# UNIT: 1

# Nature and Scope of Econometrics

## TOPICS

I) Meaning of econometrics
II) Relationship between Statistics, mathematics and econometrics
III) Economic and econometric models
IV) The aims and methodology of econometrics
V) Historical origin of the term regression and its modern interpretation
VI) Statistical vs deterministic relationship
VII) Regression vs causation
VIII) Regression vs correlation
IX) Terminology and notation
X) The nature and source of Data for econometric analysis

# What is Econometrics ?

UNIT:1

## Introduction

Econometrics is a branch of economics that applies mathematical methods to analyze economic data and test hypothesis.

It combines economic theory, mathematics, and statistical inference to quantify economic phenomena and provide empirical evidence to support economic theories or model.

Econometrics is a combine study of statistics, mathematics and economics theory it is called Econometrics.

Econometrics deal with the measurement of economic relationships.

The term 'econometrics' is formed from two words of Greek origin. οικονομία (economy) and μέτρον (measure)

Econometrics can be defined as that branch of economics concerned with the empirical estimation of economic relationships.

Econometrics is a combination of economic theory, mathematical economics and statistics.

# Relationship between Statistics, mathematics and Econometrics

## Introduction

The econometrics is a combination of economic theory mathematical economics and statistics.

## Econometrics and Mathematical Economics

* Mathematical Economics expresses the economic theory in term of Mathematical symbols.
* There is no essential differences between mathematical economics and economic theory. Both state the same relationship.
* Economic theory uses verbal exposition, mathematical economics uses mathematical symbols.
* Both express the economic relationships in an exact form.
* Econometrics provides numerical values for the coefficients of economic relationships.
* Econometrics methods provides exact numerical value of these parameters
* But mathematical economics do not provides numerical values of the coefficients.
* with the estimated coefficients, policy making and forecasting are made in Econometrics.

## Econometrics and Statistics

Econometrics differs from Statistics also.

Statistician collects data, records them, tabulate them and then describe the pattern in their occurrences over time.

Mathematical Statistics deals with methods of measurement also but these method are developed on the basis of controlled experiments in laboratories.

Such methods are not appropriate for economic relationships which cannot be measured on the basis of evidence provided by controlled experiments.

Econometrics uses statistical methods after adopting them to the problem of economic life.

## Why a separate Discipline?

Econometrics is considered a separate discipline because it combines economics theory, mathematics and statistical methods to analyze economic data.

1) Interdisciplinary Nature

Econometrics bridges economics, mathematics, and statistics creating a unique field that requires specialized knowledge from each area.

11) Empirical Analysis

It focus on testing economic theories and models using real-world data. Which is crucial for validating theoretical predictions and informing policy decisions.

III) Quantitative Methods

Econometrics employs advanced statistical techniques to estimate economic relationships, forecast economic trends, and evaluate the impact of economic policies.

## Economic and Econometric models

### Economic Models

Economic models are simplified representations of economic processes. They help economists understand, explain and predict economic behavior.

I) Theoretical Models

These models are based on economic theories and assumptions.

They often use mathematical equations to describe relationships between different economic variables eg: IS-LM

II) Graphical Models

These use graphs to illustrate economic concepts and relationships

For eg: Supply and demand curves show how prices and quantity are determined in a market

III) Computational Models

These involve complex algorithms and simulations to analyze economic scenarios. They are often used in macroeconomic forecasting and policy analysis.

## Econometric Models

Econometrics models, on the other hand are statistical models used to link economic theories and quantify economic relationships using real-world data.

i) Regression Analysis

This is the most common econometric technique. It estimates the relationship between a dependent variable and one or more independent variables.

ii) Time Series Models.

These models analyse data collected over time to identify trends, cycles and seasonal effects.

iii) Panel Data Models

These models use data that follows the same subjects over multiple time periods.

They help control for variables that do not change over time and can provide more accurate estimates.

## The aims towards methodology of econometrics

i) Statement of theory or hypothesis

ii) Specification of the mathematical model of the theory

iii) Specification of the statistical or econometric model

iv) Obtaining the Data.

v) Estimation of the parameters of the econometric model.

vi) Hypothesis testing

vii) Forecasting or prediction

viii) using the model for control or policy purposes.

1. Statement of theory or hypothesis

Keynes stated :

The fundamental psychological law ... is that men are disposed as a rule and on average to increase their consumption as their income increase, but not as much as the increase in their income.

In short, Keynes postulated that the marginal propensity to consume (MPC), the rate of change of consumption for a unit (say a rupee) change in income is greater than zero but less than 1.

2. Specification of the Mathematical Model of consumption.

$$Y = \beta_1 + \beta_2 X \quad 0 < \beta_2 < 1 \quad \text{---} (1)$$

where : Y = consumption expenditure

X = income

where $\beta_1$ and $\beta_2$, known as the parameters of the model are respectively the intercept and slope coefficients.

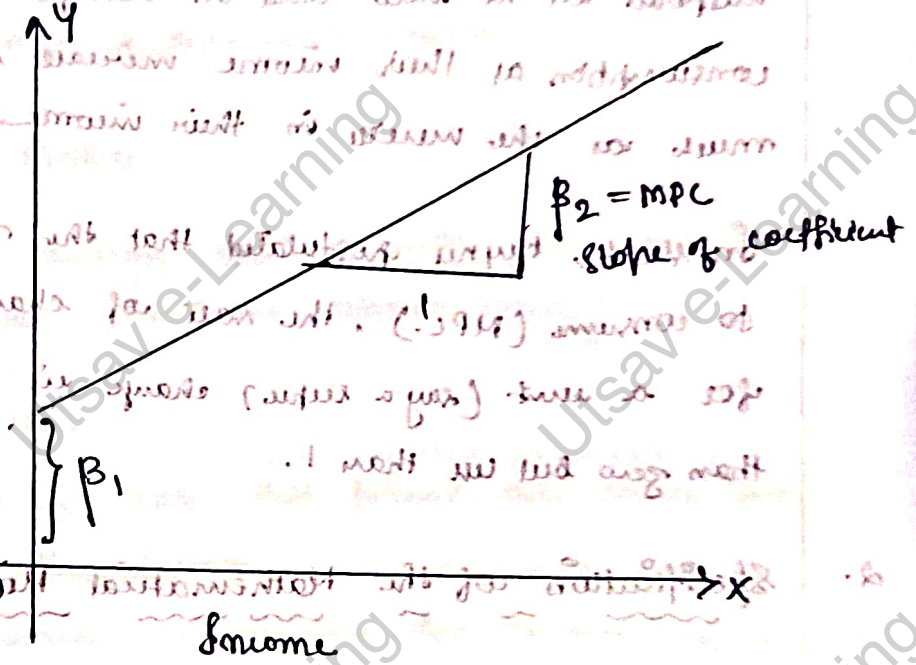
The slope coefficient $\beta_2$ measures the MPC

The relationship between consumption and income that is called the consumption function.

3) Specification of the Econometric Model of Consumption

If we assume that there is an exact or deterministic relationship between consumption and income.

But relationships between economic variables are generally inexact. Thus if we were to obtain data on consumption expenditure and disposable



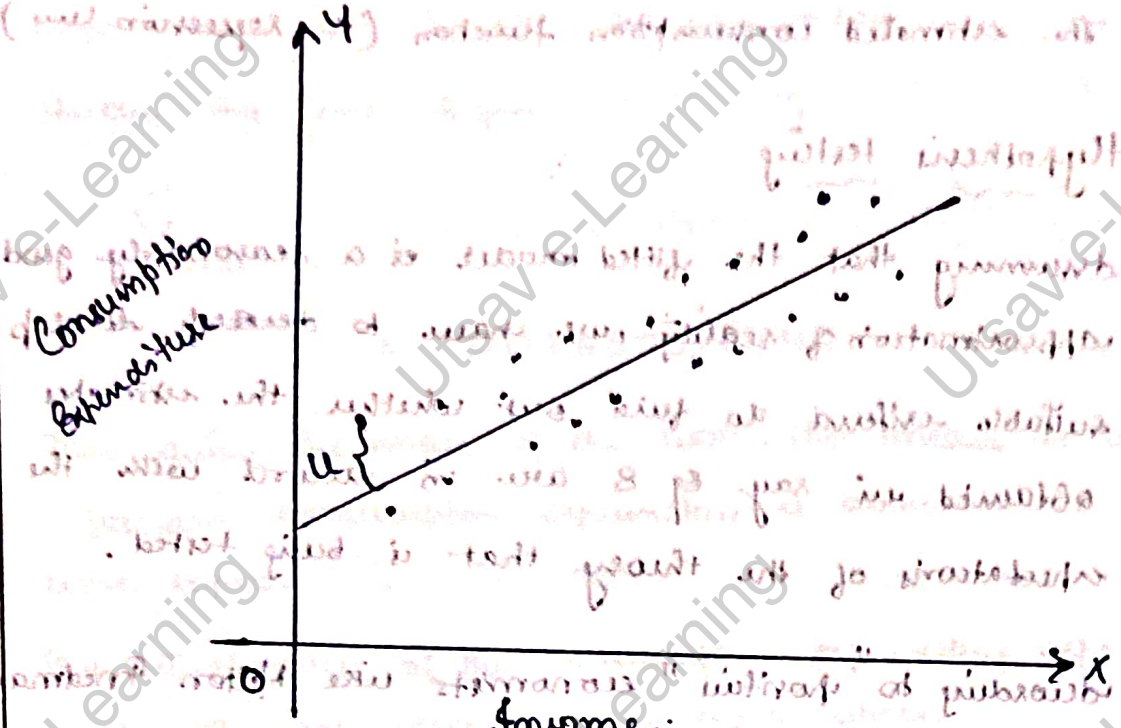

(1)

So allow for the inexact relationship between economic variables, the econometrician would modify the deterministic consumption function :

$$Y = \beta_1 + \beta_2 X + u \quad \text{---(2)}$$

where, $u$, known as the disturbance or error term is a random (stochastic) variable that has well-defined probabilistic properties.

4) Obtaining Data




5) Estimation of the Econometric Model.

Estimate the parameters of the consumption function.

The numerical estimates of the parameters give empirical content to the consumption function.

The statistical technique of regression analysis is the main tool used to obtain the estimates.

We Obtaining the following estimate $\beta_1$ and $\beta_2$ namely: 103736.0493 and 0.6303

estimated consumption function is

$$\hat{Y}_t = 103736.0493 + 0.6303 X_t \quad \text{---(3)}$$

the hat (^) on the Y indicates that it is an estimate.

The estimated consumption function (i.e., regression line)

6) Hypothesis testing

Assuming that the fitted model is a reasonably good approximation of reality, we have to develop suitable criteria to find out whether the estimates obtained in say Eq 8. are in accord with the expectations of the theory that is being tested.

According to "positive" economists like Milton Friedman, a theory or hypothesis that is not verifiable by appeal to empirical evidence may not be admissible as a part of scientific enquiry.

As noted earlier, Keynes expected the MPC to be positive but less than 1.

In our example we found the MPC to be about 0.63

But before we accept this finding as confirmation of Keynes's consumption theory

7) Forecasting or Prediction

If the chosen model does not refute the hypothesis or theory under consideration, we may use it to predict the future value(s) of the dependent or forecast variable Y on the basis of the known or expected future value(s) of the explanatory or predictor variable X.

To illustrate suppose we want to predict the mean consumption expenditure for 2007 - 2008. The GDP value for 2007 - 2008 was 31,29,717 crore rupees

Putting this GDP figure on the right-hand side of Eq (3)

$$\hat{Y}_{2007-2008} = 103736.0493 + 0.6303\,(3129717)$$

$$= 2076396.6744.$$

Thus, given the value of the GDP, the mean or average, forecast consumption expenditure is about 2076 thousand crore rupees.

The actual value of the consumption expenditure reported in 2007 - 08 was 1947 thousand crore rupees.

The estimated model Eq (3) thus overpredicted the actual consumption expenditure by about 130 thousand crore rupees. We could say the forecast error is about 130 thousand crore rupees.

B) Use of the Model for control or policy. Suppose

The government believes that consumer expenditure of about 2500 thousand crore rupees (1999 - 2000 prices) will help alleviate unemployment rate in the country.

$$2500850 = 103736.0493 + 0.6303\,(GDP)$$

$$(GDP = X)$$

$X = 3801,783$ approximately. that is an income level of about 3082 thousand crore rupees given an MPC of about 0.63 will produce an expenditure of about 2500 thousand crore rupees

Estimated model may be used for control or policy purposes. By appropriate fiscal and monetary policy mix the government can manipulate the control variable x to produce the desired level of the target variable Y.

## Types of Econometrics

- Theoretical
  - Classical
  - Bayesian
- Applied
  - Classical
  - Bayesian

Theoretical econometrics is concerned with the development of appropriate methods for measuring economic relationship specified by econometric models.

In applied econometrics we use the tools of theoretical econometrics to study some special field(s) of economics and business such as the production function, investment function, demand & supply function, portfolio theory etc.

# Historical origin of term Regression and its modern interpretation

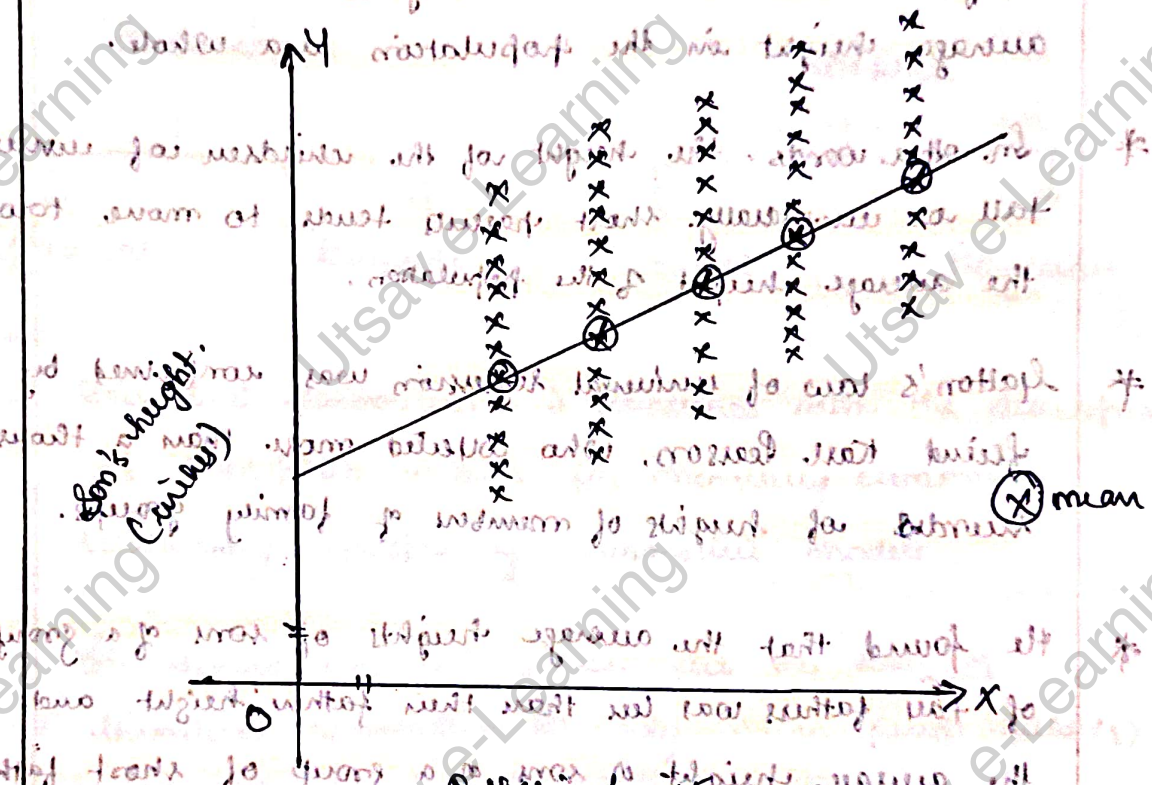
* The term regression was introduced by Francis Galton.
* In a famous paper, Galton found that, although there was a tendency for tall parents to have tall children and short parents to have short children, the average height of children born of parents of a given height tended to move or "regress" toward the average height in the population as a whole.
* In other words, the height of the children of unusually tall or unusually short parents tends to move toward the average height of the population.
* Galton's law of universal regression was confirmed by his friend Karl Pearson, who collected more than a thousand records of heights of members of family groups.
* He found that the average heights of sons of a group of tall fathers was less than their father's height and the average height of sons of a group of short fathers was greater than their fathers' height
* thus, "regressing" tall and short sons alike tended toward the average height of all the men.

As to Galton, this was "regression to mediocrity."

# The Modern Interpretation of Regression

Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables.

The explanatory variables with a view to estimating and/or predicting the (population) mean or average value of the former in terms of the known or fixed (in repeated sampling) values of the latter




Hypothetical distribution of son's heights corresponding to given heights of fathers.

# Statistical Versus Deterministic Relationship

## Statistical Variables

Here we deal with stochastic or random variables i.e. they have a probability distribution

## Deterministic Variables

Here we deal with relationships as in classical physics for example: Newton's law of gravity

## Regression and Causation

Regression does not necessarily imply causation. Any idea of causation must come from outside statistics, ultimately some theory or other

## Regression

We try to estimate or predict the average value of one variable on the basis of fixed values of other variables

Asymmetry in the relation between dependent and independent variable is treated as random / stochastic while independent variables are assumed to be fixed in repeated sampling.

## Correlation

Measure the strength of degree of linear association between two variables

Here both variables are treated symmetrically no distinction between dependent variable and explanatory variables

# Statistical Versus Deterministic Relationship

In statistical relationships among variables we essentially deal with random or stochastic variables that is variables that have probability distributions.

In functional or deterministic dependency, on the other hand, we also deal with variables but these variables are not random or stochastic.

In a deterministic relationship, the value of one variable can exactly determined by knowing the value of the other variable. In a statistical relationship there is natural variability in both measurements.

# Regression versus causation

Regression analysis deals with the dependence of one variable on other variables, it does not necessarily imply causation.

A statistical relationship however strong and however suggestive, can never establish causal connection: our ideas of causation must come from outside statistics, ultimately from some theory or other.

| | Regression | Correlation |
|---|---|---|
| Definition | Regression is a statistical method used to examine the relationship between a dependent variable and one or more independent variables. | Correlation means that a change in one variable directly causes a change in another variable. |
| Purpose | It helps predict the value of the dependent variable based on the value of the independent variable | It establishes a cause-and effect relationship |

## Regression versus Correlation

where the primary objective is to measure the strength or degree of linear association between two variables

Regression and correlation share some fundamental differences that are worth mentioning.

In regression analysis, there is an asymmetry in the way the dependent and explanatory variables are treated.

The dependent variable is assumed to be statistical random or stochastic that is to have a probability distribution.

| Regression | Correlation |
|---|---|
| It statistical method that models the relationship b/w a dependent variable & one or more independent variable (predictors). | It statistical method that quantifies the strength and direction of the relationship between two variables. |
| To predict or estimate the value of the dependent variable based on the independent variable. | To indicate whether and how strongly two variables are related. |
| $Y = a + bX$ | Correlation coefficient range $-1$ to $1$ |
| Where $Y$ = dependent variable | $r = +1$ = positive correlation |
| $X$ = Independent variable | $r = -1$ = negative correlation |
| | $r = 0$ = No correlation. |

## The Nature and Source of Data for Econometric Analysis

The success of any econometric analysis ultimately depends on the availability of the appropriate data.

It is therefore essential that we spend some time discussing the nature, source and limitations of the data that one may encounter in empirical analysis.

1) Time Series Data

A time series is a set of observations on the values that a variable takes at different times.

Such data may be collected at regular time intervals such as daily, weekly, monthly, quarterly, annually, quinquennially.

11) Cross Section Data

Cross section data are one or more variable collected at the same point in time, such as the census of population conducted by the Government of India every 10 years.

## The Accuracy of Data

* Most social science data are nonexperimental in nature. Therefore there is the possibility of observational errors, either of omission or commission.
* Even in experimentally collected data, errors of measurement arise from approximations and round offs.
* The sampling method used in obtaining the data may vary so widely that it is often difficult to compare the results obtained from various samples.
* Economic data are generally available at a highly aggregate level. For eg most macrodata (eg. GNP, employment, inflation, unemployment) are available for the economy as a whole or at the most for some broad geographical regions.

## The Measurement Scales of Variables

The variables that we will generally encounter fall into four broad categories:

→ Ratio Scale
→ Interval Scale
→ Ordinal Scale
→ Nominal Scale

i) Ratio Scale.

For a variable X, taking two values $X_1$ and $X_2$ the ratio $X_1/X_2$ and distance $(X_2 - X_1)$ are meaningful quantities.

ii) Interval Scale

An interval scale variable satisfies the last two properties of the ratio scale variable but not the first.

Thus, the distance between two time periods say (2000 - 1995) is meaningful, but not the ratio of two time periods

iii) Ordinal Scale

A variable belongs to this category only if it satisfies the third property of the ratio scale (i.e, natural ordering).

Eg. are grading system (A, B, C grades) or income class (upper, middle, lower)

For these variables the ordering exists but the distances between the categories cannot be quantified.

iv) Nominal Scale

Variables in this category share none of the features of the ratio scale variable.

Variables such as gender (male, female) and marital status (married, unmarried, divorced, separate) simply denote categories.

Econometric techniques that may be suitable for ratio, interval scale variables may not suitable for nominal scale variables.