{XX}$ is nonsingular.

2° $\left(\frac{X'e}{T}\right)$ consists of terms $\left(\sum_t x_{tk} e_t\right)/T$ ... These terms are consistent estimators of covariances between x's and e. It means that if errors are contemporaneously uncorrelated with the regressors, the term $\frac{X'e}{T}$ is a consistent estimator of zero (vector of zeros), so that

$$\text{plim}\ \frac{X'e}{T} = 0$$

hence:

$$\text{plim}\ b = \beta + \left(\text{plim}\ \frac{X'X}{T}\right)^{-1} \text{plim}\ \frac{X'e}{T}$$

$$= \beta + \Sigma_{XX}^{-1} \cdot 0$$

$$= \beta$$

! if regressors are correlated with errors, $\text{plim}\ \frac{X'e}{T} \neq 0$ and $\text{plim}\ b \neq \beta$

The estimators of regressor variances $\frac{1}{T}\sum(x_t-\bar{x}_0)^2$ are consistent, as are the estimator $\frac{1}{T-1}\sum(x_t-\bar{x})^2$

$$\text{plim}\ \frac{1}{T}\sum(x_t-\bar{x})^2 = \sigma_x^2$$

# RANDOM REGRESSOR MODELS

WE RELAX THE ASSUMPTION THAT REGRESSORS ARE NONSTOCHASTIC, FIXED IN REPEATED SAMPLES

1. The random regressors are INDEPENDENT OF ALL ERRORS
   OLS unbiased and consistent

2. The random regressors are contemporaneously uncorrelated with the error term
   OLS unbiased and consistent

Assumptions are altered to guarantee the exact properties

- the model is linear in parameters
- $[(x_{t2}, x_{t3}, \ldots x_{tK}, y_t) : t = 1 \ldots T]$ are i.i.d. arrays (random sample)
- No perfect collinearity
- $E(e \mid x_2, x_3 \ldots x_K) = 0 \Rightarrow \text{cov}(e_t x_{tK}) = 0\ \forall K$
- $\text{Var}(e \mid x_2, x_3 \ldots x_K) = \sigma^2$

which hold conditional on $X$.

OLS is biased and consistent in time series models

$y_t$ [illegible]

3. the random regressors are contemporaneously correlated with the error term.

OLS biased and inconsistent

examples

1° lagged dependent variable and correlated errors.

$$y_t = \beta_1 + \beta_2 X_t + \beta_3 y_{t-1} + \underbrace{(e_t - \lambda e_{t-1})}_{v_t}$$

$e_t$ independent, $e_t \sim (0, \sigma^2)$

two problems:

- $e_t - \lambda e_{t-1}$ is correlated with the regressor $y_{t-1}$

as $y_{t-1} = \beta_1 + \beta_2 X_{t-1} + \beta_3 y_{t-2} + e_{t-1} - \lambda e_{t-2}$

- $e_t - \lambda e_{t-1}$ is serially correlated

$$E\left[(e_t - \lambda e_{t-1})(e_{t-1} - \lambda e_{t-2})\right] = -\lambda E(e_{t-1})^2 = -\lambda \cdot \sigma^2$$

2° measurement error

explanatory variable is measured with error

3° simultaneous equations

$$C_t = \beta + \gamma y_t + e_t$$

$$Y_t = C_t + i_t$$

if substitute for $C_t$ in 2nd eq, you'll see $Y_t$ is a f. of $e_t$

For variance

given: $\text{plim } \frac{e'e}{T} = \sigma^2$

$$\hat\sigma^2 = \frac{1}{T-K} e'Me = \frac{1}{T-K} e'e - \frac{1}{T-K}\left(e'X(X'X)^{-1}X'e\right)$$

$$\text{plim } \frac{e'e}{T-K} = \text{plim } \frac{e'e}{T} = \sigma^2$$

Since $\text{plim}_{T\to\infty} \frac{X'X}{T} = \Sigma_{xx}$ finite, nonsingular

and $\text{plim } \frac{X'e}{T} = 0$ by Lemma

$$\text{plim } \left(\frac{1}{T-K} e'X\right)\left(\frac{X'X}{T}\right)^{-1} \frac{X'e}{T} = 0$$

$\hat\sigma^2 = \frac{\hat e'\hat e}{T-K}$ IS CONSISTENT

$$\left(\underset{\downarrow\, 0}{\text{pli } \frac{e'x}{T-K}} \cdot \underset{\downarrow\, \Sigma_{xx}^{-1}}{\left(\text{pli } \frac{X'X}{T}\right)^{-1}} \; \underset{\downarrow\, 0}{\text{pli } \frac{X'e}{T}}\right)$$

$$\text{pli } \hat\sigma^2 = \text{pli } \frac{1}{T-K} e'e - \text{pli } \ldots = \sigma^2$$

## ⇒ efficient and asymptotically efficient estimator

when we can derive exact small sample results we call
**efficient** unbiased estimator with minimum var (among unbiased)

in the asymptotic setup:
**asymptotic efficient** is a consistent estimator with a smaller variance of it's asymptotic distribution, or the asymptotic variance is lower.

WE ALWAYS WANT TO HAVE CONSISTENT AND ASYMPTOTICALLY EFFICIENT ESTIMATORS.

THE MLE ESTIMATOR IS
- CONSISTENT
- ASY EFFICIENT
- ASY NORMALLY DISTRIBUTED

HENCE: IT IS THE BEST ASYMPTOTIC ESTIMATOR