



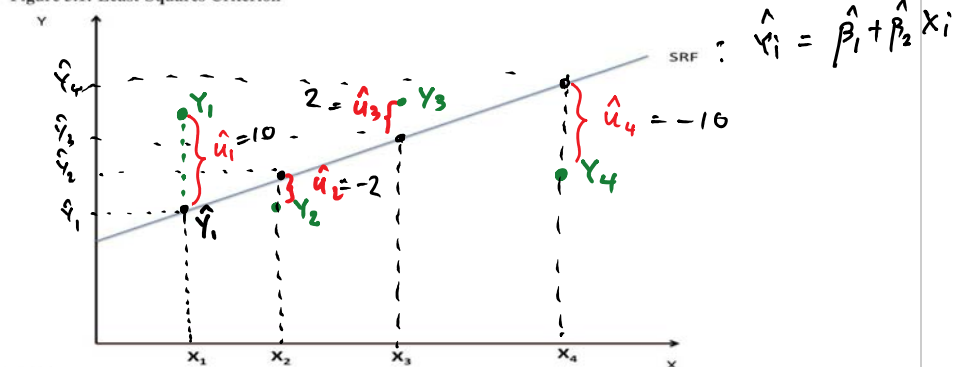
3. REGRESSION: THE PROBLEM OF ESTIMATION

As mentioned in the previous chapter, our main objective is to estimate the population regression function (PRF) based on the basis of the sample regression function (SRF) as accurately as possible.

In this chapter, we are going to discuss the method of estimation: Ordinary Least Squares (OLS)

3.1 The Method of Ordinary Least Squares (OLS)

Figure 3.1: Least-Squares Criterion



$$\begin{aligned}\hat{u}_i &= Y_i - \hat{Y}_i \\ &= Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i) \quad \text{From SRF} \\ &= Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i\end{aligned}$$

CRITERIA TO CHOOSE THE BEST SRF (THAT WE CAN USE TO ESTIMATE PRF):

OPTION 1 $\sum_{i=1}^n \hat{u}_i = 0 \Rightarrow \hat{u}_1 + \hat{u}_2 + \dots + \hat{u}_n = 0$

OPTION 2 $\sum_{i=1}^n \hat{u}_i^2 = 0 \Rightarrow \hat{u}_1^2 + \hat{u}_2^2 + \dots + \hat{u}_n^2 = 0$

We should choose the SRF such that $\left(\sum_{i=1}^n \hat{u}_i^2 \right)$ is as least as possible.
 (sum of squared errors)

Minimize

or sum of errors squared
or sum of residuals squared

3.1 The Method of Ordinary Least Squares (OLS)

49

3.1.1 The Method to Find Out the Least-Squares Estimators: $\hat{\beta}_1$ and $\hat{\beta}_2$

$$\begin{aligned} \text{Minimize}_{\hat{\beta}_1, \hat{\beta}_2} \sum_{i=1}^n \hat{u}_i^2 &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \end{aligned}$$

The method of ordinary least squares (OLS): choose $\hat{\beta}_1$ and $\hat{\beta}_2$ such that, for a given set of data, $\sum_{i=1}^n \hat{u}_i^2$ is as small as possible.

Math box: Do the first order conditions (F.O.C)

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_1} = 0 \quad \text{AND} \quad \frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_2} = 0$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_1} = \frac{\partial \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2}{\partial \hat{\beta}_1} = 2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(-1) = 0 \Rightarrow \sum \hat{u}_i = 0 \quad \text{--- ①}$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_2} = \frac{\partial \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2}{\partial \hat{\beta}_2} = 2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(-X_i) = 0 \Rightarrow \sum \hat{u}_i X_i = 0 \quad \text{--- ②}$$

$$\text{From ①} \quad \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$$

$$\sum_{i=1}^n Y_i - \sum_{i=1}^n \hat{\beta}_1 - \sum_{i=1}^n \hat{\beta}_2 X_i = 0$$

$$\sum_{i=1}^n Y_i - n \cdot \hat{\beta}_1 - \hat{\beta}_2 \sum_{i=1}^n X_i = 0$$

$$n \cdot \hat{\beta}_1 = \sum_{i=1}^n Y_i - \hat{\beta}_2 \sum_{i=1}^n X_i$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i}{n} - \frac{\hat{\beta}_2 \sum_{i=1}^n X_i}{n}$$

$$\boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}}$$

$$\text{From ②} \quad \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(X_i) = 0$$

$$\sum_{i=1}^n X_i Y_i - \hat{\beta}_1 \sum_{i=1}^n X_i - \hat{\beta}_2 \sum_{i=1}^n X_i^2 = 0$$

$$\sum_{i=1}^n X_i Y_i = \hat{\beta}_1 \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2$$

$$\sum_{i=1}^n X_i Y_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2$$

$$\sum_{i=1}^n X_i Y_i = \bar{Y} \sum_{i=1}^n X_i - \hat{\beta}_2 \bar{X} \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2$$

$$\sum_{i=1}^n x_i y_i = \bar{y} \sum_{i=1}^n x_i - \hat{\beta}_2 \bar{x} \sum_{i=1}^n x_i + \hat{\beta}_2 \sum_{i=1}^n x_i^2$$

$$\hat{\beta}_2 \sum_{i=1}^n x_i^2 - \hat{\beta}_2 \bar{x} \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i$$

$$\hat{\beta}_2 \left(\sum_{i=1}^n x_i^2 - \frac{\sum x_i}{n} \sum_{i=1}^n x_i \right) = \sum_{i=1}^n x_i y_i - \frac{\sum y_i}{n} \sum_{i=1}^n x_i$$

$$\hat{\beta}_2 \left(\frac{\sum_{i=1}^n x_i^2 - (\sum x_i)^2}{n} \right) = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum y_i}{n} \sum_{i=1}^n x_i}{n}$$

$$\hat{\beta}_2 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

*

From the SRF:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

Now, we obtain the **least-squares estimators**:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2} \\ &= \bar{Y} - \hat{\beta}_2 \bar{X}\end{aligned}\tag{3.1}$$

$$\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}\tag{3.2}$$

If we define $\bar{X}$ and $\bar{Y}$ to be the sample means of X and Y. Then:

$$\left. \begin{aligned}x_i &= (X_i - \bar{X}) \\ y_i &= (Y_i - \bar{Y})\end{aligned} \right\} \text{Deviation form}\tag{3.3}$$

We can have the alternative expressions for $\hat{\beta}_2$:

$$\begin{aligned}\hat{\beta}_2 &= \frac{\sum x_i y_i}{\sum x_i^2} \\ &= \frac{\sum x_i y_i}{\sum X_i^2 - n \bar{X}^2} \\ &= \frac{\sum X_i y_i}{\sum X_i^2 - n \bar{X}^2}\end{aligned}\tag{3.4}$$

$$x_i = X_i - \bar{X}$$

$$y_i = Y_i - \bar{Y}$$

3.1 The Method of Ordinary Least Squares (OLS)

51

Show that

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\hat{\beta}_2 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \bar{y})}{\sum [x_i^2 - 2x_i \bar{x} + (\bar{x})^2]}$$

$$= \frac{\sum x_i y_i - \bar{y} \sum x_i - \bar{x} \sum y_i + \sum \bar{x} \bar{y}}{\sum x_i^2 - 2\bar{x} \sum x_i + \sum (\bar{x})^2}$$

$$= \frac{\sum x_i y_i - \frac{\sum y_i}{n} \sum x_i - \frac{\sum x_i}{n} \sum y_i + n \frac{\sum x_i}{n} \frac{\sum y_i}{n}}{\sum x_i^2 - 2 \frac{\sum x_i}{n} \sum x_i + \sum \left[\frac{\sum x_i}{n} \right]^2}$$

$$= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n}$$

$$= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - n \cdot 2 \frac{(\sum x_i)^2}{n} + n \cdot n \frac{(\sum x_i)^2}{n^2}}$$

EXAMPLE

$$\sum x_i^2 - 2 \left(\frac{\sum x_i}{n} \right)^2 + \sum \left(\frac{\sum x_i}{n} \right)^2$$

Table 3.1: A Random Sample From the Population

X	Y
500	390
600	425
700	560
800	575
900	630
1000	679

$$= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad \#$$

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

Table 3.2: Raw Data Based on the Sample Data on Table 3.1

(1) Y_i	(2) X_i	(3) $Y_i X_i$	(4) X_i^2	(5) $x_i = X_i - \bar{X}$	(6) $y_i = Y_i - \bar{Y}$	(7) x_i^2	(8) $x_i y_i$	(9) $\hat{Y}_i$	(10) $\hat{u}_i = Y_i - \hat{Y}_i$	(11) $\hat{Y}_i \hat{u}_i$
390	500	195,000	250,000	-250	-153.17	62,500	38,291.67			
425	600	255,000	360,000	-150	-118.17	22,500	17,725			
560	700	392,000	490,000	-50	16.83	2,500	-841.67			
575	800	460,000	640,000	50	31.83	2,500	1,591.67			
630	900	567,000	810,000	150	86.83	22,500	13,025			
679	1,000	679,000	1,000,000	250	135.83	62,500	33,958.33			
Sum 3,259	4,500	2,548,000	3,550,000	0	0	175,000	103,750			
Mean 543.17	750	424,666.67	591,666.67	0	0	29,166.67	17,291.67			

52

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{103,750}{175,000} = 0.5929 \neq$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = 543.17 - 0.5929 \cdot 750 = 98.5238 \neq$$

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \rightarrow \text{SRF}$$

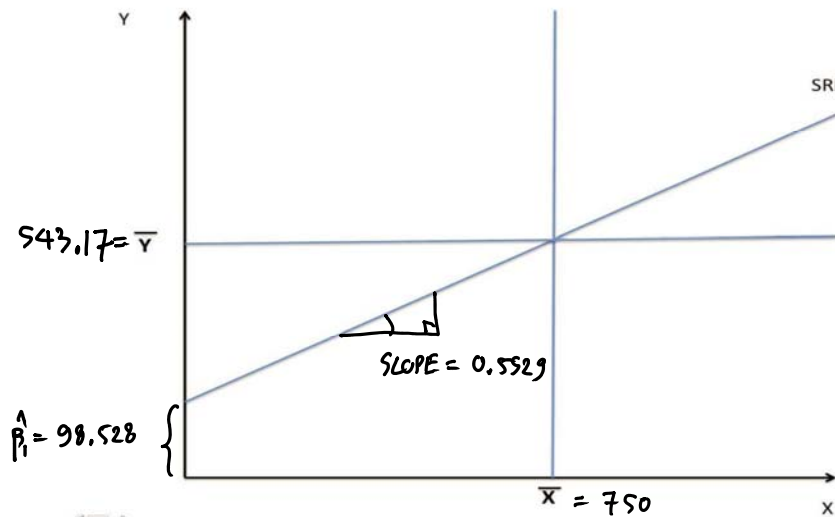
$$\text{Alternatively, } \hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{6 \cdot 2,548,000 - (4,500)(3,259)}{6 \cdot 3,550,000 - (4,500)^2}$$

$$= \frac{622,500}{1,050,000} = 0.59285$$

3.1 The Method of Ordinary Least Squares (OLS)

53

Figure 3.2: Sample Regression Line Based on the Data of Table 3.2



$$\hat{Y}_i = 98.5238 + 0.5529 X_i$$

Notice that the SRF passes through $(\bar{X}, \bar{Y})$.

sample mean of X sample mean of Y.

3.1.2 The numerical and statistical properties of OLS estimators

1. The OLS estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ are expressed solely in terms of the observable (Sample size) and quantities (i.e X and Y).

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2} \\ &= \bar{Y} - \hat{\beta}_2 \bar{X}\end{aligned}$$

(3.5)

$$\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

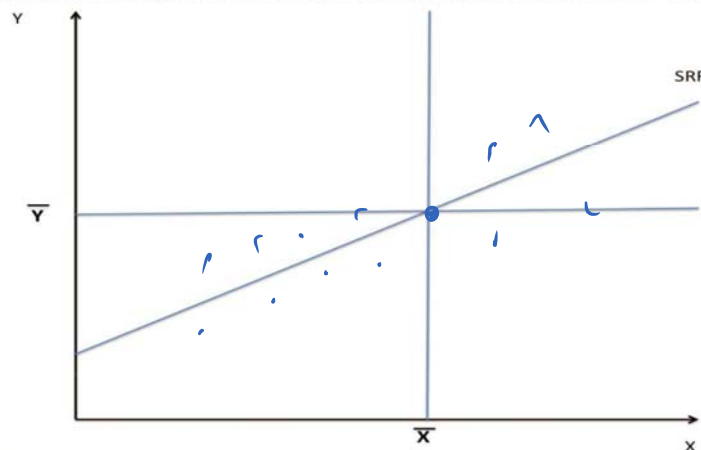
(3.6)

2. They are **point estimators**.

3. The regression line has the following properties.

3.1 The sample regression function (SRF) passes through the sample means of Y and X ($\bar{Y}$ and $\bar{X}$).

Figure 3.3: The Sample regression Line Passes through the Sample Mean Values of Y and X



$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 \hat{X}$$

$$\begin{matrix} Y_1 & \hat{Y}_1 \\ \vdots & \vdots \\ Y_n & \hat{Y}_n \end{matrix}$$

3.2 The mean value of the estimated $Y = \hat{Y}_i$ is equal to the mean value of the actual Y.

$$\overline{\hat{Y}_i} = \bar{Y}$$

Proof:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

$$\hat{Y}_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) + \hat{\beta}_2 X_i$$

$$\sum \hat{Y}_i = \sum \bar{Y} + \hat{\beta}_2 \sum (X_i - \bar{X})$$

$$\begin{aligned} &= \sum X_i - \sum \bar{X} \\ &= \sum X_i - n \bar{X} \end{aligned}$$

3.3. The mean value of the residuals $\hat{u}_i$ is zero.

$$\overline{\hat{u}_i} = 0$$

From our F.O.C:

$$-2 \sum (\underbrace{Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i}_{\hat{u}_i}) = 0$$

$$-2 \sum \hat{u}_i = 0$$

$$\begin{aligned} &= \sum X_i - n \bar{X} \\ &= \sum X_i - n \frac{\sum X_i}{n} \\ &= 0. \end{aligned}$$

[Also, $\sum x_i = 0$]

$$\sum y_i = 0$$

$$\text{So, } \overline{\hat{u}_i} = \frac{\sum \hat{u}_i}{n} = 0. \quad \#$$

3.1 The Method of Ordinary Least Squares (OLS)

55

3.4 The residuals $\hat{u}_i$ are uncorrelated with the predicted $\hat{Y}_i$. : $\sum \hat{Y}_i \hat{u}_i = 0$

Proof:

$$\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X} \quad \text{--- ①}$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i \quad \text{--- ②}$$

$$\text{②} - \text{①} \quad Y_i - \bar{Y} = \hat{\beta}_2 (X_i - \bar{X}) + \hat{u}_i \quad \text{--- ③}$$

$$\hat{Y}_i = \hat{\beta}_2 X_i + \hat{u}_i \quad \text{--- ④}$$

From $\hat{Y}_i = \hat{\beta}_2 X_i$

$$\hat{Y}_i - \hat{u}_i = \hat{\beta}_2 X_i - \hat{u}_i$$

$$\sum \hat{Y}_i \hat{u}_i = \sum \hat{\beta}_2 X_i \hat{u}_i$$

$$= \sum \hat{\beta}_2 X_i (Y_i - \hat{\beta}_2 X_i)$$

$$= \hat{\beta}_2 \left(\sum X_i Y_i - \hat{\beta}_2 \sum X_i^2 \right) = 0!$$

3.5 The residuals $\hat{u}_i$ are uncorrelated with X_i .

$$\sum \hat{u}_i X_i = 0$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_2} = 0 \quad \text{(FROM FOC)}$$

Recall that $\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$

$$\text{So } \sum X_i Y_i = \hat{\beta}_2 \sum X_i^2.$$

3.1.3 The Assumptions Underlying the Method of Least Squares

Assumption 1: Linear regression model

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

OLS

Assumption 2: X values are fixed in repeated sampling

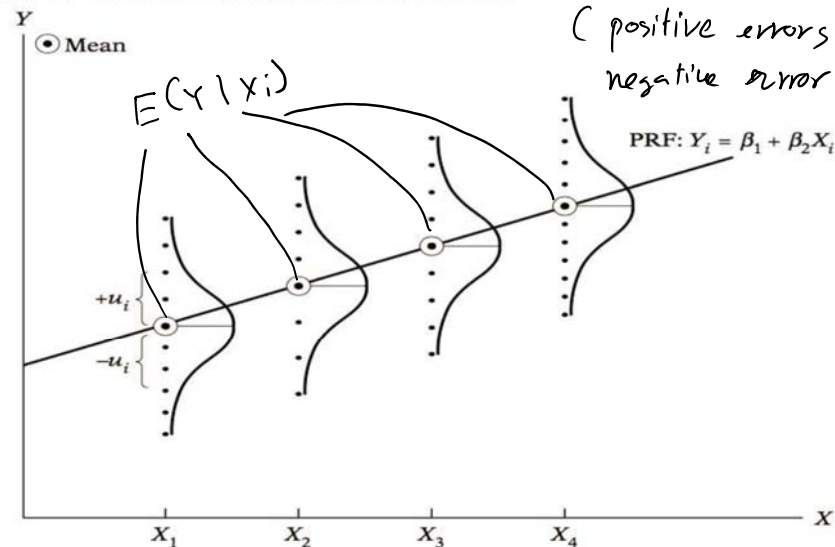
X is assumed to be nonstochastic.

Assumption 3: Zero mean value of disturbance u_i

$$E(u_i | X_i) = 0$$

average value of u_i
evaluated at any $X_i = 0$

(positive errors and
negative error are cancelled
out)

Figure 3.4: Conditional Distribution of the Disturbances u_i 

3.1 The Method of Ordinary Least Squares (OLS)

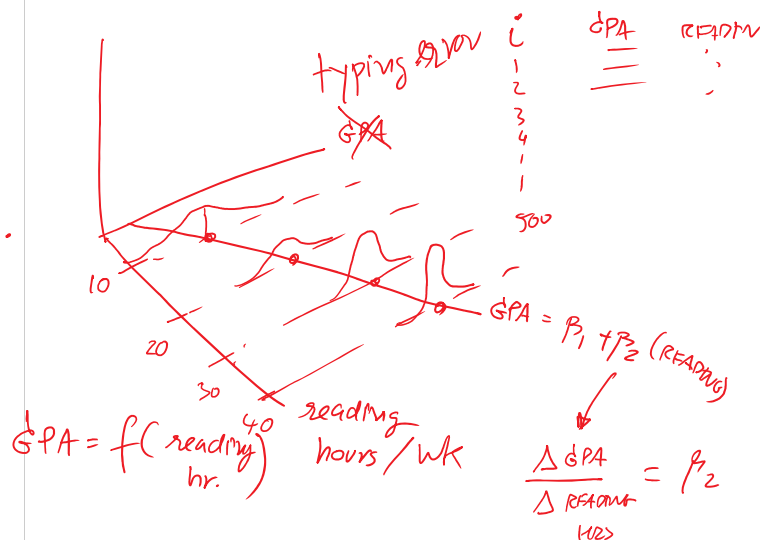
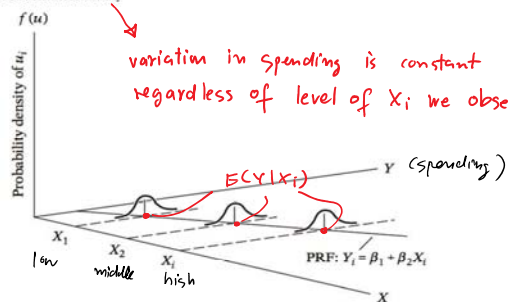
57

Assumption 4: Homoscedasticity or Equal Variance of u_i

$$\text{var}(u_i | x_i) = \sigma^2$$

equal spread

Figure 3.5: Homoscedasticity

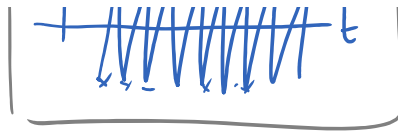


[illegible]

... $\rightarrow \hat{u}_{t-3} \rightarrow u_{t-2} \rightarrow \hat{u}_{t-1} \rightarrow \hat{u}_t$ $10 \times 12 = 120$

when $0 < \rho \leq 1 \rightarrow$ Positive autocorrelation
 when $\rho = 0 \rightarrow$ No autocorrelation
 when $-1 \leq \rho < 0 \rightarrow$ Negative autocorrelation

ch3 Page 13



vice versa.

Residual Plot → Available in EViews

$$-\infty \leq \text{cov} \leq \infty$$

statistical measure of dependence between two random variables.

3.1 The Method of Ordinary Least Squares (OLS)

59

Assumption 6: Zero Covariance Between u_i and X_i : $\text{cov}(u_i, X_i) = 0$

In other words, variation in X_i is not related with variation in u_i .

other measure of quantifying dependence is

"correlation coefficient"

$$(\rho_{xy})$$

$$-1 \leq \rho_{xy} \leq 1$$

$$\rho_{xy} = \frac{\text{cov}(x, y)}{s_x s_y}$$

This one is dimensionless.

Ex:

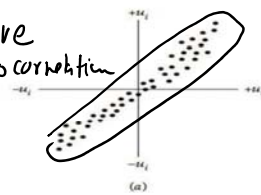
$$\begin{matrix} \text{(Spend)} & & \\ Y_i & & \hat{Y}_i \\ \vdots & & \vdots \\ 500 & & 500 \end{matrix} \quad \hat{u}_i = Y_i - \hat{Y}_i$$

scatter plot

Figure 3.7: Patterns of Correlation Among the disturbances

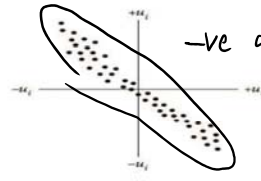
(RELATED w/ ASS.#5)

+ve autocorrelation



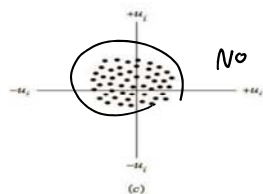
(a)

-ve autocorrelation



(b)

No autocorrelation



(c)

Assumption 7: The number of observations n must be greater than the number of parameters to be estimated.

Ex: $\beta_1, \beta_2 \xrightarrow{\text{estimated by}} \hat{\beta}_1, \hat{\beta}_2$ (# parameters = 2)

so n must be greater than 2

60

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

Assumption 8: Variability in X values.

$$\text{Var}(X) = \frac{\sum (X_i - \bar{X})^2}{n} \neq 0$$

EX: X_i 500 600 700 800 ... 1000 (income groups)

Assumption 9: The regression model is correctly specified.

we will see this after midterm.

In EViews: Misspecification Test

In Gujarati's Text: Topic of

Assumption 10: There is no perfect multicollinearity.

SPENDING INCOME WEALTH

$$EX: Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3$$

However, if

$$X_3 = 2X_2$$

i	X_2	X_3
1	100	200
2	200	400
3	300	600
4	400	800
...	...	...

Perfect
Multicollinearity

$$Y_i = \beta_1 + \beta_2 X_2 + \beta_3 (2X_2) = \beta_1 + (\beta_2 + 2\beta_3) X_2$$

Consequence: β_2 and β_3 cannot be estimated!!!

	X_2	X_3	X_3
1	100	200	$2X_2 + 0$
2	200	500	$2X_2 + 100$
3	300	600	$2X_2 + 0$
4	400	700	$2X_2 - 100$
5			

 u_i
 0
 +100
 0
 -100

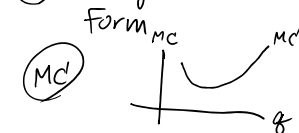
 HERE $X_3 = 2X_2 + u_i$
 imperfect
 multicollinearity

Specification Errors.

① Inclusive Errors

② Exclusive Errors

③ Wrong functional

Q: can we estimate β_2 and β_3 ?A: YES!!! BUT LESS
RELIABILITY...

3.1 The Method of Ordinary Least Squares (OLS)

61

3.1.4 Standard Errors of Least-Squares Estimates

The standard errors of the OLS estimates can be obtained as follows:

We know that

$$\hat{\beta}_2 = \frac{\sum x_i Y_i}{\sum x_i^2} = \sum k_i Y_i$$

where

$$k_i = \frac{x_i}{\sum x_i^2} = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2} \quad \text{where } x_i = X_i - \bar{X}$$

The properties of the weights k_i 1. The k_i are nonstochastic.

$$\sum k_i = 0$$

2. $\sum k_i = 0$ 3. $\sum k_i^2 = \frac{1}{\sum x_i^2}$ 4. $\sum k_i x_i = \sum k_i X_i = 1$

Since

$$\text{var}(\hat{\beta}_2) = E[\hat{\beta}_2 - E(\hat{\beta}_2)]^2$$

First Step

Find the $E(\hat{\beta}_2)$

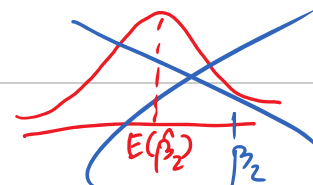
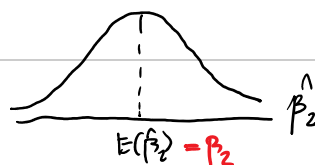
$$\begin{aligned} \hat{\beta}_2 &= \sum_{i=1}^n k_i Y_i \\ &= \sum_{i=1}^n k_i (\beta_1 + \beta_2 X_i + u_i) \\ &= \sum_{i=1}^n (k_i \beta_1 + k_i \beta_2 X_i + k_i u_i) \\ &= \beta_1 \sum_{i=1}^n k_i + \beta_2 \sum_{i=1}^n k_i X_i + \sum_{i=1}^n k_i u_i \\ \hat{\beta}_2 &= \beta_2 + \sum_{i=1}^n k_i u_i \end{aligned}$$

Taking Expectation throughout gives

$$E(\hat{\beta}_2) = \beta_2 + E\left(\sum_{i=1}^n k_i u_i\right)$$

$$\text{so, } E(\hat{\beta}_2) = \beta_2$$

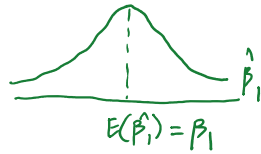
$$\sum_{i=1}^n E(k_i u_i) = 0$$

Therefore $\hat{\beta}_2$ is an unbiased estimator of β_2 . # 😊

By the same manner, we can also prove that $E(\hat{\beta}_1) = \beta_1$. # 😊

By the same manner, we can also prove that $E(\hat{\beta}_1) = \beta_1$. #

O-I-Y



In repeated sampling,
average value of $\hat{\beta}_1$ will be
equal to β_1 (the true but unknown parameter)

$$\begin{aligned}\hat{\beta}_1 &= \bar{Y} - \hat{\beta}_2 \bar{X} \\ &= \bar{Y} - \sum_{i=1}^n k_i Y_i \bar{X} \\ &= \frac{\sum_{i=1}^n Y_i}{n} - \bar{X} \sum_{i=1}^n k_i Y_i \\ &= \sum_{i=1}^n \left[\frac{1}{n} - \bar{X} k_i \right] Y_i \\ &= \sum_{i=1}^n \left[\frac{1}{n} - \bar{X} k_i \right] [\beta_1 + \beta_2 X_i + u_i]\end{aligned}$$

DIY

62

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

Second Step

Using the definition of variance

$$\text{var}(\hat{\beta}_2) = E[\hat{\beta}_2 - E(\hat{\beta}_2)]^2$$

$$\text{var}(\hat{\beta}_2) = E(\hat{\beta}_2 - \beta_2)^2$$

Recall that $\hat{\beta}_2 = \beta_2 + \sum k_i u_i$. So $\hat{\beta}_2 - \beta_2 = \sum k_i u_i$.

$$\text{So } \text{var}(\hat{\beta}_2) = E\left[\sum_{i=1}^n k_i u_i\right]^2$$

$$= E[k_1 u_1 + k_2 u_2 + \dots + k_n u_n][k_1 u_1 + k_2 u_2 + \dots + k_n u_n]$$

$$= E[k_1^2 u_1^2 + k_2^2 u_2^2 + \dots + k_n^2 u_n^2 + 2k_1 k_2 u_1 u_2 + \dots + 2k_{n-1} k_n u_{n-1} u_n]$$

$$= k_1^2 E(u_1^2) + k_2^2 E(u_2^2) + \dots + k_n^2 E(u_n^2)$$

cross terms:

$$2k_i k_j u_i u_j$$

Since $E(u_i u_j) = 0$ by assumption of no autocorrelation

Unfortunately,

 σ_u^2 is unknown.

So we have to find a recipe to "ESTIMATE" IT.

The covariance between $\hat{\beta}_1$ and $\hat{\beta}_2$

$$\text{cov}(\hat{\beta}_1, \hat{\beta}_2) = E[(\hat{\beta}_1 - E(\hat{\beta}_1))(\hat{\beta}_2 - E(\hat{\beta}_2))]$$

$$= E(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2)$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma_u^2}{\sum x_i^2}$$

$$\text{From } \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} \quad (1)$$

$$E(\hat{\beta}_1) = \bar{Y} - \beta_2 \bar{X} \quad (2)$$

$$(1) - (2) \quad \hat{\beta}_1 - E(\hat{\beta}_1) = -\bar{X}(\hat{\beta}_2 - \beta_2)$$

$$\text{So, } \text{cov}(\hat{\beta}_1, \hat{\beta}_2) = E[-\bar{X}(\hat{\beta}_2 - \beta_2)(\hat{\beta}_2 - \beta_2)]$$

$$= -\bar{X} E[(\hat{\beta}_2 - \beta_2)^2]$$

$$= -\bar{X} \cdot \text{var}(\hat{\beta}_2)$$

$$\text{cov}(\hat{\beta}_1, \hat{\beta}_2) = -\bar{X} \cdot \frac{\sigma_u^2}{\sum x_i^2}$$

$$H: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

$$t = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)}$$

3.1 The Method of Ordinary Least Squares (OLS)

63

3.1.5 The Least-Square Estimator of σ^2

$$\hat{\sigma}_u^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

see appendix 3A.5 (page 102-103)
in Gujarati

$\hat{\sigma}_u^2$ is an "UNBIASED" ESTIMATOR OF TRUE σ_u^2 .

In sum, the standard errors of the OLS estimators can be obtained as follow:

$$\begin{aligned}\text{var}(\hat{\beta}_2) &= \frac{\sigma^2}{\sum x_i^2} \\ \text{se}(\hat{\beta}_2) &= \frac{\sigma}{\sqrt{\sum x_i^2}}\end{aligned}\tag{3.7}$$

$$\begin{aligned}\text{var}(\hat{\beta}_1) &= \frac{\sum X_i^2}{n \sum x_i^2} \sigma^2 \\ \text{se}(\hat{\beta}_1) &= \sqrt{\frac{\sum X_i^2}{n \sum x_i^2}} \sigma\end{aligned}\tag{3.8}$$

We can estimate the σ^2 from the data where the formula for the estimated σ^2 is following :

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

where

$$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2^2 \sum x_i^2$$

The alternative expression for computing $\sum \hat{u}_i^2$ is

$$\sum \hat{u}_i^2 = \sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2}$$

The covariance between $\hat{\beta}_1$ and $\hat{\beta}_2$ is:

$$\begin{aligned}\text{cov}(\hat{\beta}_1, \hat{\beta}_2) &= -\bar{X} \text{var}(\hat{\beta}_2) \\ &= -\bar{X} \left(\frac{\sigma^2}{\sum x_i^2} \right)\end{aligned}\tag{3.9}$$

3.1.6 Properties of Least-Squares Estimators: The Gauss-Markov Theorem

Given the assumptions of the classical linear regression model, the least-square estimators are satisfied the optimum properties which is known as "The Gauss-Markov Theorem." To understand this theorem, we need to know the small-sample properties of an estimator first.

The Small-Sample Properties of An Estimator**1. Unbiasedness**

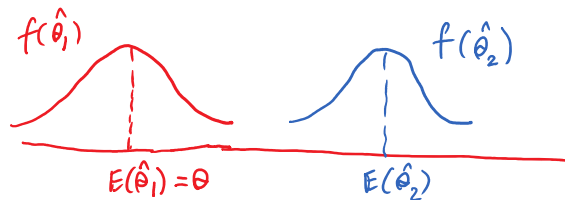
An estimator $\hat{\theta}$ is said to be an unbiased estimator of θ if the expected value of $\hat{\theta}$ is equal to the true θ .

$$E(\hat{\theta}) = \theta$$

Therefore, if the expected value of $\hat{\theta}$ is not equal to the true θ , then the estimator is said to be biased. We can calculate the biased as:

$$\text{bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$$

Figure: Biased and Unbiased Estimators



$\hat{\theta}_1$ is an unbiased estimator of θ .

$\hat{\theta}_2$ is a biased estimator of θ .

↙ b/c we see the difference between $E(\hat{\theta}_2)$ and θ !
i.e., $\text{BIAS}(\hat{\theta}_2) = E(\hat{\theta}_2) - \theta$

OLS is a clever method as its estimators are BLUE estimators!

$$E(\hat{\beta}_2) = \beta_2$$

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

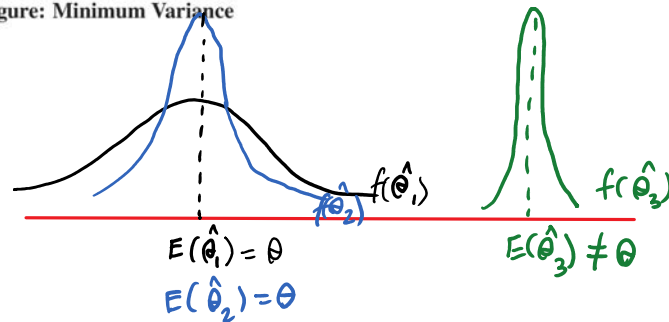
66

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

2. Minimum Variance

$\hat{\theta}_1$ is said to be a minimum variance estimator of θ if the variance of $\hat{\theta}_1$ is smaller ~~than~~ ^{than} or at most equal to the variance of $\hat{\theta}_2$, which is any other estimator of θ

Figure: Minimum Variance



To select a good estimator, unbiasedness must come first!

3. Best Unbiased or Efficient Estimator = property 1+ property 2

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators of θ and the variance of $\hat{\theta}_1$ is smaller than or at most equal to the variance of $\hat{\theta}_2$, then $\hat{\theta}_1$ is a **minimum-variance unbiased estimator** or **best unbiased estimator**.

4. Linearity

An estimator $\hat{\theta}$ is said to be a linear estimator of θ if it is a linear function of the sample observations. For example:

$$\bar{X} = \frac{1}{n} \sum X_i = \frac{1}{n} (X_1 + X_2 + \dots + X_n)$$

Thus, $\bar{X}$ is a linear estimator because it is a linear function of the X values.

3.1 The Method of Ordinary Least Squares (OLS)

67

Best Linear Unbiased Estimators : BLUE

The estimator $\hat{\theta}$ is called as the Best Linear Unbiased Estimator **BLUE** if it is satisfied the properties 1,2,4 that is $\hat{\theta}$ is linear, is unbiased, and has the minimum variance in the class of all linear unbiased estimators of θ .

Minimum Mean-Square-Error (MSE) Estimator

The MSE measures dispersion around the true value of the parameter. It is defined as:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

estimator true parameter

However, the variance of $\hat{\theta}$ measures the dispersion of the distribution of the distribution of $\hat{\theta}$ around its mean or expected value.

$$var(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2$$

The relationship between the MSE($\hat{\theta}$) and the var($\hat{\theta}$) is as follows:

$$\begin{aligned} MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 \\ &= E[(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)]^2 \\ &= E[(\hat{\theta} - E(\hat{\theta}))^2 + 2E[(\hat{\theta} - E(\hat{\theta}))][E(\hat{\theta}) - \theta] + E[E(\hat{\theta}) - \theta]^2] \\ MSE(\hat{\theta}) &= var(\hat{\theta}) + [BIASNESS(\hat{\theta})]^2 \end{aligned}$$

$(a+b)^2 = a^2 + 2ab + b^2$
 $\swarrow$ BIASNESS $E(\hat{\theta}) = \theta$

If BIASNESS = 0, then $MSE(\hat{\theta}) = var(\hat{\theta}) !!!$

In other words, we can estimate $MSE(\hat{\theta})$ by using $var(\hat{\theta})$, given that $E(\hat{\theta}) = \theta$

check point: so far we obtain 6 things + BLUE PROPERTY

$$\min_{\hat{\beta}_1, \hat{\beta}_2} \sum_{i=1}^n \hat{u}_i^2 \longrightarrow \begin{aligned} &- \hat{\beta}_1 \\ &- \hat{\beta}_2 \\ &- var(\hat{\beta}_1) \\ &- var(\hat{\beta}_2) \\ &- se(\hat{\beta}_1) \\ &- se(\hat{\beta}_2) \\ &- cov(\hat{\beta}_1, \hat{\beta}_2) \end{aligned}$$

68

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

An estimator $\hat{\beta}_2$ is said to be a best linear unbiased estimator (BLUE) of β_2 if the following hold:

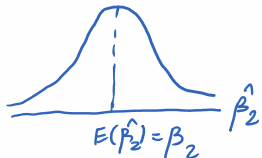
♣ It is linear. It is the linear function of a random variable.

L
$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \sum k_i Y_i \quad \text{where } k_i = \frac{x_i}{\sum x_i^2}$$

random variable

♣ It is unbiased. That is $E(\hat{\beta}_2)$ is equal to the true value, β_2

U
$$E(\hat{\beta}_2) = \beta_2$$



♣ It has the minimum variance in the class of all such linear unbiased estimators.

Suppose we have 2 candidates :

B

① $\hat{\beta}_2 = \sum k_i Y_i$ From method of OLS

② $\hat{\beta}_2^* = \sum w_i Y_i$ from any other methods of finding linear unbiased estimator.

This property says that
$$\text{Var}(\hat{\beta}_2) \leq \text{Var}(\hat{\beta}_2^*)$$

Note: Before we make the comparison above, we must check first if $E(\hat{\beta}_2) = \beta_2$ and $E(\hat{\beta}_2^*) = \beta_2$.

3.1 The Method of Ordinary Least Squares (OLS)

69

Gauss-Markov Theorem: Given the assumptions of the classical linear regression model, the least-squares estimators, in the class of unbiased linear estimators, have minimum variance, that is, they are BLUE.

3.1.7 A measure of goodness of fit: r^2

In this section, we are going to study the goodness of fit of the fitted regression line to a set of data. Let us consider the following example:

Suppose we were to estimate the family expenditure (Y) based on our information from a random sample (as in Table 3.2).

What will happen if we set the estimated Y to be $\bar{Y}$?

Table 3.3: Estimating the expenditure of the household

Family Number (i)	Actual Y_i	Estimate $\hat{Y}_i = \bar{Y}$	Error in Estimation $Y_i - \bar{Y}$	Errors Squared $(Y_i - \bar{Y})^2$
1	390	543	-153	23460.03
2	425	543	-118	13963.36
3	560	543	17	283.36
4	575	543	32	1013.36
5	630	543	87	7540.03
6	679	543	136	18450.69
Sum	3259	3259	0	64710.83

$$\bar{Y} = 543$$

We can see all this graphically:

$$= \sum_{i=1}^n (Y_i - \bar{Y})^2$$

= Total sum of squares (TSS)

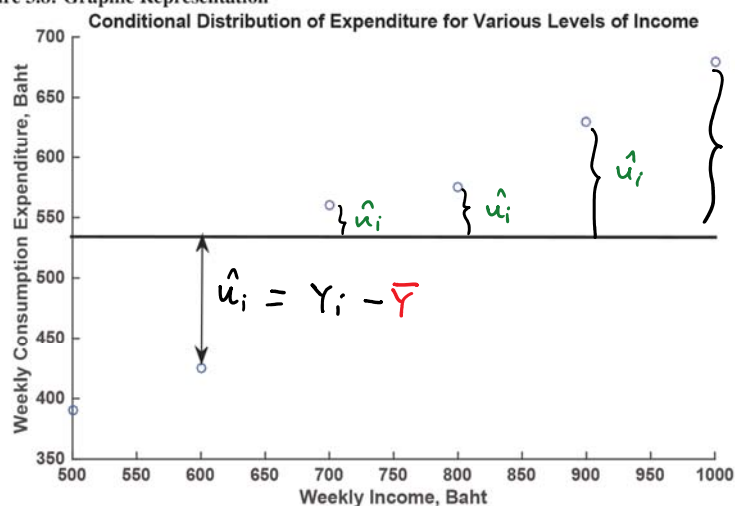
↓

Total estimation errors we committed WHEN we use just sample mean ($\bar{Y}$) as our predictor.

70

Chapter 3. REGRESSION: THE PROBLEM OF ESTIMATION

Figure 3.8: Graphic Representation



Question: Can we determine the total estimation error for this sample data?

Answer: Yes, we can calculate the total (combined) amount of estimation error for all observations in the sample when **using the mean as the estimate** as following:

$$TSS = \sum (Y_i - \bar{Y})^2$$

It is called the total sum of squares (TSS) which is the total variation of the actual Y values about their sample mean.

Since our objective in estimation is to minimize error (maximize precision), we need to cut down the amount of the estimation error (TSS).

We can achieve this by using information about other variables suspected to be strong predictors (strongly related to) the expenditure of the families.

We now can attempt to estimate the expenditure from the information on the income level of the family, rather than from its own mean.

3.1 The Method of Ordinary Least Squares (OLS)

71

Table 3.4: Estimating the expenditure of the household with income

Family (i)	Actual Y_i	Income X_i	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})(Y - \bar{Y})$	$(X - \bar{X})^2$
1	390	500	-250	-153.17	38291.67	62500
2	425	600	-150	-118.17	17725.00	22500
3	560	700	-50	16.83	-841.67	2500
4	575	800	50	31.83	1591.67	2500
5	630	900	150	86.83	13025.00	22500
6	679	1000	250	135.83	33958.33	62500
Sum	3259	4500	0	0	103750	175000

From the table 8, we can calculate the simple regression as following:

$$\hat{\beta}_1 = ? \quad \hat{\beta}_2 = ? \rightarrow \hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{103750}{175000} = 0.5928$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = 543.16 - (0.5928) 750 = 98.524$$

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \rightarrow \hat{Y}_i = 98.524 + 0.5928 X_i$$

$$\text{When } X_i = 500, \hat{Y}_i = 98.524 + 0.5928 (500) = 394.2 \text{ baht}$$

$\hat{\beta}_2 = 0.5928 \rightarrow$ Holding all other factors that may affect expenditure constant, when income rises by 1 baht, expenditure would rise by 0.5928 baht, on average.

Figure 3.9: Breakdown of the variation of Y_i into two components

Conditional Distribution of Expenditure for Various Levels of Income

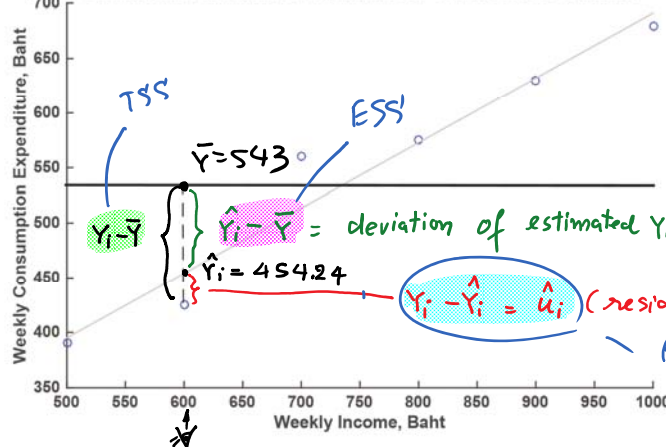


Table 3.5: Estimating the expenditure of the household with income

Family (i)	Actual Y_i	Income X_i	Regression Estimate $\hat{Y}_i$	Residual $Y_i - \hat{Y}_i$	Residual squared $(Y_i - \hat{Y}_i)^2$
1	390	500	394.95	-4.95	24.53
2	425	600	454.24	-29.24	854.87
3	560	700	513.52	46.48	2160.04
4	575	800	572.81	2.19	4.80
5	630	900	632.10	-2.10	4.39
6	679	1000	691.38	-12.38	153.29
Sum	3259	4500	0	0	3201.90

From the table 9, we can calculate the estimation error we have committed by using the regression line as:

$$RSS = \sum (Y_i - \hat{Y}_i)^2 = \sum \hat{u}_i^2$$

where RSS stands for the residual sum of squares, which is the unexplained variation of the Y values about the regression line.

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \text{ or } = \sum_{i=1}^n \hat{u}_i^2 \text{ actually}$$

our remaining errors after we use "income" to help us predict family expenditure 😊

$$TSS \xrightarrow{\text{by OLS}} RSS$$

(64710.83) (3201.90)

when using $\bar{Y}$

$$TSS - RSS = ESS$$

explained sum of squares

3.1 The Method of Ordinary Least Squares (OLS)

73

Total Baseline Error using the mean (SS Total) = $\sum (Y_i - \bar{Y})^2$ New or Remaining Error (SS Error or SS Residual) = $\sum (Y_i - \hat{Y}_i)^2$ **QUESTION:** How much of the original estimation error have we explained away (eliminated) by using the regression model (instead of the mean)?

ANS

$$ESS = TSS - RSS$$

$$= 64710.83 - 3201.90$$

$$= \underline{61508.93} \rightarrow \text{errors that is eliminated by using regression model.}$$

QUESTION: What % of estimation error have we explained (eliminated by using the regression model)?

ANS

$$TSS = ESS + RSS$$

$$64710.83 \quad 61508.93 \quad 3201.90$$

$$\frac{ESS}{TSS} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{61508.93}{64710.83} = 0.9505 \rightarrow \text{we call this } R^2 \text{ (R-squared)}$$

QUESTION: What does the remaining % represent?

ANS

Percent of variation (differences) in expenditures that can be accounted for by: (a) all other potential predictors not included in the model, beyond income levels, and (b) unexplainable random/chance variations.

$$1 - R^2 = 1 - 0.9505 = 0.0495 \text{ or } \underline{4.95\%}$$

(Part of variation in expenditure that is left unexplained as there are some variables apart from income that may affect expenditure BUT we didn't include into our model.)

$$r^2 = \frac{ESS}{TSS} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

♣ r^2 is a measure of our success regarding accuracy of our estimation effort.

♣ r^2 = % of estimation error that we have been able to explain away by using the regression model, instead of using the mean.

♣ r^2 indicates how much better we can predict Y from information about Xs, rather than from using its own mean.

♣ r^2 = % of differences (variations) in Y values that is explained by (attributable to) differences in X values.

Note: with 2-variable regression model, R^2 is equal to correlation coefficient:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\sum (x_i y_i)^2}{\sum x_i^2 \sum y_i^2}$$

let's define him as r^2

correlation coefficient.

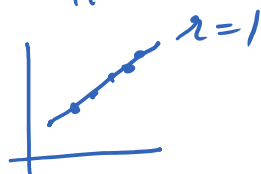
measurement of the degree of association between two variables.

$$r = \pm \sqrt{r^2}$$

$$= \frac{\sum x_i y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}}$$

sample correlation coefficient

$$= \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2] \cdot [n \sum Y_i^2 - (\sum Y_i)^2]}}$$

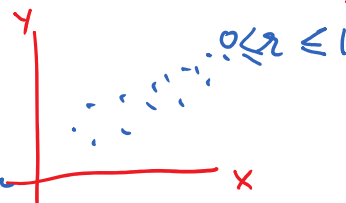


negative correlation

positive correlation



$$\sqrt{[n \sum X_i^2 - (\sum X_i)^2] \cdot [n \sum Y_i^2 - (\sum Y_i)^2]}$$



$$-1 \leq r \leq 1$$



Remark: r tells nothing about "CAUSATION,"

($X \rightarrow Y$)

$$\text{or } Y \rightarrow X$$

$$\text{or } X \leftrightarrow Y$$