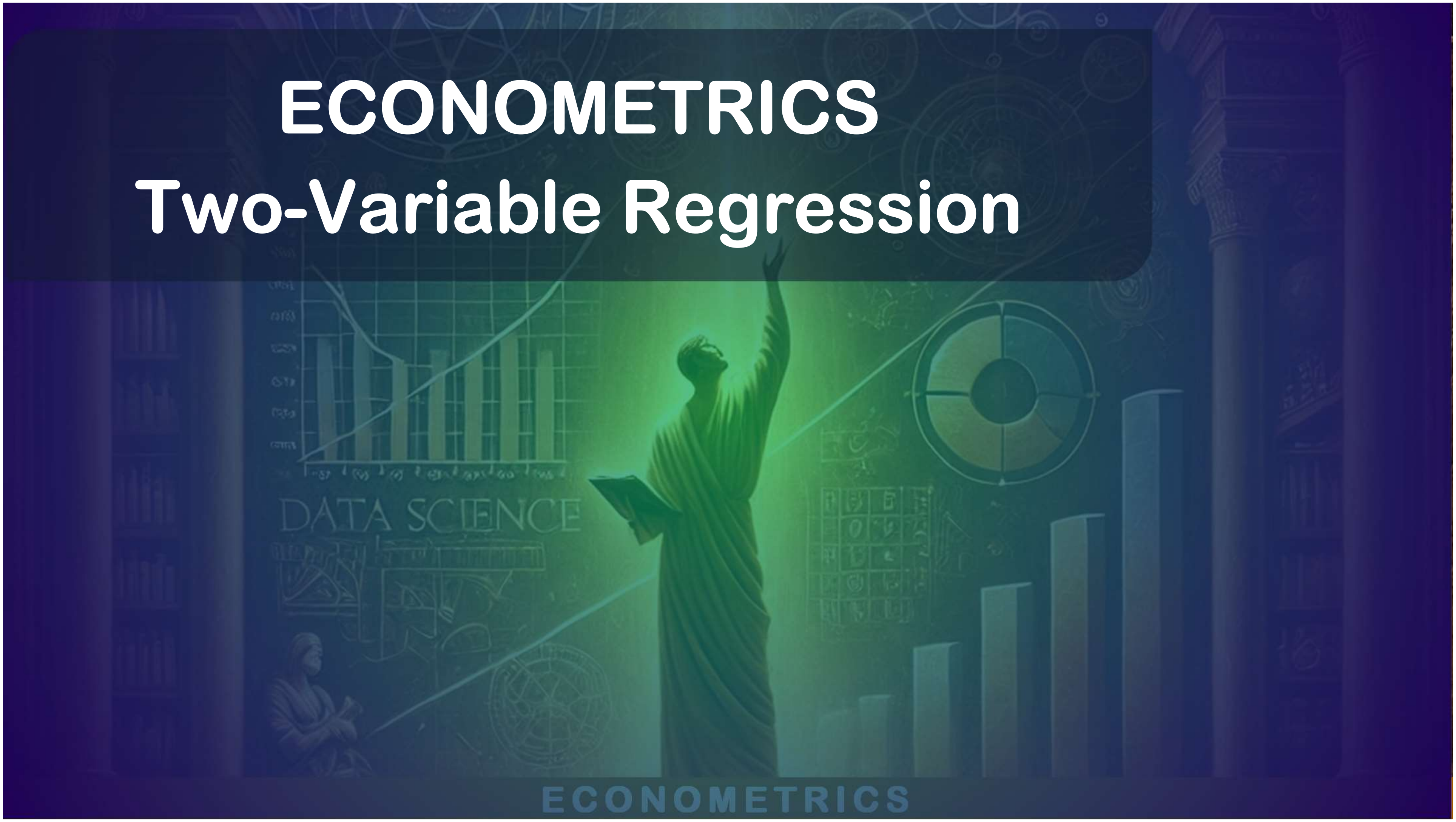
ECONOMETRICS
Two-Variable Regression



Lesson Goal

- Learn the concepts of the two-variable regression analysis.

Population Regression Function

- In an econometric model:

Dependent Variable ← **independent variable(s)**

estimate/predict the mean — known/fixed values

- Expectation: estimate Y with the known values of X_i.

  $E(Y|X_i)$ – "the expectation of Y given the values of X_i"

Stochastic Error Term

$E(Y|X)$ – "the expectation of Y given the values of X"

Functionally related to X.

- Notation:

$$E(Y|X_i) = f(X_i)$$

- This is called "Conditional Expectation Function (CEF)" or "Population Regression Function (PRF)".

Population Regression Function

- If PRF is a linear function of X_i, then:

$$E(Y|X_i) = \beta_1 + \beta_2 X_i$$

Regression coefficients
(intercept and slope coefficients)

PRF – Concept of Linearity

- Linearity in the variables

$E(Y|X_i)$ is a linear function of X_i

$E(Y|X_i) = \beta_1 + \beta_2 X_i^2$ - nonlinear in variables

$E(Y|X_i) = \beta_1 + \beta_2 \sqrt{X_i}$- nonlinear in variables

$E(Y|X_i) = \beta_1 + \beta_2 X_i$- linear in variables

PRF – Concept of Linearity

- Linearity in the parameters

$E(Y|X_i)$ is a linear function of the parameters β_1, β_2

$E(Y|X_i) = \beta_1 + \beta_2^2 X_i$ - nonlinear in parameters

$E(Y|X_i) = \beta_1^3 + \sqrt[3]{\beta_2} X_i$- nonlinear in parameters

$E(Y|X_i) = \beta_1 + \beta_2 X_i$- linear in parameters

PRF – Concept of Linearity

- Linearity in the parameters is the preferred choice for regression analysis.

- Linear regression means:

Linear in the parameters

- May or may not be linear in X_i's.

PRF – Stochastic Specification

- For a given population Y_i, we estimate $E(Y|X_i)$ which is also called the "mean" (estimated/predicted value) of Y_i.
- Deviation around this expected value (mean):

$$u_i = Y_i - E(Y|X_i)$$

or

$$Y_i = \underbrace{E(Y|X_i)}_{\text{systematic/deterministic component}} + \underbrace{u_i}_{\text{stochastic error or disturbance terms}}$$

non-systematic component

systematic/deterministic component

stochastic error or disturbance terms

Sample Regression Function (SRF)

- Samples are best approximations to the population.
- Let's develop the concept of SRF:

$$\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2} \cdot X_i$$

- where: $\widehat{Y_i}$ ("Y-hat" or "Y-cap") is an estimator of Y_i.

  $\widehat{\beta_1}$ is an estimator of β_1

  $\widehat{\beta_2}$ is an estimator of β_2

Sample Regression Function (SRF) – Stochastic Specification

- Stochastic form of SRF:

$$Y_i = \underbrace{\widehat{\beta_1} + \widehat{\beta_2} \cdot X_i}_{\widehat{Y_i}} + \widehat{u_i}$$

residual term

- Primary aim is to estimate PRF using SRF.

$$Y_i = \widehat{\beta_1} + \widehat{\beta_2} \cdot X_i + u_i$$

- Our analysis is based on a sample from a population.

THANK YOU!

Next Lesson: Method of Ordinary Least Squares (OLS)