GENERAL LINEAR MODEL

- y_t : DEPENDENT VAR (REGRESSAND)
- $x_{t1}, x_{t2}, x_{t3}, \dots x_{tK}$ EXPLANATORY VARIABLES (REGRESSORS), K variables
- $\beta_1, \beta_2, \beta_3 \dots \beta_K$ coefficients
- x_{t1} USUALLY IS A VECTOR OF ONES, HENCE β_1 is INTERCEPT, while $\beta_2 \dots \beta_K$ are SLOPE COEFICIENTS : $\beta_K = \frac{\partial y_t}{\partial x_{tK}}$

MODEL:

- $$\boxed{y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_K x_{tK} + e_t}$$

ASSUMPTIONS:

1. vector of obs on y can be expressed as a linear combination of the sample obs on the

explanatory X plus a disturbance term.

for every obs t we have

$$y_1 = \beta_1 + \beta_2 x_{12} + \beta_3 x_{13} + \dots + \beta_K x_{1K} + e_1$$
$$y_2 = \beta_1 + \beta_2 x_{22} + \beta_3 x_{23} + \dots + \beta_K x_{2K} + e_2$$
$$y_3 = \beta_1 + \beta_2 x_{32} + \beta_3 x_{33} + \dots + \beta_K x_{3K} + e_3$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$y_T = \beta_1 + \beta_2 x_{T2} + \beta_3 x_{T3} + \dots + \beta_K x_{TK} + e_T$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} 1 & x_{12} & x_{13} & \dots \\ 1 & x_{22} & x_{23} & \dots \\ 1 & x_{32} & x_{33} & \dots \\ \vdots & \vdots & \vdots & \\ 1 & x_{T2} & x_{T3} & \dots \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_T \end{bmatrix}$$

$$\underbrace{Y}_{T\times 1} = \underbrace{X}_{T\times K} \underbrace{\beta}_{K\times 1} + \underbrace{e}_{T\times 1}$$

within $X = [x_1 \;\; x_2 \;\; x_3 \,..\, x_K]$ where

$$\mathbf{X} = \begin{bmatrix} | & | & | \\ x_1 & x_2 & x_3 \\ | & | & | \end{bmatrix}$$

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \\ \vdots \\ x_{T2} \end{bmatrix}, \quad x_3 = \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \\ \vdots \\ x_{T3} \end{bmatrix} \text{ etc.}$$

2. $E_x(e) = 0$, i.e. $E_x(y) = X\beta$

3. $E_x(ee') = \sigma^2 I_T$ VARIANCE MATRIX

$$e_x \sim (0, \sigma^2 I_T)$$

and

$$E_x[(y - X\beta)(y - X\beta)'] = E_x[ee'] = \sigma^2 I_T$$

$$y|x \sim (X\beta, \sigma^2 I_T)$$

$$\mathrm{var}_x(e) = E_x(ee') = \begin{bmatrix} \underline{\mathrm{var}\ e_1} & \mathrm{cov}(e_1, e_2) & \cdots & \mathrm{cov}(e_1 e_K) \\ \mathrm{cov}(e_2, e_1) & \underline{\mathrm{var}\ e_2} & & \\ \vdots & & \ddots & \\ \mathrm{cov}(e_T, e_1) & & & \underline{\mathrm{var}\ e_T} \end{bmatrix}$$

- each e has the same variance σ^2 and mean 0.
- all disturbances are pairwise uncorrelated.

4. $\rho(X) = K$

the explanatory variables (including $x_1 = \begin{matrix} 1 \\ 1 \\ \vdots \\ 1 \end{matrix}$) do not form a linearly dependent set.

for example, if we had just two explanatory variables X_2 and X_3 and this assumption was not fulfilled, there would exist an exact relationship

$$X_3 = c_1 + c_2 X_2$$

which $\Rightarrow$

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + e$$

$$Y = (\beta_1 + \beta_3 c_1) + (\beta_2 + \beta_3 c_2) X_2 + e$$

the constants c_1 and c_2 can be determined exactly and we can estimate the intercept and slope of * but there is NO WAY to obtain estimates of β parameters

5. X is a NONSTOCHASTIC MATRIX

fixed in repeated samples, then the only source of variation is in the e vector and hence in Y.

additionally we may assume that conditional on X (fixed)

7. $e|X \sim N(0, \sigma^2 I_T)$ and hence

$y|X \sim N(X\beta, \sigma^2 I_T)$

ESTIMATION BY OLS

OLS estimate for β minimizes $\sum e_t^2$, i.e:

$$S(\beta_1, \beta_2, \dots \beta_K) = \sum_{t=1}^{T} \left[y_t - E_X(y_t) \right]^2$$

$$= \sum_{t=1}^{T} \left(y_t - \beta_1 - \beta_2 x_{t2} - \dots \right)^2$$

we differentiate S w.r. to $\beta_1, \beta_2 \dots \beta_K$ and set the derivatives $= 0$.

The partial derivatives are equal to zero for $b_1, b_2 \dots b_K$ which are ESTIMATES OF $\beta_1, \beta_2 \dots \beta_K$, at the minimum of $S(\beta)$.

$$\frac{\partial S}{\partial \beta_1} = 2T\beta_1 + 2\beta_2 \sum x_{t2} + 2\beta_3 \sum x_{t3} - 2 \sum y_t$$

$$\frac{\partial S}{\partial \beta_2} = 2\beta_1 \sum x_{t2} + 2\beta_2 \sum x_{t2}^2 + 2\beta_3 \sum x_{t2} x_{t3} - 2 \sum x_{t2}$$

$$\frac{\partial S}{\partial \beta_3} = 2\beta_1 \sum x_{t3} + 2\beta_2 \sum x_{t2}x_{t3} + 2\beta_3 \sum x_{t3}^2 - 2\sum x_{t3}y$$

$$\frac{\partial S}{\partial \beta_1} = 0 \Rightarrow \quad T\,b_1 + \sum x_{t2} \cdot b_2 + \sum x_{t3}\, b_3 = \sum y_t$$

$$\frac{\partial S}{\partial \beta_2} = 0 \Rightarrow \quad \sum x_{t2}\, b_1 + \sum x_{t2}^2 \cdot b_2 + \sum x_{t2}x_{t3}\, b_3 = \sum x_{t2}y$$

$$\frac{\partial S}{\partial \beta_3} = 0 \Rightarrow \quad \sum x_{t3}\, b_1 + \sum x_{t2}x_{t3}\, b_2 + \sum x_{t3}^2\, b_3 = \sum x_{t3}y$$

$\sum x_{t1}^2$ $\qquad\qquad$ $\sum x_{t1}\cdot y_t$

$$\begin{bmatrix} T & \sum x_{t2} & \sum x_{t3} \\ \sum x_{t2} & \sum x_{t2}^2 & \sum x_{t2}x_{t3} \\ \sum x_{t3} & \sum x_{t2}x_{t3} & \sum x_{t3}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \sum y_t \\ \sum x_{t2}y \\ \sum x_{t3}y \end{bmatrix}$$

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

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

again b has the "blue" property, in particular:

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

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

$$E(b) = \beta \quad \text{since}$$

$$E((X'X)^{-1}X'e) = (X'X)^{-1}X'E(e) = 0$$

by Assumptions 2 and 5. $E(X'e) = X'E(e) = 0$

estimation of σ^2

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_t^2}{T-K} = \frac{\hat{e}'\hat{e}}{T-K} = \frac{(Y - Xb)'(Y - Xb)}{T-K}$$

K: K degrees of freedom are lost = I can pin down K among T residuals: one from the condition $\sum \hat{e}_t = 0$, one for $\sum x_{t1}\hat{e}_t = 0$, one for $\sum x_{t2}\hat{e}_t = 0$ etc till $\sum x_{tK}\hat{e}_t = 0$

we have shown that b is unbiased. It's variance, as in the simple regression is:

$$var(b) = E\left[(b-\beta)(b-\beta)'\right] =$$

$$\begin{bmatrix} \underline{var\ b_1} & cov(b_1 b_2) & cov(b_1, b_3) & \cdots \\ cov(b_2 b_1) & \underline{var\ b_2} & cov(b_2 b_3) & \cdots \\ cov(b_3 b_1) & cov(b_3 b_2) & \underline{var\ b_3} & \cdots \\ \vdots & \vdots & \vdots & \\ & & & \underline{var\ b_K} \end{bmatrix}$$

$$= \sigma^2 (X'X)^{-1} = \sigma^2 \begin{bmatrix} T & \sum x_{t2} & \sum x_{t3} \\ \sum x_{t2} & \sum x_{t2}^2 & \sum x_{t2} x_{t3} \\ \sum x_{t3} & \sum x_{t2} x_{t3} & \sum x_{t3}^2 \end{bmatrix}^{-1}$$

$$\widehat{var}(b) = \hat{\sigma}^2 (X'X)^{-1} = \hat{\sigma}^2 \begin{bmatrix} \quad X'X \quad \end{bmatrix}^{-1}$$

R^2 : COEFFICIENT OF DETERMINATION

measures the proportion of variation in the dependent var. that is explained by variation in the explanatory variable (regressors)

$$R^2 = \frac{\text{variation in } Y \text{ accomodated by explanatory var}}{\text{total variation in } Y} = \frac{ESS}{TSS}$$

$TSS = RSS + ESS$

$$= \frac{\sum(\hat{y}_t - \bar{y})^2}{\sum(y_t - \bar{y})^2} = 1 - \frac{\text{error variation}}{\text{total variation}} = 1 - \frac{RSS}{TSS}$$

$$= 1 - \frac{\sum \hat{e}_t^2}{\sum(y_t - \bar{y})^2} = 1 - \frac{\hat{e}'\hat{e}}{y'y - T\bar{y}^2}$$

Sometimes it is useful to compute the adjusted R^2:

$$\bar{R}^2 = 1 - \frac{\hat{e}'\hat{e}/T-K}{y'y - T\bar{y}^2/T-1} = 1 - \frac{RSS/T-K}{TSS/T-1}$$

Since $R^2 = 1 - \dfrac{\hat{e}'\hat{e}/T}{y'y - T\bar{y}^2/T}$, we have

$$\bar{R}^2 = 1 - \frac{T-1}{T-K}(1-R^2) = \frac{1-K}{T-K} + \frac{T-1}{T-K} R^2$$

when we add a regressor $\Sigma \hat{e}_t^2 \downarrow$ But does the addition of an extra regressor improve the explanatory power of regressors?

well, R^2 always $\uparrow$ when we add a regressor, hence we can be mislead and believe that an extra regressor improves the fit.

if an additional var. produces too small a reduction in $(1-R^2)$ to compensate for the increase in $\frac{(T-1)}{(T-K)}$!

$\bar{R}^2$ will decline.

note: for $R^2 = 0$, $\bar{R}^2 = 1 - \frac{T-1}{T-K}$ and can be negative.

note also that if in a regression we omit the constant, without a constant

$$\Sigma(y_t - \bar{y})^2 \neq \Sigma(\hat{y}_t - \bar{y})^2 + \Sigma \hat{e}_t^2$$

$$TSS \neq ESS + RSS$$

UNLESS: WE REGESS $\tilde{y} = y - \bar{y}$ on $\tilde{x} = x - \bar{x}$!