

Simple linear Regression Model

Specify simple linear regression model

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

β_0 & β_1 are unknown parameters, and we want to estimate these parameters

they are called regression coefficients

Y_t & X_t are observable r.v.s but u_t is not
↳ disturbance term
(also estimated)

$\beta_0 \rightarrow$ intercept term

$\beta_1 \rightarrow$ slope term

$$C_t = \beta_0 + \beta_1 Y_t^d + u_t, \quad t = 1, 2, \dots$$

Vectors of observations

$$\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \begin{bmatrix} Y_t \\ \vdots \\ \end{bmatrix} \quad \begin{bmatrix} X_t \\ \vdots \\ \end{bmatrix} \begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix}$$

$$\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rightarrow \text{constant term}$$
$$\underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}}_{\text{we want to predict}} \rightarrow \hat{u}_t$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t \quad t = 1, \dots, T$$

The model in eqn ① is subject to following assumptions

1- Expected or mean value of U_t is zero

$$E(U_t) = 0 \text{ for all } t \text{ } (\forall t)$$

this assumption is violated in various ways

(i) - left out variable,
assume you left out an important variable z_t , the correct model is

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 z_t + v_t \quad t = 1, \dots, T$$

$$Y_t = \beta_0 + \beta_1 X_t + U_t$$

$$U_t = \beta_2 z_t + v_t$$

$$E(U_t) = E(\beta_2 z_t) + E(v_t)$$

$$= \beta_2 E(z_t) + E(v_t)$$

$$\text{Since } E(v_t) = 0 \\ E(U_t) = \beta_2 E(z_t)$$

(ii) - usage of incorrect variable
instead of correct variable w_t , x_t is
used incorrectly

correct model is

$$Y_t = \alpha_0 + \alpha_1 w_t + v_t, \quad t = 1, \dots, T$$

$$E(U_t) = (\beta_0 - \alpha_0) + (\beta_1 (x_t - \alpha_1 w_t)) \neq 0$$

2- Variance of U_t is a constant and equal to σ^2 and does not systematically vary with t .

$$E[U_t - E(U_t)]^2 = E(U_t)^2 = \sigma^2 \quad \forall t$$

If variance of U_t was not same for all observations this means that all observations would not be equally reliable

this problem is called $\Rightarrow$ Heteroscedasticity

Homoscedasticity

3- Value of t^{th} disturbance is independent of value of any other disturbance term, like s^{th}

$$\text{Cov}(u_t, u_s) = E(u_t - E(u_t))(u_s - E(u_s))$$

$$= E(u_t u_s) = 0 \quad \text{for } t \neq s$$

$$\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \left[\begin{matrix} \uparrow \\ u_1 \\ u_2 \\ \vdots \end{matrix} \right]$$

Auto correlation

If this assumption is violated then u_t together with x_t becomes a member of systematic errors

4- x_t maybe regarded as

fixed $\nwarrow$
non stochastic

$\swarrow$ observed element
of a random variable
which do not vary
to sample to sample

5- If we assume x_t is a random variable we need to assume that u_t are independent of all T values of regressor x

$$\text{Cov}(u_t, x_s) = E(u_t x_s) = E(u_t) E(x_s)$$

$$E(u_t x_s) = 0, \quad \forall t$$

6- Intercept term γ & x must be linearly independent
All values of x_t should not be same at least one value of x_t must differ from others

$$\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \left[\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \right] \quad \begin{matrix} x_t \\ s \\ s \\ \vdots \\ s \end{matrix} \left[\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \right]$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + U_t$$

$$\begin{matrix} 1990 \\ 2020 \end{matrix} \begin{bmatrix} X_t \\ 5 \\ 7 \\ 14 \\ \vdots \end{bmatrix} \quad \begin{bmatrix} Z_t \\ 10 \\ 14 \\ 28 \\ \vdots \end{bmatrix} \rightarrow \text{multi collinearity}$$

If this condition is not satisfied we cannot estimate β_0 & β_1 separately

$$\text{let } e_t = 1 \quad \& \quad X_t = X_0$$

$$Y_t = \beta_0 e_t + \beta_1 X_0 \cdot e_t + U_t$$

$$= \underbrace{(\beta_0 + \beta_1 X_0)}_{\beta_0^*} e_t + U_t$$

$$= \beta_0^* e_t + U_t$$

Least Squares method

Estimation Problem

$$\textcircled{1} Y_t = \beta_0 + \beta_1 X_t + v_t, \quad t = 1, \dots, T$$

Since Y_t is a random variable, information about its behaviour is contained in its probability distr.

Since Y_t & U_t is linearly related therefore mean, variance & covariance of Y_t is related to distribution of U_t

$$E(Y_t) = E(\beta_0 + \beta_1 X_t + U_t)$$

$$= E(\beta_0) + E(\beta_1 X_t) + E(U_t)$$

$$= \beta_0 + \beta_1 E(X_t) + E(U_t)$$

Since $E(X_t) = x_t$ & $E(U_t) = 0$ by assumption 1

$$\textcircled{2} E(Y_t) = \beta_0 + \beta_1 x_t$$

$$\text{Var}(Y_t) = E[Y_t - E(Y_t)]^2$$

$$\text{Var}(Y_t) = E(Y_t - \beta_0 - \beta_1 x_t)^2$$

Since from eqn 1

$$Y_t - \beta_0 - \beta_1 x_t = U_t$$

$$\text{Var}(Y_t) = E(U_t^2) = \sigma^2$$

Covariance b/w. Y_t & Y_s ($t \neq s$)

$$\text{cov}(Y_t, Y_s) = E[Y_t - E(Y_t)][Y_s - E(Y_s)]$$

$$= E \left[\underbrace{Y_t - \beta_0 - \beta_1 x_t}_{U_t} \right] \left[\underbrace{Y_s - \beta_0 - \beta_1 x_s}_{U_s} \right]$$

$$= E(U_t U_s) = 0 \text{ by our assumption 3}$$

Problem of estimating parameters of model maybe viewed as one of estimating parameters of prob distribution of dependent variable Y_t

Since $E(Y_t) = \beta_0 + \beta_1 x_t$ if we estimate β_0 & β_1 ($\hat{\beta}_0 + \hat{\beta}_1$) we estimate mean of Y_t

Since $\text{Var}(Y_t) = \sigma^2$ estimating variance of Y_t is equal to estimating variance of U_t (σ^2)

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t \rightarrow \text{this is the estimated version of } U_t \text{ that's why it's not here}$$

where $\hat{\beta}_0$ & $\hat{\beta}_1$ denote least squares (LS) estimators of β_0 & β_1

Since $\hat{\beta}_0$ and $\hat{\beta}_1$ are random variables $\Rightarrow$

$\hat{Y}_t$ is also a random variable

There will be a discrepancy between actual values of Y_t & fitted values of $\hat{Y}_t$

$$\begin{matrix} 1990 \\ \vdots \\ 2020 \end{matrix} \begin{bmatrix} Y_t \\ 50 \\ 60 \\ \vdots \\ 100 \end{bmatrix} \quad \begin{bmatrix} X_t \\ 100 \\ 110 \\ \vdots \\ 250 \end{bmatrix} = \begin{matrix} Y_t^1 \\ 1990 \\ \vdots \\ 2020 \end{matrix} \quad \begin{bmatrix} \hat{Y}_t \\ 1990 \\ \vdots \\ 2020 \end{bmatrix}$$

$$\hat{Y}_t = 20 + 0.7 X_t$$

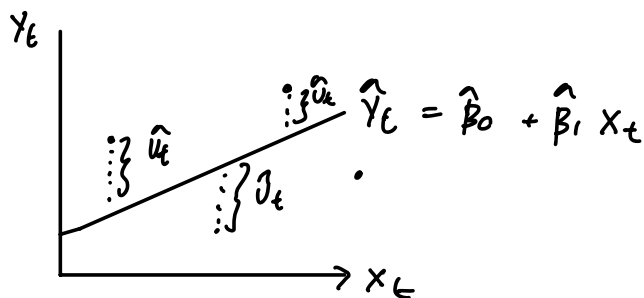
$$\hat{Y}_{1990} = 20 + 0.7 (100) = 90$$

These discrepancies are called residuals and called

$$\hat{U}_t$$

$$\hat{U}_{1990} = 50 - 90 = -40$$

$$\hat{U}_t = Y_t - \bar{Y}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t$$



Method of LS proposes to find values of $\hat{\beta}_0$ & $\hat{\beta}_1$ that will make sum of squares of these distances from the line

$$\bar{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t \text{ as small as possible}$$

$$\min \sum_{t=1}^T \hat{U}_t^2$$

SSR $\rightarrow$ sum of squares residuals
 $\hookrightarrow S$

$$\min S = \sum_{t=1}^T \hat{u}_t^2 = \sum (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)^2$$

with respect to $\hat{\beta}_0$ & $\hat{\beta}_1$

$$\left. \begin{aligned} \textcircled{1a} \partial S / \partial \hat{\beta}_0 &= -2 \sum_{t=1}^T (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t) = 0 \\ \textcircled{1b} \partial S / \partial \hat{\beta}_1 &= -2 \sum_{t=1}^T x_t (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t) = 0 \end{aligned} \right\} \text{normal equations}$$

$$\textcircled{2a} \sum_{t=1}^T y_t - \sum_{t=1}^T \hat{\beta}_0 - \sum_{t=1}^T \hat{\beta}_1 x_t = 0$$

$$\sum y_t - \hat{\beta}_0 T - \hat{\beta}_1 \sum x_t = 0$$

$$\textcircled{2b} -\hat{\beta}_0 \sum x_t - \hat{\beta}_1 \sum x_t^2 + \sum x_t y_t = 0$$

from 2a

$$\hat{\beta}_0 T = \sum y_t - \sum \hat{\beta}_1 x_t$$

$$\hat{\beta}_0 = \frac{\sum y_t}{T} - \hat{\beta}_1 \frac{\sum x_t}{T}$$

$$\textcircled{3} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

substitute 3 into eqn 2b

$$\sum x_t y_t - (\bar{y} - \hat{\beta}_1 \bar{x}) \sum x_t - \hat{\beta}_1 \sum x_t^2 = 0$$

$$\sum x_t y_t - \bar{y} \sum x_t - \hat{\beta}_1 \bar{x} \sum x_t - \hat{\beta}_1 \sum x_t^2 = 0$$

$$\sum x_t y_t - \bar{y} \sum x_t - \hat{\beta}_1 (\bar{x} \sum x_t + \sum x_t^2) = 0$$

$$\hat{\beta}_1 = \frac{(\sum x_t y_t + \bar{x} T \bar{y})}{\sum (x_t^2 + T \bar{x}^2)}$$

$$\hat{\beta}_1 = \frac{[T(\sum x_t y_t) - (\sum x)(\sum y)]}{[T(\sum x^2) - (\sum x)^2]}$$

estimator of β_1 in raw data form

$$\hat{\beta}_1 = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2}$$

$$\textcircled{4} \quad \hat{\beta}_1 = \frac{\sum_t x_t y_t}{\sum_t x_t^2}$$

$$x_t = x_t - \bar{x}$$

$$y_t = y_t - \bar{y}$$

estimator of β_1
in mean-deviation
form

from 1a, 1b

$$\text{sa} \quad \sum \hat{u}_t = 0$$

$$\text{sb} \quad \sum x_t \hat{u}_t = 0$$

$$\textcircled{7} \quad \sum_t \hat{y}_t \hat{u}_t = \beta_0 \sum \hat{u}_t + \beta_1 \sum x_t \hat{u}_t = 0$$

$$\sum_t \hat{y}_t^2 = \sum_t \hat{y}_t (y_t - \hat{u}_t) = \sum_t \hat{y}_t y_t - \sum_t \hat{y}_t \hat{u}_t$$

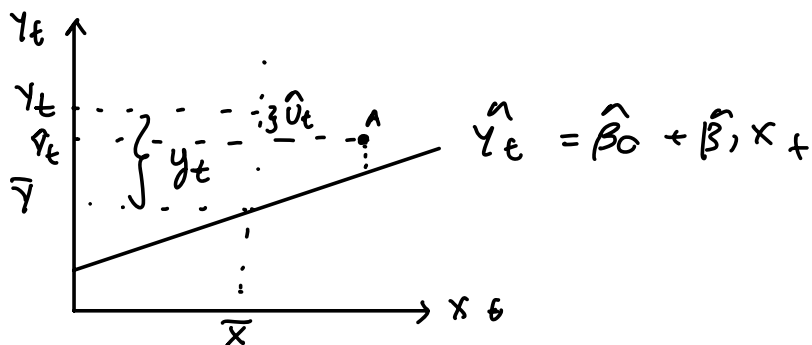
by using 7

$$\sum_t \hat{y}_t^2 = \sum_t \hat{y}_t y_t$$

$\hat{y}_t$ = fitted
 y_t = actual

Goodness of fit

how good is the fit of $\hat{y}$ to scatter of observations
on y_t & x_t ?



$$\hat{y}_t = \hat{Y} - \bar{Y}$$

We may define three different distances for ordering y_t of point A

1- Distance y_t from origin

$$y_t - 0 = y_t$$

2- Distance of y_t from its sample mean

$$y_t - \bar{Y} = y_t$$

3- Distance of y_t from fitted regression line

$$y_t - \hat{y}_t = \hat{u}_t$$

Distance to use to form a measure of goodness of fit is variation of y_t around its sample mean

$$\sum_t (y_t - \bar{Y})^2 = \sum y_t^2$$

This distance can be decomposed as

↳ part explained by $\hat{y}_t$
($\hat{y}_t - \bar{Y} = \hat{y}_t$)

↳ part explained by $\hat{u}_t$
($y_t - \hat{y}_t$)

$$\begin{aligned} \sum y_t^2 &= \sum (\hat{y}_t + \hat{u}_t)^2 = \sum \hat{y}_t^2 + \sum \hat{u}_t^2 - \cancel{2 \sum \hat{y}_t \hat{u}_t} \\ &= \textcircled{9} \sum \hat{y}_t^2 + \sum \hat{u}_t^2 \end{aligned}$$

$= 0$

Dividing both sides of eqn 9 by $\sum y_t^2$

gives definition of coefficient of determination

$$\textcircled{10} R^2 = \frac{\sum \hat{y}_t^2}{\sum y_t^2} = 1 - \left(\frac{\sum \hat{u}_t^2}{\sum y_t^2} \right)$$

$$0 \leq R^2 \leq 1$$

If $R^2 = 0 \Rightarrow$ observation points are scattered around regression line in a purely random fashion

If $R^2 = 1 \Rightarrow$ all observations lie on regression line

Low values of $R^2 \Rightarrow$ a poor fit of regression
 High values of $R^2 \Rightarrow$ a good fit of regression

$$R^2 = \frac{[T \sum XY - (\sum X)(\sum Y)]^2}{[T \sum X^2 - (\sum X)^2][T \sum Y^2 - (\sum Y)^2]} \quad \left. \vphantom{\frac{[T \sum XY - (\sum X)(\sum Y)]^2}{[T \sum X^2 - (\sum X)^2][T \sum Y^2 - (\sum Y)^2]}} \right\} R^2 \text{ in raw data form}$$

$$R^2 = r_{\hat{y}y}^2$$

$\sqrt{r_{\hat{y}y}^2} \Rightarrow$ does not lie in interval $[-1, 1]$

$$\text{but } 0 \leq \sqrt{r_{\hat{y}y}^2} \leq 1$$

Since y_t & $\hat{y}_t$ cannot move in opposite directions

$\sqrt{r_{xy}^2} \Rightarrow$ is in interval $[-1, 1]$

Finite sample Properties of $\hat{\beta}_0$ & $\hat{\beta}_1$

define

$$a_{t1} = \frac{x_t}{\sum x_t^2}$$

$$\hat{\beta}_1 = \frac{\sum x_t y_t}{\sum x_t^2} = \sum_t a_{t1} y_t$$

$$x_t = x_t - \bar{x} \quad y_t = y_t - \bar{y}$$

$$\begin{aligned}\hat{\beta}_1 &= \sum a_{t1} (Y_t - \bar{Y}) = \sum a_{t1} Y_t - \bar{Y} \sum a_{t1} \\ &= \sum a_{t1} Y_t \quad (14)\end{aligned}$$

$$\text{Since } \sum x_t = 0 \Rightarrow \sum a_{t1} = 0$$

For $\hat{\beta}_0$ substitute (14) into (3) eqn

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\begin{aligned}\hat{\beta}_0 &= \frac{\sum Y_t}{T} - \hat{\beta}_1 \left(\frac{\sum x_t}{T} \right) \\ &= \sum Y_t T^{-1} - \hat{\beta}_1 \sum x_t T^{-1} \\ &= \sum Y_t T^{-1} - \sum_t a_{t1} Y_t \left(\sum x_t T^{-1} \right) \\ &= \sum Y_t T^{-1} - \sum_t a_{t1} Y_t \bar{X} \\ &= \sum Y_t T^{-1} - \bar{X} \sum a_{t1} Y_t \\ &= \sum Y_t \underbrace{(T^{-1} - \bar{X} a_{t1})}_{a_{t0}}\end{aligned}$$

$$\hat{\beta}_0 = \sum a_{t0} Y_t \quad \text{where } a_{t0} = T^{-1} - \bar{X} a_{t1}$$

We show that $\hat{\beta}_0$ & $\hat{\beta}_1$ belong to set of linear estimators

For unbiasedness, let define

$$\tilde{\beta}_i = \sum c_{ti} Y_t$$

be any linear estimator of two regression coefficients

$$Y_t = \beta_0 + \beta_1 X_t + U_t \quad \text{into } \tilde{\beta}_i$$

$$\tilde{\beta}_i = \sum c_{ti} (\beta_0 + \beta_1 X_t + U_t)$$

$$= \sum c_{ti} \beta_0 + \sum c_{ti} \beta_1 X_t + \sum \cancel{c_{ti}}^0 U_t$$

$$E(\tilde{\beta}_i) = \beta_0 \sum c_{ti} + \beta_1 \sum c_{ti} X_t$$

In order for $\tilde{\beta}_i$ to be unbiased

$$\begin{aligned}\sum c_{t0} &= 1 & \sum c_{t0} x_t &= 0 \\ \sum c_{t1} &= 0 & \sum c_{t1} x_t &= 1\end{aligned}$$

$$\sum a_{t1} = \frac{\sum x_t}{\sum x_t^2} = 0$$

$$\sum a_{t1} x_t = \sum a_{t1} x_t = \frac{\sum x_t^2}{\sum x_t^2} = 1$$

$$\text{since } \sum_t a_{t1} = \frac{\sum x_t}{\sum x_t^2} \Rightarrow \sum a_{t1} x_t = \frac{\sum x_t^2}{\sum x_t^2} = 1$$

$$\text{and } \sum a_{t0} = 1 - \bar{x} \underbrace{\sum a_{t1}}_0 = 1$$

$$\begin{aligned}a_{t0} x_t &= \bar{x} (x_t) + (T^{-1} - \bar{x} a_{t1}) \\ &= \bar{x} - \bar{x} \underbrace{\sum a_{t1} x_t}_0 = 0\end{aligned}$$

a_{ti} satisfy conditions in (LS)

Thus, LS estimators are both linear and unbiased estimators

$$E(\hat{\beta}_1) = \beta_1 \quad E(\hat{\beta}_0) = \beta_0$$

For efficiency we find expressions for any linear unbiased estimator of β_i 's

by using the expression given in LS

$$\tilde{\beta}_i = \beta_i + \sum c_{ti} u_t$$

$$\text{Var } \tilde{\beta}_i = \sigma_{\tilde{\beta}_i}^2 = E[(\tilde{\beta}_i - \beta_i)]^2 = E[(\sum c_{ti} u_t)^2]$$

$$E(u_t^2) = \sigma^2 \quad \text{and} \quad \text{var}(\tilde{\beta}_i) = \sigma^2 \sum c_{ti}^2$$

$$\sigma_{\hat{\beta}_i}^2 = \sigma^2 \sum a_{ti}^2 = \sigma^2 \frac{\sum_t x_t^2}{(\sum_t x_t^2)^2} = \frac{\sigma^2}{\sum x_t^2}$$

$$\begin{aligned} \sigma_{\hat{\beta}_i}^2 &= \sigma^2 \sum a_{ti}^2 = \sigma^2 \sum_t (T^{-1} - \bar{x} a_{ti})^2 \\ &= \sigma^2 (T^{-1} + \bar{x}^2 \sum a_{ti}^2) \\ &= \sigma^2 (T^{-1} + \frac{(\bar{x}^2)}{\sum x_t^2}) \end{aligned}$$

$$\text{since } \sigma^2 \sum_t a_{ti}^2 = \frac{\sigma^2}{\sum x_t^2}$$

$$\sigma_{\hat{\beta}_i}^2 = \sigma^2 \frac{\sum x_t^2}{T \sum x_t^2}$$

$$\text{since } \sum x_t^2 = \sum x_t^2 - T \bar{x}^2$$

consider the difference bw $\tilde{\beta}_i$ & $\hat{\beta}_i$

$\tilde{\beta}_i$ is any estimator, I try to show that $\hat{\beta}_i$ have variance as at least as any other estimator, if I can show this I end up with efficiency property

$$\tilde{\beta}_i - \hat{\beta}_i = \sum \underbrace{(\sum c_{ti} - a_{ti})}_{d_t} y_t = \sum d_{ti} y_t$$

$$\sigma_{\tilde{\beta}_i}^2 = \underbrace{\sigma^2 \sum a_{ti}^2}_{\sigma_{\hat{\beta}_i}^2} + \sigma^2 \sum d_{ti}^2 + 2\sigma^2 \sum a_{ti} d_{ti}$$

$$\begin{aligned} \sum a_{ti} d_{ti} &= \sum (T^{-1} - \bar{x} a_{ti}) d_{ti} \\ &= T^{-1} \sum d_{ti} - \bar{x} \sum a_{ti} d_{ti} \end{aligned}$$

But for unbiasedness

$$\sum d_{t0} = \sum a_{t0} - \sum c_{t0} = (-1) = 0 \text{ for unbiasedness}$$

$$\sum a_{t1} d_{t0} = \frac{(\sum x_t c_{t0} - \sum x_t a_{t0})}{\sum_t x_t^2} = 0$$

$$\text{since } \sum x_t c_{t0} = \frac{\sum x_t c_{t0}}{1} - \bar{X} \frac{\sum_t c_{t0}}{1} = -\bar{X}$$

$$\sum x_t a_{t0} = -\bar{X}$$

$$\sum a_{t0} d_{t0} = 0$$

finally

$$a_{t1} d_{t1} = \frac{(\sum c_{t1} x_t - \sum a_{t1} x_t)}{\sum x_t^2} = 0$$

Thus,

$$\sigma_{\hat{\beta}_i}^2 - \sigma_{\hat{\beta}_i}^2 = \sigma^2 \sum d_{ti}^2 \geq 0$$

this implies that $\hat{\beta}_i$ is at least as efficient as any other linear unbiased estimator of β_i

Difference between $\sigma_{\hat{\beta}_i}^2$ & $\sigma_{\hat{\beta}_i}^2$ will be zero only when $d_{ti} = c_{ti} - a_{ti} = 0$

that is only when estimator is LS estimator

$\hat{\beta}_i$ are best efficient, linear and unbiased estimators (BLUE)

This result is known as Gauss-Markov Theorem & least-square estimators are robust!

Estimating σ^2

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t$$

If disturbances u_t were observable a quadratic unbiased estimator of σ^2 would be obtained as

$$s^2 = \frac{\sum_t (u_t - \bar{u})^2}{(T-1)}$$

Since u_t 's are unobservable we need sample counterparts of them. Suppose residuals

$\hat{u}_t$ qualify as these counterparts

$$\hat{Y}_t - Y_t = \hat{u}_t$$

Substitute these into (16)

$$\sum_t (\hat{u}_t - \bar{\hat{u}})^2 = \sum_t \hat{u}_t^2 \quad \text{since} \quad \bar{\hat{u}} = T^{-1} \underbrace{\sum_t \hat{u}_t}_{=0} = 0$$

$$\hat{\sigma}^2 = \frac{\sum_t \hat{u}_t^2}{T-2}$$

$$\hat{\sigma}^2 = \frac{\sum_t \hat{u}_t^2}{(T-k-1)}$$

$T \rightarrow$ number of observations
 $k \rightarrow$ number of explanatory (independent) variables

$$k=1 \rightarrow (T-2)$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t$$

$$(T-3) \rightarrow k=2$$

Let prove that $\hat{\sigma}^2$ is an unbiased estimator of σ^2

$$\bar{\hat{u}} = \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x}$$

Subtract this from

$$\hat{u}_t = y_t - \hat{y}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t$$

$$\hat{u}_t = y_t - \hat{\beta}_1 x_t \quad y_t = y_t - \bar{y} \\ x_t = x_t - \bar{x}$$

Substitute for $y_t = \beta_1 x_t + u_t - \bar{u}$

$$\hat{u}_t = -(\hat{\beta}_1 - \beta_1) x_t + u_t - \bar{u}$$

$$\sum_t \hat{u}_t^2 = (\hat{\beta}_1 - \beta_1)^2 \sum_t x_t^2 + \sum_t (u_t - \bar{u})^2 - 2 \sum_t x_t (\hat{\beta}_1 - \beta_1) (u_t - \bar{u})$$

take the expected value of $\hat{u}_t^2$,

$$(17) E[(\hat{\beta}_1 - \beta_1)^2] \sum_t x_t^2 = \sigma_{\hat{\beta}_1}^2 \sum_t x_t^2 = \sigma^2$$

$$(18) E[E(u_t - \bar{u})^2] = (T-1) E(s^2) = (T-1) \sigma^2$$

since s^2 is an unbiased estimator of σ^2

finally

$$-2 E[\sum_t x_t (\hat{\beta}_1 - \beta_1) (u_t - \bar{u})] = 2 \sum_t x_t (\hat{\beta}_1 - \beta_1) u_t$$

$$\text{since } E[(\hat{\beta}_1 - \beta_1) \bar{u}] \sum_t x_t = 0$$

$$\text{But } E[(\hat{\beta}_1 - \beta_1) u_t] = \sum_s \sigma_{s1} E(u_s u_t) = \sigma^2 a_{t1}$$

$$(19) 2 E \sum_t x_t (\hat{\beta}_1 - \beta_1) (u_t - \bar{u}) = 2 \sigma^2 \underbrace{\sum_t a_{t1} x_t}_{=1} = 2 \sigma^2$$

by combining 17-18-19 we can write

$$E(\hat{\sigma}^2) = \frac{\sigma^2 [1 + (T-1) - 2]}{T-2}$$

$$= \underbrace{[(T-2) / (T-2)]}_{1''} \sigma^2$$

$$E(\hat{\sigma}^2) = \sigma^2$$

thus we prove that $\hat{\sigma}^2$ is an unbiased estimator of σ^2

we may write numerator of $\hat{\sigma}^2$ as

$$\begin{aligned}\sum_t \hat{u}_t^2 &= \sum_t y_t^2 - \sum_t \hat{y}_t^2 = \sum_t y_t^2 - \sum_t y_t \hat{y}_t \\ &= \sum_t y_t^2 - \hat{\beta}_1 \sum_t y_t x_t\end{aligned}$$

This is convenient for computations since it does not require separate calculations of each residuals

Moreover by using $\hat{\sigma}^2$ we can obtain estimators for $\sigma_{\beta_0}^2$ & $\sigma_{\beta_1}^2$

$$\sigma_{\beta_0}^2 = \frac{\hat{\sigma}^2 \sum_t x_t^2}{T \sum_t x_t^2} = \frac{(\sum_t \hat{u}_t^2) (\sum_t x_t^2)}{T(T-2) \sum_t x_t^2}$$

$$\sigma_{\beta_1}^2 = \frac{\hat{\sigma}^2}{\sum_t x_t^2} = \frac{\sum_t \hat{u}_t^2}{(T-2) \sum_t x_t^2}$$