$y = X\beta + e$

$y' = (y_1, y_2 \dots, y_T)$, $e' = (e_1, e_2 \dots e_T)$, $\beta' = (\beta_1, \beta_2 \dots \beta_K)$

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1K} \\ x_{21} & x_{22} & & x_{2K} \\ \vdots & \vdots & & \vdots \\ x_{T1} & x_{T2} & & x_{TK} \end{bmatrix}$$

Assump

1. X nonstoch. rank $K \leq T$

$\lim\limits_{T \to \infty} \frac{X'X}{T} = Q$ finite, nonsingular ← no collinearity in the limit

2. e are random errors: $E(e) = 0$, $Eee' = \sigma^2 I$

i.i.d.

OLS minimizes the criterion:

$$S = \sum_{t=1}^{T} (y_t - \beta_1 x_{t1} - \dots - \beta_K x_{tK})^2 = (y - X\beta)'(y - X\beta)$$

$$\frac{\partial S}{\partial \beta} = -2X'y + 2X'X\beta = 0$$

$$X'X\,b = X'y$$

at the maximum replace β

if X has rank K $\Rightarrow (X'X)^{-1}$ exists

$$b = (X'X)^{-1} X'y$$

$$\frac{\partial^2 S}{\partial\beta\partial\beta'} = 2X'X \quad \text{pos. det.}$$

OLS : $S = (y - X\beta)'(y - X\beta)$

GLS : $S^* = (y - X\beta)' W (y - X\beta)$ where W pos. det.

GAUSS MARKOV: OLS estimator b is unbiased and efficient in the class of linear, unbiased estimators (under assumptions of the classical linear regression)

linear estimator $\beta^* = HY$, a linear transform of Y

proof of no bias:

$$b = (X'X)^{-1}X'y$$

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

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

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

$$E(b) = \beta + (X'X)^{-1}X'Ee$$

$$\underline{E(b) = \beta}$$

An estimator is efficient with respect to a class of estimators if it has a variance no greater than the variance of any other estimator in that class

class of linear estimators:

$$E(b-Eb)(b-Eb)' = E(b-\beta)(b-\beta)'$$

$$= E\,(X'X)^{-1}X'ee'X(X'X)^{-1}$$

$$= (X'X)^{-1}X'\,E\,ee'\,X(X'X)^{-1}$$

$$= \sigma^2 (X'X)^{-1}X'X(X'X)^{-1}$$

$$= \sigma^2 (X'X)^{-1}$$

compare to another arbitrary linear estimator

$$\beta^* = Hy$$

$$H = (X'X)^{-1}X' + C$$

C is a matrix of constants.

$$\beta^* = Hy = HX\beta + He$$

$$E\beta^* = HX\beta$$

There's no bias $\Leftrightarrow$ $HX = I$

but:

$$HX = (X'X)^{-1}X'X + CX = I + CX$$

$HX = I \quad \Leftrightarrow \quad \underline{CX = 0}$

Then:

$$E(\beta^* - \beta)(\beta^* - \beta)' = E(H\,ee'\,H')$$

$$= E[(X'X)^{-1}X' + C]ee'[X(X'X)^{-1} + C']$$

$$= \sigma^2 [(X'X)^{-1}X'X(X'X)^{-1} + (X'X)^{-1}\underbrace{X'C'}_{=0}$$

$$+ \underbrace{CX}_{=0}(X'X)^{-1} + CC']$$

$$= \sigma^2 (X'X)^{-1} + \sigma^2 CC'$$

but CC' is pos. semidef: $\Rightarrow$ $\underline{VAR(\beta^*) \geq Var(b)}$

$$\underset{(K,n)(n,K)}{CC'} \gg 0$$

u of dim $(K,1)$

$\underset{(1\times K)(K,n)(n\times K)(K\times 1)}{u'CC'u} \overset{?}{\geq} 0 \quad \forall u$; yes because

$$\underset{(1\times n)}{(C'u)'}\underset{(n\times 1)}{(C'u)} = \|C'u\|^2 \geq 0$$

GAUSS MARKOV IS A SMALL SAMPLE RESULT

⇒ VALID FOR ANY SAMPLE $T \geq K$ under assumptions.

⇒ FOR ANY DISTRIBUTION OF e. IF e are Normal, OLS IS THE BEST AMONG ALL (LINEAR & NONLINEAR) ESTIMATORS

— • —

THE VARIANCE ESTIMATOR IS QUADRATIC!

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

is unbiased of σ^2 with $\hat{e} = y - Xb$

no bias:

$$\hat{e} = y - Xb = y - X(X'X)^{-1}X'y$$

$$= \left[I - X(X'X)^{-1}X'\right]y = My$$

$$= \left[I - X(X'X)^{-1}X'\right](X\beta + e)$$

$$= X\beta - X(\cancel{X'X})^{-1}\cancel{X'X}\beta + Me = Me$$

$$\hat{e} = Me$$

M symmetric, idempotent: $M' = M$ and $MM = M^2 = M$

trace M = rank M

$$E\,\hat{e}'\hat{e} = E(e'M'Me) = E(e'Me)$$

$\underset{1\times n}{e'}\underset{n\times n}{M}\underset{n\times 1}{e}$ is a scalar

$$E(e'Me) = E\,\mathrm{tr}\,(e'Me)$$

$$= E\,\mathrm{tr}(M\,e'e) \qquad \mathrm{tr}(AB) = \mathrm{tr}(BA)$$

$$= \mathrm{tr}\,E(M\,ee') \qquad \text{tr is a linear function}$$

$$= \mathrm{tr}\,M \cdot E\,ee' \qquad M \text{ nonstoch.}$$

$$= \mathrm{tr}\,M \cdot \sigma^2 I$$

$$= \sigma^2\,\mathrm{tr}\,M$$

$$\mathrm{tr}\,M = \mathrm{tr}\,I_T - \mathrm{tr}\left(X(X'X)^{-1}X'\right) \qquad \mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$$

$$= \mathrm{tr}\,I_T - \mathrm{tr}\left(X'X(X'X)^{-1}\right) \qquad \mathrm{tr}\,AB = \mathrm{tr}\,BA$$

$$= \mathrm{tr}\,I_T - \mathrm{tr}\,I_K = T-K$$

$$E\,\hat{e}'\hat{e} = \sigma^2(T-K)$$

$$\frac{E\,\hat{e}'\hat{e}}{T-K} = \sigma^2$$

We showed that the OLS estimator of the general linear model

$$Y = X\beta + e$$

$E(e)=0$, $var(e) = \sigma^2 I_T$ and nonstochastic X matrix is BLUE : Best Linear Unbiased

IF WE ADD THE ASSUMPTION:

$$e_t \sim N(0, \sigma^2) \quad \forall t = 1, \dots T$$

then, OLS is BUE

- WE KNOW THE EXACT DISTRIBUTION OF b:

$$b \sim N(\hat{\beta}, \sigma^2 (X'X)^{-1})$$

- WE KNOW THE EXACT DISTRIBUTION OF the "t-student" statistic under the null $\beta_j = \beta_j^\circ$

$$t = \frac{b_j - \beta_j^\circ}{s.e(b_j)} \sim t_{(T-K)}$$

which is the t-distribution with T-K degrees of freedom under the $H_0: \beta_j = \beta_j^\circ$.

The basic hypothesis: $H_0: \beta_j = 0$ is tested against $H_A: \beta_j \neq 0$ at level α.
We reject H_0 if $|t| > t_c\,(T-K, \frac{\alpha}{2})$

- WE KNOW THE EXACT DISTRIBUTION OF the F statistic

$$F = \frac{ESS/(K-1)}{\hat{\sigma}^2}$$

Under $H_0: \beta_2 = \beta_3 = \cdots = \beta_K = 0$ (all slopes $=0$) against H_A: at least one slope is not zero. Under H_0, $F \sim F(K-1, T-K)$ i.e. is Fisher distributed with $K-1$ and $T-K$ degrees of freedom.

WHEN THE CLASSICAL ASSUMPTIONS ON X and errors DON'T HOLD, WE CAN'T DETERMINE THE ~~EXACT~~ DISTRIBUTION OF THE ESTIMATOR, WE DONT KNOW THE ESTIMATORS ~~EXACT~~ PROPERTIES, AND DON'T KNOW THE EXACT DISTRIBUTIONS OF THE TEST STATISTICS.

IN THESE CASES WE WILL DERIVE

- APPROXIMATE PROPERTIES OF ESTIMATORS
- APPROXIMATE DISTRIBUTIONS OF TEST STATISTICS

THIS APPROXIMATION IS OBTAINED BY ASSUMING

$$T \to \infty$$

i.e. the sample size goes to infinity.

ASYMPTOTIC RESULTS WE HAVE SEEN:

1. CENTRAL LIMIT THEOREM:

- for $X_1, X_2, \ldots, X_T$ i.i.d with mean μ and variance σ^2, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_T - T \cdot \mu}{\sqrt{T} \cdot \sigma} \text{ goes to } N(0,1) \ \underline{\text{as } T \to \infty}$$

equivalently

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{T}} \text{ goes to } N(0,1) \ \underline{\text{as } T \to \infty}$$