

DegreeLive

# **FOUNDATIONS FOR BUSINESS ANALYTICS**

BBA Semester 2 – Calicut University

## MODULE 1: INTRODUCTION TO BUSINESS ANALYTICS & PROBABILITY

### Business Analytics

1. Business Analytics is the use of data, statistics and analytical techniques to support business decisions.
2. Helps organizations identify patterns, solve problems and improve performance.
3. Converts raw data into useful information.

### Why Analytics?

1. Better decision making.
2. Improved efficiency.
3. Risk reduction.
4. Competitive advantage.
5. Forecasting future trends.

### Types of Business Analytics

1. **Descriptive Analytics:** Explains what happened in the past.
2. **Diagnostic Analytics:** Explains why something happened.
3. **Predictive Analytics:** Predicts future outcomes.
4. **Prescriptive Analytics:** Suggests best actions to take.

### Random Experiment, Sample Space and Events

#### Random Experiment

1. An experiment whose outcome cannot be predicted with certainty.
2. Example: Tossing a coin, Rolling a dice.

#### Sample Space (S)

1. Set of all possible outcomes.
2. Example:
  - a. Coin Toss  $\rightarrow \{H, T\}$
  - b. Dice  $\rightarrow \{1,2,3,4,5,6\}$

#### Event

1. One or more outcomes from a sample space.
2. Example: Getting an even number on a dice.

#### Probability by Relative Frequency

*Probability = Number of times event occurs  $\div$  Total number of trials*

### Algebra of Events

1. **Union ( $A \cup B$ ):** Occurrence of A or B or both.

2. **Intersection ( $A \cap B$ ):** Occurrence of both A and B.
3. **Complement ( $A'$ ):** Event not occurring.

### Fundamental Concepts in Probability

#### Probability

1. Numerical measure of likelihood of an event.
2. Value lies between 0 and 1.

#### Axioms of Probability

1. Probability is never negative.
2. Probability of sample space = 1.
3. Probability of mutually exclusive events can be added.

#### Joint Probability

1. Probability of two events occurring together.
2. Example: Probability of drawing a red card and an ace.

### Marginal, Conditional Probability and Bayes' Theorem

#### Marginal Probability

1. Probability of a single event without considering other events.

#### Independent Events

1. Occurrence of one event does not affect another.
2. Example: Two coin tosses.

#### Conditional Probability

1. Probability of an event given that another event has occurred.
2. Example: Probability of rain given cloudy weather.

#### Applications of Probability

1. Risk analysis.
2. Insurance.
3. Business forecasting.
4. Investment decisions.

#### Bayes' Theorem

1. Calculates revised probability when new information becomes available.
2. Widely used in prediction and decision making.

## MODULE 2: THEORETICAL DISTRIBUTIONS

### Random Variables, PDF and CDF

#### Random Variable

1. Variable whose value depends on outcomes of a random experiment.
2. Example: Number obtained on a dice.

#### Types

1. **Discrete Random Variable:**
  - a. Takes specific values.
  - b. Example: Number of customers.
2. **Continuous Random Variable:**
  - a. Can take any value within a range.

b. Example: Height, weight, income.

### Probability Density Function (PDF)

1. Describes probability distribution of a continuous random variable.

### Cumulative Distribution Function (CDF)

1. Gives probability that a variable is less than or equal to a given value.

### Binomial Distribution

1. Probability distribution for situations having only two outcomes.
2. Examples: Success or Failure, Pass or Fail, Yes or No.

### Conditions

1. Fixed number of trials.
2. Independent trials.
3. Constant probability of success.

### Applications

1. Quality control.
2. Market surveys.
3. Customer response analysis.

### Poisson Distribution

1. Used to measure occurrence of rare events within a fixed interval.

2. Examples: Number of accidents per day, Number of customer complaints, Number of machine breakdowns.

### Characteristics

1. Events occur independently.
2. Average rate remains constant.

## Normal, Chi-Square, t and F Distributions

### Normal Distribution

1. Bell-shaped and symmetrical distribution.
2. Mean = Median = Mode.
3. Most common distribution in statistics.
4. Examples: Heights, IQ scores.

### Chi-Square Distribution

1. Used in hypothesis testing.
2. Useful for testing association between variables.

### Student's t-Distribution

1. Used when sample size is small and population variance is unknown.

### F-Distribution

1. Used to compare variances.
2. Important in Analysis of Variance (ANOVA).

## MODULE 3: SAMPLING, CORRELATION AND REGRESSION

### Sampling and Estimation

#### Sampling

1. Selecting a portion of a population for study.

#### Importance

1. Saves time.
2. Reduces cost.
3. Easier data collection.

#### Estimation

1. Using sample data to estimate population values.

### Population Parameters and Sample Statistics

#### Population

1. Entire group being studied.

#### Sample

1. Small part of population.

#### Population Parameters

1. Numerical measures of population.
2. Examples: Population Mean, Population Variance.

#### Sample Statistics

1. Numerical measures calculated from sample data.
2. Examples: Sample Mean, Sample Standard Deviation.

### Types of Sampling

#### Probability Sampling

1. Every member has a known chance of selection.
2. Types: Simple Random Sampling, Stratified Sampling, Systematic Sampling, Cluster Sampling.

#### Non-Probability Sampling

1. Selection based on convenience or judgment.
2. Types: Convenience Sampling, Judgment Sampling, Quota Sampling, Snowball Sampling.

### Sample Size Estimation and Population Estimation

#### Sample Size Estimation

1. Determining how many observations are required for reliable results.

#### Factors Affecting Sample Size

1. Population size.
2. Accuracy required.
3. Confidence level.
4. Variability in data.

#### Estimation of Population Parameters

1. Population Mean.
2. Population Proportion.
3. Population Variance.

## Central Limit Theorem (CLT)

1. One of the most important concepts in statistics.
2. States that the sampling distribution of sample means tends to become normal as sample size increases.
3. Works even when original population is not normally distributed.

### Importance

1. Basis for statistical inference.
2. Supports hypothesis testing and estimation.

## Correlation

### Meaning

1. Measures relationship between two variables.

### Significance

1. Helps identify strength and direction of relationship.
2. Useful in prediction and business decisions.

### Types

1. **Positive Correlation:** Both variables move in same direction. Example: Income and expenditure.
2. **Negative Correlation:** Variables move in opposite directions. Example: Price and demand.
3. **Zero Correlation:** No relationship exists.

## Methods of Simple Correlation

### Karl Pearson's Coefficient of Correlation

1. Measures degree of linear relationship.
2. Value ranges from -1 to +1.
3. Interpretation:
  - a. +1 → Perfect positive correlation.
  - b. -1 → Perfect negative correlation.
  - c. 0 → No correlation.

## Spearman's Rank Correlation

1. Based on ranks instead of actual values.
2. Used when data is ordinal.

## Regression

### Meaning

1. Statistical technique used to predict one variable from another.

### Significance

1. Forecasting.
2. Trend analysis.
3. Business planning.

### Regression vs Correlation

Correlation	Regression
Measures relationship	Predicts relationship
No dependent variable	Has dependent variable
Symmetrical	One-way relationship

## Linear Regression and Regression Lines

### Linear Regression

1. Relationship represented by a straight line.

### Regression Line of X on Y

1. Predicts X using Y.

### Regression Line of Y on X

1. Predicts Y using X.

### Standard Error of Estimate

1. Measures accuracy of prediction.
2. Lower value indicates better prediction.

## MODULE 4: TIME SERIES AND INDEX NUMBERS

## Time Series

### Meaning

1. Data collected over different periods of time.
2. Examples: Monthly sales, Annual profit, Daily stock prices.

### Significance

1. Forecasting future trends.
2. Planning and decision making.
3. Performance analysis.

## Components of Time Series

1. **Trend:** Long-term movement.
2. **Seasonal Variation:** Repeating pattern within a year.
3. **Cyclical Variation:** Long-term business cycle movements.
4. **Irregular Variation:** Unexpected fluctuations.

## Measurement of Trend

### Method of Least Squares

1. Mathematical method used to determine trend line.
2. Provides best-fit line through data points.

### Advantages

1. Objective.
2. Accurate.
3. Useful for forecasting.

## Parabolic Trend and Logarithmic Trend

### Parabolic Trend

1. Trend follows a curved pattern.
2. Used when growth accelerates or decelerates over time.

### Logarithmic Trend

1. Growth increases rapidly at first and then slows down.

2. Common in business growth and technology adoption.

## Index Numbers

---

### Meaning

1. Statistical measures showing changes in a variable over time.

### Significance

1. Measure price changes.
2. Measure economic performance.
3. Compare different periods.

### Examples

1. Consumer Price Index (CPI).
2. Wholesale Price Index (WPI).

## Construction of Index Numbers

---

### Problems in Construction

1. Selection of base year.
2. Selection of items.
3. Selection of weights.
4. Availability of data.

## Methods of Construction

1. **Unweighted Index Numbers:** All items treated equally.
2. **Weighted Index Numbers:** Different weights assigned based on importance.

## Test of Adequacy

A good index number should satisfy:

1. Time Reversal Test.
2. Factor Reversal Test.

## Chain Index Numbers

---

1. Index numbers where each year is compared with the immediately preceding year.
2. Uses continuously changing base year.

## Advantages

1. Reflects recent changes.
2. More flexible than fixed base index.

## Limitation

1. More complex calculations.
2. Errors may accumulate over time.