ECO 344 - Applied Econometrics I

Chapter 2

Lecture Slides

Muhammad Salman Khalid

School of Economics & Social Sciences

September 16, 2026



Outline

1. The Simple Regression Model
2. Deriving Ordinary Least Square (OLS) Estimates
3. Algebraic Properties of OLS Estimators
4. Units of Measurement & Functional Form
5. Expected Values and Variance of the OLS Estimators
6. Regression Through Origin

The Simple Regression Model

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR
- studying how y varies/changes with the variation/change in x.

$$y = \beta_0 + \beta_1 x + \mu \qquad [2.1]$$

μ = error term

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR
- studying how y varies/changes with the variation/change in x.

$$y = \beta_0 + \beta_1 x + \mu \qquad [2.1]$$

μ = error term

- There is no exact relationship between y & x.

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR
- studying how y varies/changes with the variation/change in x.

$$y = \beta_0 + \beta_1 x + \mu \qquad [2.1]$$

μ = error term

- There is no exact relationship between y & x.
- How do we allow other factors to affect y.

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR
- studying how y varies/changes with the variation/change in x.

$$y = \beta_0 + \beta_1 x + \mu \qquad [2.1]$$

μ = error term

- There is no exact relationship between y & x.
- How do we allow other factors to affect y.
- What is a functional relationship between y & x.

The Simple Regression Model

The Simple Regression Model allows us to study the relationship between two variables.

Let x and y be two variables representing some population. We are interested in;

- explaining y in terms of x OR
- studying how y varies/changes with the variation/change in x.

$$y = \beta_0 + \beta_1 x + \mu \qquad [2.1]$$

μ = error term

- There is no exact relationship between y & x.
- How do we allow other factors to affect y.
- What is a functional relationship between y & x.
- How do we capture ceteris paribus effect (main goal).

The Simple Regression Model

Equation 2.1 represents the simple regression model which is assumed to hold in the population of interest. The simple regression model is between two variables, however later we will extend it to multiple regression model.

The Simple Regression Model

Equation 2.1 represents the simple regression model which is assumed to hold in the population of interest. The simple regression model is between two variables, however later we will extend it to multiple regression model.

Table: Terminology for simple regression model

Y	**X**
Dependent Variable	Independent Variable
Explained Variable	Explanatory Variable
Response Variable	Control Variable
Predicted Variable	Predictor Variable
Regressand	Regressor

The Simple Regression Model

- The variable μ is called the error term or disturbance in the relationship between x and y.

The Simple Regression Model

- The variable μ is called the error term or disturbance in the relationship between x and y.
  - A simple regression treats all other factors affecting y as unobserved and are included in μ.

The Simple Regression Model

- The variable μ is called the error term or disturbance in the relationship between x and y.
  - A simple regression treats all other factors affecting y as unobserved and are included in μ.
- The β_0 is called the intercept parameter or a constant term indicating the average value of y when the value of x is zero.

The Simple Regression Model

- The variable μ is called the error term or disturbance in the relationship between x and y.
  - A simple regression treats all other factors affecting y as unobserved and are included in μ.
- The β_0 is called the intercept parameter or a constant term indicating the average value of y when the value of x is zero.
- The β_1 is called the slope parameter and it shows how one unit change in x, leads to $\hat{\beta}_1$ unit change in y.

$$\frac{d}{dx}y = \beta_1$$

The Simple Regression Model

- The variable μ is called the error term or disturbance in the relationship between x and y.
  - A simple regression treats all other factors affecting y as unobserved and are included in μ.
- The β_0 is called the intercept parameter or a constant term indicating the average value of y when the value of x is zero.
- The β_1 is called the slope parameter and it shows how one unit change in x, leads to $\hat{\beta}_1$ unit change in y.

$$\frac{d}{dx}y = \beta_1$$

- The relationship between y and x is assumed to be linear. It is important to note that it is linear in parameters and not in variables themselves.

Deriving Ordinary Least Square (OLS) Estimates

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want
  1. should we minimize the sum of error?

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want
  1. should we minimize the sum of error?
  2. should we minimize the sum of the absolute value of error?

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want
  1. should we minimize the sum of error?
  2. should we minimize the sum of the absolute value of error?
  3. should we minimize the sum of square error?

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want
  1. should we minimize the sum of error?
  2. should we minimize the sum of the absolute value of error?
  3. should we minimize the sum of square error?

Deriving OLS Estimates

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$

- we want
  1. should we minimize the sum of error?
  2. should we minimize the sum of the absolute value of error?
  3. should we minimize the sum of square error?

$$Minimize_{\beta_0,\beta_1} : \sum_{i=1}^{n} Error_i^2 = \sum_{i=1}^{n} \mu_i^2$$

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} \hat{\mu}_i{}^2$$

As $\hat{\mu}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$, so

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} \hat{\mu}_i{}^2$$

As $\hat{\mu}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$, so

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Taking derivative w.r.t. $\hat{\beta}_0$ and setting it equal to zero to minimize,

$$\sum_{i=1}^{n} -2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} \hat{\mu}_i^2$$

As $\hat{\mu}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$, so

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Taking derivative w.r.t. $\hat{\beta}_0$ and setting it equal to zero to minimize,

$$\sum_{i=1}^{n} -2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Deriving OLS Estimates (continued)

Starting from Equation I,

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Deriving OLS Estimates (continued)

Starting from Equation I,

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Expanding the summation,

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{\beta}_0 - \sum_{i=1}^{n} \hat{\beta}_1 x_i = 0$$

Deriving OLS Estimates (continued)

Starting from Equation I,

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Expanding the summation,

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{\beta}_0 - \sum_{i=1}^{n} \hat{\beta}_1 x_i = 0$$

$$\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

Deriving OLS Estimates (continued)

Starting from Equation I,

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Expanding the summation,

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{\beta}_0 - \sum_{i=1}^{n} \hat{\beta}_1 x_i = 0$$

$$\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

Dividing both sides by n,

$$\bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0$$

Deriving OLS Estimates (continued)

Starting from Equation I,

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \qquad (I)$$

Expanding the summation,

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{\beta}_0 - \sum_{i=1}^{n} \hat{\beta}_1 x_i = 0$$

$$\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

Dividing both sides by n,

$$\bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad (II)$$

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Similarly, taking derivative w.r.t. $\hat{\beta}_1$ and setting it equal to zero to minimize,

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Similarly, taking derivative w.r.t. $\hat{\beta}_1$ and setting it equal to zero to minimize,

$$\sum_{i=1}^{n} -2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0$$

Deriving OLS Estimates

$$Minimize_{\hat{\beta}_0,\hat{\beta}_1} : \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Similarly, taking derivative w.r.t. $\hat{\beta}_1$ and setting it equal to zero to minimize,

$$\sum_{i=1}^{n} -2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0$$

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \qquad (III)$$

Deriving OLS Estimates

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \qquad (III)$$

Putting the value of $\hat{\beta}_0$ from Equation II, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$,

Deriving OLS Estimates

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \qquad (III)$$

Putting the value of $\hat{\beta}_0$ from Equation II, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$,

$$\sum_{i=1}^{n}(y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}))x_i = 0$$

Deriving OLS Estimates

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \qquad (III)$$

Putting the value of $\hat{\beta}_0$ from Equation II, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$,

$$\sum_{i=1}^{n}(y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}))x_i = 0$$

Distributing x_i and splitting the sum,

$$\sum_{i=1}^{n}(y_i - \bar{y})x_i - \sum_{i=1}^{n}\hat{\beta}_1(x_i - \bar{x})x_i = 0$$

Deriving OLS Estimates (continued)

Starting from where we left off,

$$\sum_{i=1}^{n}(y_i - \bar{y})x_i - \sum_{i=1}^{n}\hat{\beta}_1(x_i - \bar{x})x_i = 0$$

Deriving OLS Estimates (continued)

Starting from where we left off,

$$\sum_{i=1}^{n}(y_i - \bar{y})x_i - \sum_{i=1}^{n}\hat{\beta}_1(x_i - \bar{x})x_i = 0$$

Solving for $\hat{\beta}_1$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})x_i}$$

Deriving OLS Estimates (continued)

Starting from where we left off,

$$\sum_{i=1}^{n}(y_i - \bar{y})x_i - \sum_{i=1}^{n}\hat{\beta}_1(x_i - \bar{x})x_i = 0$$

Solving for $\hat{\beta}_1$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})x_i}$$

Using the fact that $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$, this is algebraically the same as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (IV)$$

Deriving OLS Estimates

Equations II and IV give us $\hat{\beta}_0$ and $\hat{\beta}_1$, respectively, minimizing the error term.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})x_i} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (IV)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \qquad (II)$$

Deriving OLS Estimates

Equations II and IV give us $\hat{\beta}_0$ and $\hat{\beta}_1$, respectively, minimizing the error term.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})x_i} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (IV)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \qquad (II)$$

It can be shown by putting II & IV into the second derivative w.r.t $\hat{\beta}_0$ and $\hat{\beta}_1$ that they do minimize the error and not maximize it. (Do it at Home)

Interpretation

- $\hat{\beta}_0$ represents the average value of y given x is zero. In our wage, education example $\hat{\beta}_0$ will be the average wage of the workers without any formal education.

Interpretation

- $\hat{\beta}_0$ represents the average value of y given x is zero. In our wage, education example $\hat{\beta}_0$ will be the average wage of the workers without any formal education.

- $\hat{\beta}_1$ represents the change in y with the 1 unit change in x keeping all other factors constant. In our wage, education example one extra year of schooling will lead to $\hat{\beta}_1$ PKR change in wage keeping all other factors constant.

Algebraic Properties of OLS Estimators

Fitted Values & Residuals

Now, we have a formula to estimate $\hat{\beta}_0$ and $\hat{\beta}_1$, we can predict the dependent variable given independent variable.

Going back to our wage and education problem; Suppose we have following relationship:

$$wage = 45000 + 9000 Education + \mu$$

What will be the interpretation of this equation?

- Not only, we can tell the relationship of y and x, we can also predict the wage of an unknown individual given his/her years of schooling?

Fitted Values & Residuals

Now, we have a formula to estimate $\hat{\beta}_0$ and $\hat{\beta}_1$, we can predict the dependent variable given independent variable.

Going back to our wage and education problem; Suppose we have following relationship:

$$wage = 45000 + 9000 Education + \mu$$

What will be the interpretation of this equation?

- Not only, we can tell the relationship of y and x, we can also predict the wage of an unknown individual given his/her years of schooling?
- What will be the wage of a PhD (23 years of schooling), according to this model? - Prediction

Fitted Values & Residuals

Now, we have a formula to estimate $\hat{\beta}_0$ and $\hat{\beta}_1$, we can predict the dependent variable given independent variable.

Going back to our wage and education problem; Suppose we have following relationship:

$$wage = 45000 + 9000 Education + \mu$$

What will be the interpretation of this equation?

- Not only, we can tell the relationship of y and x, we can also predict the wage of an unknown individual given his/her years of schooling?
- What will be the wage of a PhD (23 years of schooling), according to this model? - Prediction
- what if we use the model to find the y values (used to estimate model) we already have? - **Fitted Values.**

Algebraic Properties of OLS Estimates

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Where $\hat{y}_i$ is known as fitted value or predicted value.
Similarly,

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Where $\hat{\mu}_i$ is known as **residual** or error.

Algebraic Properties of OLS Estimates

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Where $\hat{y}_i$ is known as fitted value or predicted value.
Similarly,

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Where $\hat{\mu}_i$ is known as **residual** or error.
By Definition;

1. $\sum_{i=1}^{n} \hat{\mu}_i = 0$ From Equation I

Algebraic Properties of OLS Estimates

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Where $\hat{y}_i$ is known as fitted value or predicted value.
Similarly,

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Where $\hat{\mu}_i$ is known as **residual** or error.
By Definition;

1. $\sum_{i=1}^{n} \hat{\mu}_i = 0$ From Equation I
2. $\sum_{i=1}^{n} \hat{\mu}_i x_i = 0$ From Equation III

Algebraic Properties of OLS Estimates

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Where $\hat{y}_i$ is known as fitted value or predicted value.
Similarly,

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Where $\hat{\mu}_i$ is known as **residual** or error.
By Definition;

1. $\sum_{i=1}^{n} \hat{\mu}_i = 0$ From Equation I
2. $\sum_{i=1}^{n} \hat{\mu}_i x_i = 0$ From Equation III
3. $\bar{y} = \bar{\hat{Y}}$ From Equation II

Algebraic Properties of OLS Estimates

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Where $\hat{y}_i$ is known as fitted value or predicted value.
Similarly,

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Where $\hat{\mu}_i$ is known as **residual** or error.
By Definition;

1. $\sum_{i=1}^{n} \hat{\mu}_i = 0$ From Equation I
2. $\sum_{i=1}^{n} \hat{\mu}_i x_i = 0$ From Equation III
3. $\bar{y} = \bar{\hat{Y}}$ From Equation II
4. SST = SSE + SSR

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

By Definition SST = SSE + SSR - Proof in Book*

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- R-Squared is the fraction of variation explained by our model.

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- R-Squared is the fraction of variation explained by our model.
- Variation in Y that can be explained by variation in X.

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- R-Squared is the fraction of variation explained by our model.
- Variation in Y that can be explained by variation in X.
- $0 \leq R^2 \leq 1$ - R Squared is between 0 and 1 by construction.

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- R-Squared is the fraction of variation explained by our model.
- Variation in Y that can be explained by variation in X.
- $0 \leq R^2 \leq 1$ - R Squared is between 0 and 1 by construction.
- Going back to our wage and education problem, the R-Squared of 0.0132 reveals 1.32 percent of variation in wage can be explained by the variation in education.

Goodness of Fit - R Squared

- SST = Total Sum of Square = $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- SSE = Explained Sum of Square = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
- SSR = Residual Sum of Square = $\sum_{i=1}^{n}\hat{\mu}_i{}^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

  By Definition SST = SSE + SSR - Proof in Book*

- we define R-Squared $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- R-Squared is the fraction of variation explained by our model.
- Variation in Y that can be explained by variation in X.
- $0 \leq R^2 \leq 1$ - R Squared is between 0 and 1 by construction.
- Going back to our wage and education problem, the R-Squared of 0.0132 reveals 1.32 percent of variation in wage can be explained by the variation in education.
- Is a low R-Squared bad? Should we aim to increase R-squared?

Units of Measurement & Functional Form

Changing Units in OLS Regression

How does changing unit of either independent variable (X) or dependent variable (Y) affect our estimates $\hat{\beta}_0$ & $\hat{\beta}_1$.

- if dependent variable (y) is multiplied by the constant c, then the OLS intercept $\hat{\beta}_0$ and slope intercept $\hat{\beta}_1$ is also multiplied by the constant c.

Changing Units in OLS Regression

How does changing unit of either independent variable (X) or dependent variable (Y) affect our estimates $\hat{\beta}_0$ & $\hat{\beta}_1$.

- if dependent variable (y) is multiplied by the constant c, then the OLS intercept $\hat{\beta}_0$ and slope intercept $\hat{\beta}_1$ is also multiplied by the constant c.
- if independent variable (x) is multiplied by the constant c, then the OLS intercept $\hat{\beta}_0$ will not change. whereas the slope intercept $\hat{\beta}_1$ is divided by the constant c.

Change in Units of Dependent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Change in Units of Dependent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply both sides by c, so that $y_i^{new} = c \times y_i$,

$$cy_i = c\hat{\beta}_0 + c\hat{\beta}_1 x_i + c\hat{\mu}_i$$

Change in Units of Dependent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply both sides by c, so that $y_i^{new} = c \times y_i$,

$$cy_i = c\hat{\beta}_0 + c\hat{\beta}_1 x_i + c\hat{\mu}_i$$

Therefore, $\hat{\beta}_0^{new} = c\hat{\beta}_0$ and $\hat{\beta}_1^{new} = c\hat{\beta}_1$, so

$$y_i^{new} = \hat{\beta}_0^{new} + \hat{\beta}_1^{new} x_i + \hat{\mu}_i^{new}$$

Change in Units of Dependent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply both sides by c, so that $y_i^{new} = c \times y_i$,

$$cy_i = c\hat{\beta}_0 + c\hat{\beta}_1 x_i + c\hat{\mu}_i$$

Therefore, $\hat{\beta}_0^{new} = c\hat{\beta}_0$ and $\hat{\beta}_1^{new} = c\hat{\beta}_1$, so

$$y_i^{new} = \hat{\beta}_0^{new} + \hat{\beta}_1^{new} x_i + \hat{\mu}_i^{new}$$

Does $E(\hat{\mu}_i^{new}) = 0$ and $E(\hat{\mu}_i^{new}|x_i) = 0$ still holds true? *

Change in Units of Independent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Change in Units of Independent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply x by c, so that $x_i^{new} = c \times x_i$. Multiplying and dividing $\hat{\beta}_1 x_i$ by c,

$$y_i = \hat{\beta}_0 + \frac{c}{c}\hat{\beta}_1 x_i + \mu_i$$

Change in Units of Independent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply x by c, so that $x_i^{new} = c \times x_i$. Multiplying and dividing $\hat{\beta}_1 x_i$ by c,

$$y_i = \hat{\beta}_0 + \frac{c}{c}\hat{\beta}_1 x_i + \mu_i$$

$$y_i = \hat{\beta}_0 + \frac{\hat{\beta}_1}{c} c x_i + \mu_i$$

Change in Units of Independent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply x by c, so that $x_i^{new} = c \times x_i$. Multiplying and dividing $\hat{\beta}_1 x_i$ by c,

$$y_i = \hat{\beta}_0 + \frac{c}{c}\hat{\beta}_1 x_i + \mu_i$$

$$y_i = \hat{\beta}_0 + \frac{\hat{\beta}_1}{c} c x_i + \mu_i$$

If $\hat{\beta}_1^{new} = \frac{\hat{\beta}_1}{c}$, then

$$y_i = \hat{\beta}_0 + \hat{\beta}_1^{new} x_i^{new} + \mu_i$$

Change in Units of Independent Variable

Suppose we have the following model;

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \mu_i$$

Let's multiply x by c, so that $x_i^{new} = c \times x_i$. Multiplying and dividing $\hat{\beta}_1 x_i$ by c,

$$y_i = \hat{\beta}_0 + \frac{c}{c}\hat{\beta}_1 x_i + \mu_i$$

$$y_i = \hat{\beta}_0 + \frac{\hat{\beta}_1}{c} c x_i + \mu_i$$

If $\hat{\beta}_1^{new} = \frac{\hat{\beta}_1}{c}$, then

$$y_i = \hat{\beta}_0 + \hat{\beta}_1^{new} x_i^{new} + \mu_i$$

Does $E(\hat{\mu}_i^{new} | x_i^{new}) = 0$ still holds true? *

Examples

Example 1:
Suppose we have the following model where payment to IPPs is measured in USD.

$$Payment_i = 10000 + 20 powergen_i$$

we want to estimate the model, where payments are measured in PKR (1 USD = 280 PKR). What will be the new $\hat{\beta}_0$ and $\hat{\beta}_1$?

Examples

Example 1:
Suppose we have the following model where payment to IPPs is measured in USD.

$$Payment_i = 10000 + 20 powergen_i$$

we want to estimate the model, where payments are measured in PKR (1 USD = 280 PKR). What will be the new $\hat{\beta}_0$ and $\hat{\beta}_1$?

Example 2:
Suppose we have the following model for child mortality rate where weight is measured in pounds (lbs).

$$Mortality_i = 0.2 + 0.1 birthweight_i$$

What is the effect of birth weight (in KGs) on the chances of mortality. (1 KG = 2.2 lbs)

Incorporating Non-Linearity

It is important to note the linear relationship in our simple regression is only with respect to βs. We can use any non-linear form of y and x.

There are three useful non-linear relationship in y and x that we use.

1. level-log - $y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$

Incorporating Non-Linearity

It is important to note the linear relationship in our simple regression is only with respect to βs. We can use any non-linear form of y and x.

There are three useful non-linear relationship in y and x that we use.

1. level-log - $y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$
2. log-level - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$

Incorporating Non-Linearity

It is important to note the linear relationship in our simple regression is only with respect to βs. We can use any non-linear form of y and x.

There are three useful non-linear relationship in y and x that we use.

1. level-log - $y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$
2. log-level - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$
3. log-log - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$

Incorporating Non-Linearity

It is important to note the linear relationship in our simple regression is only with respect to βs. We can use any non-linear form of y and x.

There are three useful non-linear relationship in y and x that we use.

1. level-log - $y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$
2. log-level - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$
3. log-log - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$

Incorporating Non-Linearity

It is important to note the linear relationship in our simple regression is only with respect to βs. We can use any non-linear form of y and x.

There are three useful non-linear relationship in y and x that we use.

1. level-log - $y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$
2. log-level - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$
3. log-log - $log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$

Everything stays the same, only the interpretation of $\hat{\beta}_1$ changes.

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

To find the interpretation of $\hat{\beta}_1$, let's take the total differential.

$$dy = d\hat{\beta}_0 + d\hat{\beta}_1 log(x) + d\hat{\mu}$$

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

To find the interpretation of $\hat{\beta}_1$, let's take the total differential.

$$dy = d\hat{\beta}_0 + d\hat{\beta}_1 log(x) + d\hat{\mu}$$

we know $d\hat{\mu} = 0$ and $d\hat{\beta}_0 = 0$ so

$$dy = \hat{\beta}_1 \frac{dx}{x}$$

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

To find the interpretation of $\hat{\beta}_1$, let's take the total differential.

$$dy = d\hat{\beta}_0 + d\hat{\beta}_1 log(x) + d\hat{\mu}$$

we know $d\hat{\mu} = 0$ and $d\hat{\beta}_0 = 0$ so

$$dy = \hat{\beta}_1 \frac{dx}{x}$$

Multiplying and dividing the R.H.S. by 100,

$$dy = \frac{\hat{\beta}_1}{100} \frac{dx}{x} 100$$

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

To find the interpretation of $\hat{\beta}_1$, let's take the total differential.

$$dy = d\hat{\beta}_0 + d\hat{\beta}_1 log(x) + d\hat{\mu}$$

we know $d\hat{\mu} = 0$ and $d\hat{\beta}_0 = 0$ so

$$dy = \hat{\beta}_1 \frac{dx}{x}$$

Multiplying and dividing the R.H.S. by 100,

$$dy = \frac{\hat{\beta}_1}{100} \frac{dx}{x} 100$$

Since $\frac{dx}{x}100$ is the percentage change in x, $dy = \frac{\hat{\beta}_1}{100}$ (%Δ in x).

Level-Log Form

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

To find the interpretation of $\hat{\beta}_1$, let's take the total differential.

$$dy = d\hat{\beta}_0 + d\hat{\beta}_1 log(x) + d\hat{\mu}$$

we know $d\hat{\mu} = 0$ and $d\hat{\beta}_0 = 0$ so

$$dy = \hat{\beta}_1 \frac{dx}{x}$$

Multiplying and dividing the R.H.S. by 100,

$$dy = \frac{\hat{\beta}_1}{100} \frac{dx}{x} 100$$

Since $\frac{dx}{x}100$ is the percentage change in x, $dy = \frac{\hat{\beta}_1}{100}$ (%Δ in x).

Therefore, one percent change in x will lead to $\frac{\hat{\beta}_1}{100}$ unit change in y. (Def of Semi-Elasticity)

Log-level Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$$

Log-level Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$$

Following similar steps, we can show that (%Δ in y) = 100 $\hat{\beta}_1 dx$ *

Log-level Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\mu}_i$$

Following similar steps, we can show that (%Δ in y) = 100 $\hat{\beta}_1 dx$ *
Therefore, one unit change in x will lead to 100$\hat{\beta}_1$ percent change in y.
(Def of Semi-Elasticity)

Log-Log Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

Log-Log Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

Following similar steps, we can show that (%Δ in y) = $\hat{\beta}_1$(%Δ in x) *

Log-Log Form

$$log(y_i) = \hat{\beta}_0 + \hat{\beta}_1 log(x_i) + \hat{\mu}_i$$

Following similar steps, we can show that (%Δ in y) = $\hat{\beta}_1$(%Δ in x) * Therefore, one percent change in x will lead to $\hat{\beta}_1$ percent change in y. (Def of Elasticity)

Summary

TABLE 2.3 Summary of Functional Forms Involving Logarithms

Model	Dependent Variable	Independent Variable	Interpretation of β_1
Level-level	y	x	$\Delta y = \beta_1 \Delta x$
Level-log	y	$\log(x)$	$\Delta y = (\beta_1/100)\%\Delta x$
Log-level	$\log(y)$	x	$\%\Delta y = (100\beta_1)\Delta x$
Log-log	$\log(y)$	$\log(x)$	$\%\Delta y = \beta_1 \%\Delta x$

Examples

Example 1:

$$wage = 50000 + 900000 log(education) + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

Examples

Example 1:

$$wage = 50000 + 900000 log(education) + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

1 % increase in education will lead to 9000 ($\frac{900000}{100}$) PKR increase in wages.

Examples

Example 2:

$$log(wage) = 0.584 + 0.083education + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

Examples

Example 2:

$$log(wage) = 0.584 + 0.083education + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

1 unit increase in education will lead to 8.3 (100×0.083) percent increase in wages.

Examples

Example 3:

$$log(salary) = 4.822 + 1.2log(sales) + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

Examples

Example 3:

$$log(salary) = 4.822 + 1.2log(sales) + \mu$$

What is the interpretation of slope coefficient ($\hat{\beta}_1$)?

1 percent increase in sales will lead to 1.2 percent increase in salary.

Expected Values and Variance of the OLS Estimators

Unbiasedness of OLS Estimator

Unbiasedness explores the relationship between parameter and estimator. In our Population model, we have β_0 and β_1 that represent the relationship between y and x in population.

$$y = \beta_0 + \beta_1 x + \mu$$

Unbiasedness of OLS Estimator

Unbiasedness explores the relationship between parameter and estimator. In our Population model, we have β_0 and β_1 that represent the relationship between y and x in population.

$$y = \beta_0 + \beta_1 x + \mu$$

However, we usually do not have the population but a sample from that population. We want to estimate the β_s of the population using the sample data. So, we estimate the following model from our sample data;

$$y = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\mu}$$

Unbiasedness of OLS Estimator

Unbiasedness explores the relationship between parameter and estimator. In our Population model, we have β_0 and β_1 that represent the relationship between y and x in population.

$$y = \beta_0 + \beta_1 x + \mu$$

However, we usually do not have the population but a sample from that population. We want to estimate the β_s of the population using the sample data. So, we estimate the following model from our sample data;

$$y = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\mu}$$

we want the $\hat{\beta}_i$ to be as close to β_i. The $\hat{\beta}_i$ will be unbiased estimator of β_i if

$$E[\hat{\beta}_i] = \beta_i$$

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I
2. The sample is a random sample from the population - SLR II

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I
2. The sample is a random sample from the population - SLR II
3. There is a sample variation in independent variable. (All x_i are not same) - SLR III

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I
2. The sample is a random sample from the population - SLR II
3. There is a sample variation in independent variable. (All x_i are not same) - SLR III
4. Zero Conditional Mean $E(\mu|x) = 0$ - SLR IV

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I
2. The sample is a random sample from the population - SLR II
3. There is a sample variation in independent variable. (All x_i are not same) - SLR III
4. Zero Conditional Mean $E(\mu|x) = 0$ - SLR IV

Unbiasedness of OLS Estimator

The good thing is you do not need to prove it at this stage. The brilliant econometricians have proved it for you under some assumptions. If

1. The population function is linear in parameters. $y = \beta_0 + \beta_1 x + \mu$ - SLR I
2. The sample is a random sample from the population - SLR II
3. There is a sample variation in independent variable. (All x_i are not same) - SLR III
4. Zero Conditional Mean $E(\mu|x) = 0$ - SLR IV

Under Assumption I through IV, our $\hat{\beta}_i$ will be unbiased estimator of population β_i

Variance of the OLS Estimators

Why do we need the variance of the OLS estimators?

- To do the statistical inference (Hypothesis Testing) we need the variance of the OLS estimators. Recall you used to do that using t-test statistics;

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

(μ_0 is the hypothesized population mean, s the sample standard deviation)

Variance of the OLS Estimators

Why do we need the variance of the OLS estimators?

- To do the statistical inference (Hypothesis Testing) we need the variance of the OLS estimators. Recall you used to do that using t-test statistics;

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

(μ_0 is the hypothesized population mean, s the sample standard deviation)

- To do something similar with the estimators of simple linear regression, we need the variance of OLS estimators.

Variance of the OLS Estimators

Why do we need the variance of the OLS estimators?

- To do the statistical inference (Hypothesis Testing) we need the variance of the OLS estimators. Recall you used to do that using t-test statistics;

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

(μ_0 is the hypothesized population mean, s the sample standard deviation)

- To do something similar with the estimators of simple linear regression, we need the variance of OLS estimators.
- The main source of variation in a linear regression comes from $(\mu|x)$. Therefore, let's start with variance of $(\mu|x)$

Variance of OLS Estimators

Let's define the error variance $Var[\mu|x] = \sigma^2$.

Variance of OLS Estimators

Let's define the error variance $Var[\mu|x] = \sigma^2$.

$$Var[\mu|x] = E[\mu^2|x] - (E[\mu|x])^2 = E[\mu^2|x]$$

Variance of OLS Estimators

Let's define the error variance $Var[\mu|x] = \sigma^2$.

$$Var[\mu|x] = E[\mu^2|x] - (E[\mu|x])^2 = E[\mu^2|x]$$

Now if the error variance does not depend on x then we call it homoskedasticity - SLR V.

$$Var[\mu|x] = E[\mu^2|x] = E[\mu^2] = \sigma^2$$

Variance of OLS Estimators

Let's define the error variance $Var[\mu|x] = \sigma^2$.

$$Var[\mu|x] = E[\mu^2|x] - (E[\mu|x])^2 = E[\mu^2|x]$$

Now if the error variance does not depend on x then we call it homoskedasticity - SLR V.

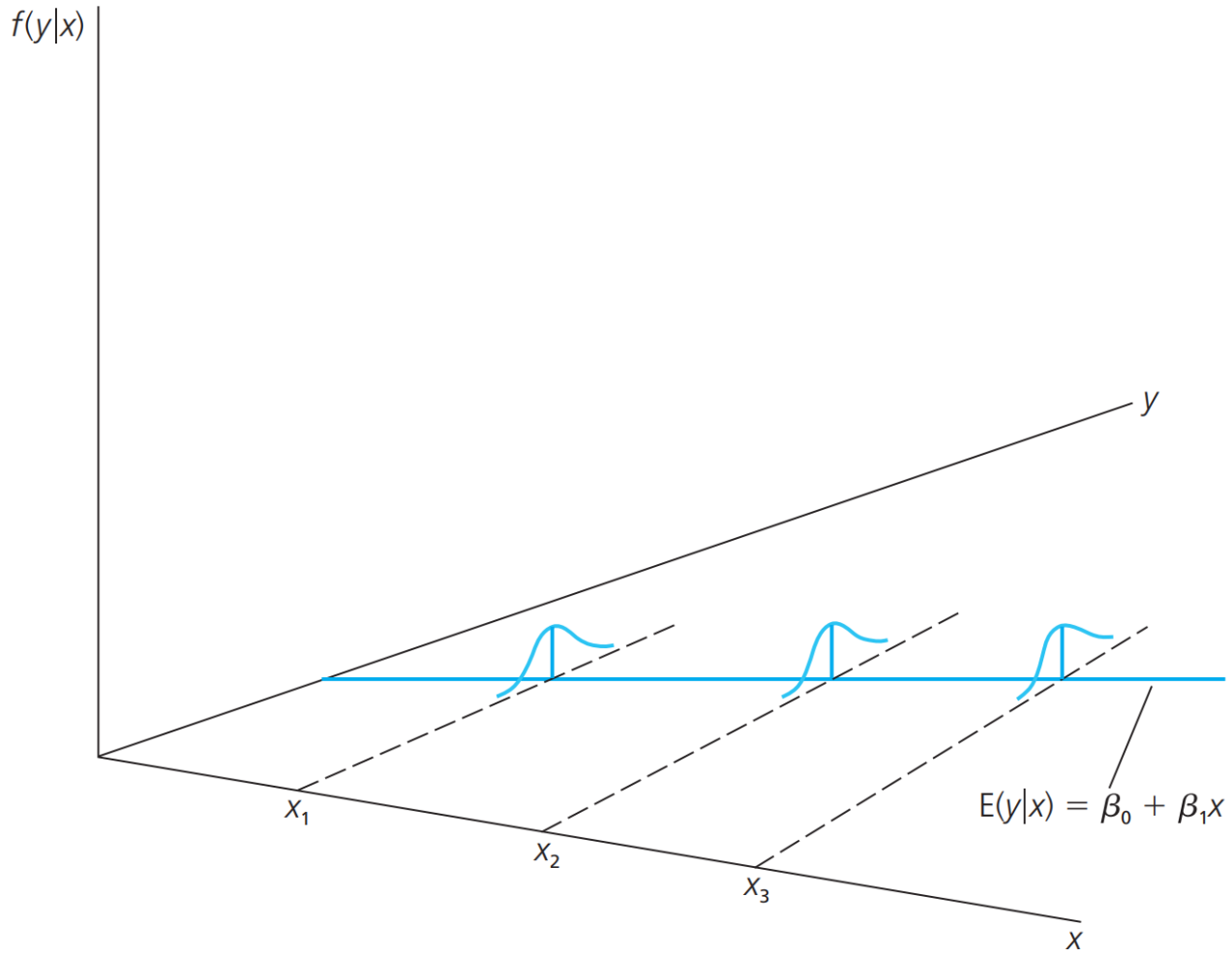
$$Var[\mu|x] = E[\mu^2|x] = E[\mu^2] = \sigma^2$$

Under the assumption of homoskedasticity (SLR V), our OLS estimators are also BLUE, Best Linear Unbiased Estimators, meaning they have the minimum variance among all linear unbiased estimators (Gauss-Markov Theorem).

- What does homoskedasticity mean?

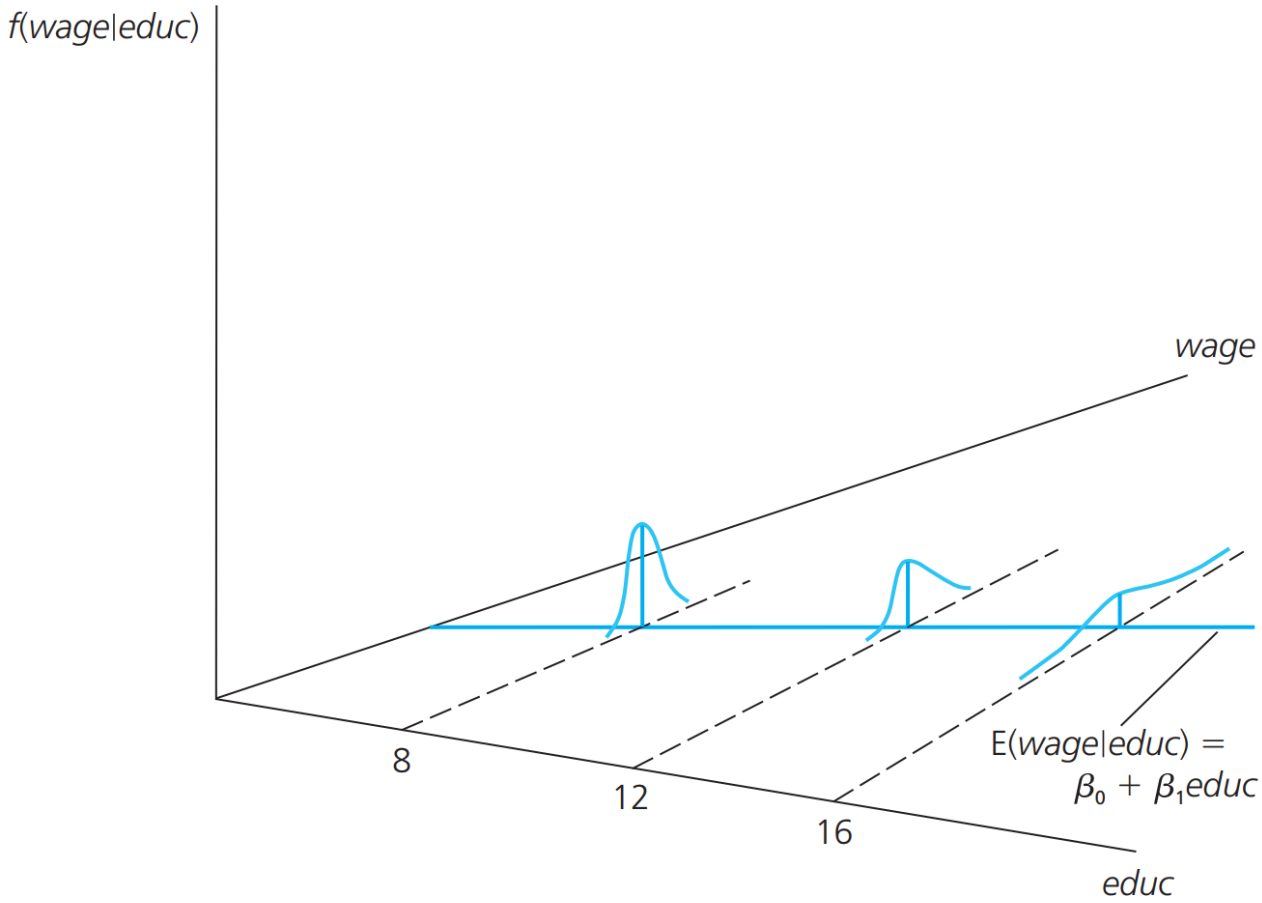
Homoskedasticity

FIGURE 2.8 The simple regression model under homoskedasticity.



Heteroskedasticity

FIGURE 2.9 Var(*wage*|*educ*) increasing with *educ*.



Variance of OLS Estimators

Now we have the variance of error, let's come back to the variance of OLS estimators.

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x}$$

Variance of OLS Estimators

Now we have the variance of error, let's come back to the variance of OLS estimators.

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x}$$

Similarly,

$$Var(\hat{\beta}_0) = \frac{\sigma^2 n^{-1} \sum_{i=1}^{n}(x_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{Var(\hat{\beta}_1)}{n} \sum_{i=1}^{n}(x_i)^2$$

Variance of OLS Estimators

Now we have the variance of error, let's come back to the variance of OLS estimators.

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x}$$

Similarly,

$$Var(\hat{\beta}_0) = \frac{\sigma^2 n^{-1} \sum_{i=1}^{n}(x_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{Var(\hat{\beta}_1)}{n} \sum_{i=1}^{n}(x_i)^2$$

The square root of these variances is the standard deviation of $\hat{\beta}_1$ and $\hat{\beta}_0$, which depends on the unknown σ. Once we replace σ with its estimate $\hat{\sigma}$ obtained from the residuals, we get the standard error of $\hat{\beta}_1$ and $\hat{\beta}_0$, which is what we actually use in practice.

Regression Through Origin

Regression Through Origin

Regression through origin means forcing $\hat{\beta}_0 = 0$. Therefore, we estimate the following model;

$$y = \tilde{\beta}x + \tilde{\mu}$$

Regression Through Origin

Regression through origin means forcing $\hat{\beta}_0 = 0$. Therefore, we estimate the following model;

$$y = \tilde{\beta}x + \tilde{\mu}$$

Following the same steps we can get the following

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

Regression Through Origin

Regression through origin means forcing $\hat{\beta}_0 = 0$. Therefore, we estimate the following model;

$$y = \tilde{\beta}x + \tilde{\mu}$$

Following the same steps we can get the following

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

However, if the true population model actually has an intercept $\beta_0 \neq 0$, then $\tilde{\mu} = \beta_0 + \mu$, so

$$E[\tilde{\mu}|x] = \beta_0 \neq 0$$

Regression Through Origin

Regression through origin means forcing $\hat{\beta}_0 = 0$. Therefore, we estimate the following model;

$$y = \tilde{\beta}x + \tilde{\mu}$$

Following the same steps we can get the following

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

However, if the true population model actually has an intercept $\beta_0 \neq 0$, then $\tilde{\mu} = \beta_0 + \mu$, so

$$E[\tilde{\mu}|x] = \beta_0 \neq 0$$

This violates SLR-IV, which requires the conditional mean of the error to be exactly zero, not just constant across x, making $\tilde{\beta}$ a Biased Estimator.

Example

wage	exper
3.5	7
3.75	7
5.9	9
3.24	14
2.3	2
2.6	4
6.5	8

Table: Regression Data

Example

wage	exper
3.5	7
3.75	7
5.9	9
3.24	14
2.3	2
2.6	4
6.5	8

Table: Regression Data

Question 1: Estimate the Following Model?

$$wage_i = \hat{\beta}_0 + \hat{\beta}_1 exper_i + \hat{\mu}_i$$

Example

wage	exper
3.5	7
3.75	7
5.9	9
3.24	14
2.3	2
2.6	4
6.5	8

Table: Regression Data

Question 1: Estimate the Following Model?

$$wage_i = \hat{\beta}_0 + \hat{\beta}_1 exper_i + \hat{\mu}_i$$

Question 2: Find the R-Squared of the above model?