# BCA SEMESTER 2 — STATISTICS (Complete Notes)

I can't generate a downloadable PDF directly, but below is the **complete, topic-wise, print-ready content** covering the entire syllabus. You can **copy all of this into Google Docs / MS Word  Export as PDF**.

---

# UNIT 1: INTRODUCTION TO STATISTICS

---

## 1.1 Definition of Statistics

**Statistics** (Plural sense): Numerical facts collected in a systematic manner. **Statistics** (Singular sense): The science of collecting, organizing, analyzing, interpreting, and presenting data.

> **Croxton and Cowden**: "Statistics is the science which deals with collection, analysis and interpretation of numerical data."

---

## 1.2 Types of Data

### A. Primary Data

Data collected **first-hand** by the investigator for a specific purpose.

**Methods of collection:**

- Direct personal interview
- Indirect oral investigation
- Questionnaire method
- Schedule method
- Observation method

### B. Secondary Data

Data collected by **someone else** and used by the investigator.

**Sources:**

- Government publications
- Journals, newspapers
- Websites, databases

### Difference Table

| Feature | Primary Data | Secondary Data |
|---|---|---|
| Originality | Original | Already collected |
| Cost | Expensive | Economical |
| Time | Time-consuming | Quick |
| Reliability | More reliable | Less reliable |
| Suitability | Specific purpose | May not fit exactly |

---

## 1.3 Classification of Data

**Classification** = Arranging data into groups based on similarities.

### Types:

1. **Geographical** – By location (state, city)
2. **Chronological/Temporal** – By time (year, month)
3. **Qualitative** – By attributes (gender, religion)
4. **Quantitative** – By numerical values (income, marks)

---

## 1.4 Tabulation

**Tabulation** = Systematic arrangement of data in rows and columns.

### Parts of a Table:

- Table number
- Title
- Column headings (caption)
- Row headings (stub)
- Body of the table
- Footnotes
- Source

---

## 1.5 Frequency Distribution

**Frequency** = Number of times a value occurs.

### Types:

1. **Ungrouped (Discrete) Frequency Distribution**
2. **Grouped (Continuous) Frequency Distribution**

### Key Terms:

- **Class Interval**: Range of values in a class (e.g., 10–20)
- **Class Width**: Upper limit – Lower limit = 20 – 10 = 10
- **Class Mark (Mid-value)**: (Lower limit + Upper limit) / 2
- **Cumulative Frequency**: Running total of frequencies

### Example:

| Marks | Frequency (f) | Cumulative Frequency (CF) |
|---|---|---|
| 0–10 | 5 | 5 |
| 10–20 | 8 | 13 |
| 20–30 | 12 | 25 |
| 30–40 | 10 | 35 |
| 40–50 | 5 | 40 |

---

## 1.6 Diagrammatic & Graphical Representation

### Diagrams:

- **Bar Diagram** (Simple, Multiple, Subdivided)
- **Pie Chart** (Circle divided into sectors)

### Graphs:

- **Histogram** – For continuous data (bars touching)
- **Frequency Polygon** – Join mid-points of histogram tops
- **Ogive (Cumulative Frequency Curve)** – Less than / More than type
- **Frequency Curve** – Smooth version of frequency polygon

### Pie Chart Formula:

```
Angle of sector = (Component Value / Total Value) × 360°
```

**Example**: If expenditure on food = 400 out of total 2000

```
Angle = (400/2000) × 360° = 72°
```

---

---

# UNIT 2: MEASURES OF CENTRAL TENDENCY

## 2.1 Introduction

A **Measure of Central Tendency** is a single value that represents the center of the data.

### Properties of a Good Average:

1. Rigidly defined
2. Easy to understand and compute
3. Based on all observations
4. Capable of further algebraic treatment
5. Not affected much by extreme values

## 2.2 Arithmetic Mean (AM)

### A. For Raw Data (Ungrouped):

$\bar{x} = \Sigma x \ / \ n$

**Example**: Find mean of 5, 10, 15, 20, 25

$\bar{x} = (5 + 10 + 15 + 20 + 25) / 5 = 75/5 = 15$

### B. For Discrete Frequency Distribution:

$\bar{x} = \Sigma f \ x \ / \ \Sigma f$

**Example**:

| x | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| f | 3 | 5 | 7 | 4 | 1 |

$\Sigma fx = 30 + 100 + 210 + 160 + 50 = 550$
$\Sigma f = 20$
$\bar{x} = 550/20 = 27.5$

### C. For Grouped (Continuous) Data:

**Direct Method:**

$\bar{x} = \Sigma f \ m \ / \ \Sigma f$ (where m = mid-value of class)

**Shortcut (Assumed Mean) Method:**

```
x̄ = A + (Σf  d   / Σf  )    where d   = m   - A
```

**Step Deviation Method:**

```
x̄ = A + (Σf  u   / Σf  ) × h    where u   = (m   - A)/h
```

**Solved Example (Step Deviation)**:

| Class | f | Mid (m) | u = (m-25)/10 | fu |
|---|---|---|---|---|
| 0–10 | 5 | 5 | -2 | -10 |
| 10–20 | 8 | 15 | -1 | -8 |
| 20–30 | 12 | 25 | 0 | 0 |
| 30–40 | 10 | 35 | 1 | 10 |
| 40–50 | 5 | 45 | 2 | 10 |

A = 25, h = 10, Σf = 40, Σfu = 2

```
x̄ = 25 + (2/40) × 10 = 25 + 0.5 = 25.5
```

### Combined Mean:

$$\bar{x}_{12} = (n_1\bar{x}_1 + n_2\bar{x}_2) / (n_1 + n_2)$$

### Weighted Mean:

```
x̄w = Σw  x   / Σw
```

### Properties of Arithmetic Mean:

1. Σ(x   - x̄) = 0 (sum of deviations from mean is zero)
2. $\Sigma(x - \bar{x})^2$ is minimum
3. If each value is multiplied by k, new mean = k × x̄
4. If constant c is added to each value, new mean = x̄ + c

---

## 2.3 Median

**Median** = The middle value when data is arranged in ascending/descending order.

### A. For Raw Data:

Arrange in order. If n is odd:

```
Median = ((n+1)/2)th value
```

If n is even:

```
Median = Average of (n/2)th and (n/2 + 1)th values
```

**Example (Odd)**: 3, 5, 7, 9, 11  n = 5

```
Median = (5+1)/2 = 3rd value = 7
```

**Example (Even)**: 2, 4, 6, 8  n = 4

```
Median = (4/2)th + (4/2+1)th / 2 = (2nd + 3rd)/2 = (4+6)/2 = 5
```

### B. For Discrete Frequency Distribution:

Find cumulative frequency. Median = value corresponding to CF ≥ N/2.

### C. For Grouped Data:

```
Median = L + [(N/2 - CF) / f] × h
```

Where:

- L = Lower limit of median class
- N = Total frequency (Σf)
- CF = Cumulative frequency of class preceding median class
- f = Frequency of median class
- h = Class width

**Solved Example**:

| Class | f | CF |
|---|---|---|
| 0–10 | 5 | 5 |
| 10–20 | 8 | 13 |
| 20–30 | 12 | 25 |
| 30–40 | 10 | 35 |
| 40–50 | 5 | 40 |

N = 40, N/2 = 20

Median class = 20–30 (CF just ≥ 20)

```
L = 20, CF = 13, f = 12, h = 10
Median = 20 + [(20-13)/12] × 10
       = 20 + (7/12) × 10
       = 20 + 5.83
       = 25.83
```

---

## 2.4 Mode

**Mode** = The value that occurs most frequently.

### A. For Raw Data:

Simply find the most repeated value.

**Example**: 2, 3, 5, 5, 5, 7, 8  Mode = 5

### B. For Grouped Data:

$$Mode = L + [(f_1 - f_0) / (2f_1 - f_0 - f_2)] \times h$$

Where:

- L = Lower limit of modal class (class with highest frequency)
- $f_1$ = Frequency of modal class
- $f_0$ = Frequency of class preceding modal class
- $f_2$ = Frequency of class succeeding modal class
- h = Class width

**Solved Example**:

| Class | f |
|---|---|
| 0–10 | 5 |
| 10–20 | 8 |
| 20–30 | 12 |
| 30–40 | 10 |
| 40–50 | 5 |

$L = 20, f_1 = 12, f_0 = 8, f_2 = 10, h = 10$
Mode = 20 + [(12-8)/(2×12-8-10)] × 10
= 20 + [4/6] × 10
= 20 + 6.67
= 26.67

---

## 2.5 Empirical Relationship

Mode = 3 Median - 2 Mean

or equivalently:

Mean - Mode = 3(Mean - Median)

---

## 2.6 Geometric Mean (GM)

### For Raw Data:

$GM = (x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n)^\wedge(1/n)$

or equivalently:

log(GM) = (1/n) Σlog(x )

### For Frequency Distribution:

log(GM) = Σf log(x ) / Σf

**Example**: Find GM of 2, 4, 8

GM = (2 × 4 × 8)^(1/3) = (64)^(1/3) = 4

**Use**: Best for rates of growth, ratios, percentages.

---

## 2.7 Harmonic Mean (HM)

### For Raw Data:

HM = n / Σ(1/x )

### For Frequency Distribution:

HM = Σf / Σ(f /x )

**Example**: Find HM of 2, 4, 8

HM = 3 / (1/2 + 1/4 + 1/8) = 3 / (0.5 + 0.25 + 0.125) = 3/0.875 = 3.43

**Use**: Best for speed, rates (time-based problems).

---

## 2.8 Relationship Between AM, GM, HM

$AM \geq GM \geq HM$ (for positive values)

Also:

$$GM^2 = AM \times HM$$

---

## 2.9 Quartiles, Deciles, and Percentiles

### Quartiles (divide data into 4 parts):

$$Q_1 = L + [(N/4 - CF)/f] \times h$$
$$Q_2 = \text{Median}$$
$$Q_3 = L + [(3N/4 - CF)/f] \times h$$

### Deciles (divide data into 10 parts):

$$D_k = L + [(kN/10 - CF)/f] \times h \quad (k = 1,2,\ldots,9)$$

### Percentiles (divide data into 100 parts):

$$P_k = L + [(kN/100 - CF)/f] \times h \quad (k = 1,2,\ldots,99)$$

---

---

# UNIT 3: MEASURES OF DISPERSION

---

## 3.1 Introduction

**Dispersion** measures the **spread or variability** of data around the central value.

### Types:

1. **Absolute Measures**: Range, QD, MD, SD
2. **Relative Measures**: Coefficient of Range, Coefficient of QD, Coefficient of MD, Coefficient of Variation (CV)

---

## 3.2 Range

Range = Maximum Value - Minimum Value
= L - S

Coefficient of Range = (L - S) / (L + S)

**Example**: Data: 5, 10, 15, 20, 25

Range = 25 - 5 = 20
Coefficient of Range = (25-5)/(25+5) = 20/30 = 0.667

---

## 3.3 Quartile Deviation (Semi-Interquartile Range)

$QD = (Q_3 - Q_1) / 2$

Coefficient of $QD = (Q_3 - Q_1) / (Q_3 + Q_1)$

**Example**: If $Q_1 = 20$, $Q_3 = 40$

QD = (40-20)/2 = 10
Coefficient of QD = (40-20)/(40+20) = 20/60 = 0.333

---

## 3.4 Mean Deviation (Average Deviation)

### About Mean:

$MD(\bar{x}) = \Sigma|x - \bar{x}| / n$ (Raw data)
$MD(\bar{x}) = \Sigma f |x - \bar{x}| / \Sigma f$ (Frequency distribution)

### About Median:

$MD(Median) = \Sigma|x - Median| / n$

Coefficient of MD = MD / Mean (or Median)

**Solved Example**: Find MD about mean for: 2, 4, 6, 8, 10

$\bar{x} = (2+4+6+8+10)/5 = 30/5 = 6$

|2-6| + |4-6| + |6-6| + |8-6| + |10-6|

= 4 + 2 + 0 + 2 + 4 = 12

MD = 12/5 = 2.4
Coefficient of MD = 2.4/6 = 0.4

---

## 3.5 Standard Deviation (σ) and Variance (σ²)

**Standard Deviation** is the most important and widely used measure of dispersion.

### A. For Raw Data:

**Direct Method**:

$\sigma = \sqrt{[\Sigma(x_i - \bar{x})^2 / n]}$

**Shortcut Method**:

$\sigma = \sqrt{[(\Sigma x_i^2/n) - (\Sigma x_i/n)^2]}$
$= \sqrt{[(\Sigma x^2/n) - \bar{x}^2]}$

### B. For Discrete Frequency Distribution:

$\sigma = \sqrt{[(\Sigma f_i x_i^2/\Sigma f_i) - (\Sigma f_i x_i/\Sigma f_i)^2]}$

### C. For Grouped Data (Step Deviation):

$\sigma = h \times \sqrt{[(\Sigma f_i u_i^2/N) - (\Sigma f_i u_i/N)^2]}$

where $u_i = (m_i - A)/h$, $N = \Sigma f$

### Variance:

Variance ($\sigma^2$) = (Standard Deviation)$^2$

---

### Solved Example (Raw Data):

Find σ for: 4, 6, 8, 10, 12

$\bar{x}$ = (4+6+8+10+12)/5 = 40/5 = 8

$(4-8)^2 + (6-8)^2 + (8-8)^2 + (10-8)^2 + (12-8)^2$
= 16 + 4 + 0 + 4 + 16 = 40

$\sigma^2 = 40/5 = 8$
$\sigma = \sqrt{8} = 2.83$

---

**Solved Example (Grouped Data):**

| Class | f | m | u=(m-25)/10 | fu | fu² |
|---|---|---|---|---|---|
| 0-10 | 5 | 5 | -2 | -10 | 20 |
| 10-20 | 8 | 15 | -1 | -8 | 8 |
| 20-30 | 12 | 25 | 0 | 0 | 0 |
| 30-40 | 10 | 35 | 1 | 10 | 10 |
| 40-50 | 5 | 45 | 2 | 10 | 20 |
| **Total** | **40** | | | **2** | **58** |

$$\sigma = h \times \sqrt{[(\Sigma fu^2/N) - (\Sigma fu/N)^2]}$$
$$= 10 \times \sqrt{[(58/40) - (2/40)^2]}$$
$$= 10 \times \sqrt{[1.45 - 0.0025]}$$
$$= 10 \times \sqrt{1.4475}$$
$$= 10 \times 1.203$$
$$= 12.03$$

---

## 3.6 Properties of Standard Deviation

1. $\sigma$ is always **non-negative** ($\sigma \geq 0$)
2. If all values are same, $\sigma = 0$
3. $\sigma$ is **independent of change of origin** (adding constant doesn't change $\sigma$)
4. $\sigma$ is **dependent on change of scale** (multiplying by k gives new $\sigma = |k|\sigma$)
5. **Combined SD** of two groups:

$$\sigma_{12}^2 = (n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)) / (n_1 + n_2)$$

where $d_1 = \bar{x}_1 - \bar{x}_{12}$, $d_2 = \bar{x}_2 - \bar{x}_{12}$

---

## 3.7 Coefficient of Variation (CV)

$$CV = (\sigma / \bar{x}) \times 100\ \%$$

**Use**: Comparing variability of two or more distributions.

**The distribution with smaller CV is more consistent/uniform.**

**Example**:

- Group A: $\bar{x} = 50$, $\sigma = 10$    CV = (10/50) × 100 = 20%

- Group B: $\bar{x} = 60$, $\sigma = 15$    CV = (15/60) × 100 = 25%

**Group A is more consistent** (lower CV).

---

## 3.8 Relationship between Measures of Dispersion

```
QD = (2/3) × SD    (approximately)
MD = (4/5) × SD    (approximately)
```

More precisely:

```
4 SD = 5 MD = 6 QD    (approximately)
```

---

# UNIT 4: SKEWNESS AND KURTOSIS

---

## 4.1 Skewness

**Skewness** measures the **asymmetry** of a frequency distribution.

### Types:

1. **Symmetrical Distribution**: Mean = Median = Mode, Skewness = 0
2. **Positively Skewed**: Mean > Median > Mode (tail extends to right)
3. **Negatively Skewed**: Mean < Median < Mode (tail extends to left)

---

## 4.2 Measures of Skewness

### A. Karl Pearson's Coefficient of Skewness:

$$Sk_p = (\text{Mean} - \text{Mode}) / \sigma$$

or (when mode is ill-defined):

$$Sk_p = 3(\text{Mean} - \text{Median}) / \sigma$$

**Range**: Usually between -3 and +3

### B. Bowley's Coefficient of Skewness:

$Sk = (Q_3 + Q_1 - 2Q_2) / (Q_3 - Q_1)$

where $Q_2$ = Median

**Range**: $-1 \leq Sk \leq +1$

### C. Kelly's Coefficient of Skewness:

$Sk_k = (P_{90} + P_{10} - 2P_{50}) / (P_{90} - P_{10})$

---

### Solved Example:

Given: Mean = 45, Median = 40, σ = 10

$Sk_p = 3(45 - 40)/10 = 15/10 = 1.5$ (Positively Skewed)

---

## 4.3 Moments

### Raw Moments (about origin):

$\mu' = \Sigma x / n$ (raw data)
$\mu' = \Sigma f x / \Sigma f$ (frequency distribution)

- $\mu'_1$ = Mean
- $\mu'_2 = \Sigma x^2/n$

### Central Moments (about mean):

$\mu = \Sigma(x - \bar{x}) / n$

- $\mu_1 = 0$ (always)
- $\mu_2$ = Variance = $\sigma^2$
- $\mu_3$ = Used for skewness
- $\mu_4$ = Used for kurtosis

### Relationship between Central and Raw Moments:

$\mu_1 = 0$
$\mu_2 = \mu'_2 - (\mu'_1)^2$

$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3$
$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4$

---

# 4.4 Skewness using Moments

### Karl Pearson's $\beta_1$ and $\gamma_1$:

$\beta_1 = \mu_3^2 / \mu_2^3$
$\gamma_1 = \sqrt{\beta_1} = \mu_3 / \mu_2$^(3/2)

- If $\gamma_1 = 0$ Symmetrical
- If $\gamma_1 > 0$ Positively skewed
- If $\gamma_1 < 0$ Negatively skewed

---

# 4.5 Kurtosis

**Kurtosis** measures the **peakedness** or **flatness** of a distribution.

### Karl Pearson's $\beta_2$ and $\gamma_2$:

$\beta_2 = \mu_4 / \mu_2^2$
$\gamma_2 = \beta_2 - 3$

### Types:

1. **Mesokurtic**: $\beta_2 = 3$, $\gamma_2 = 0$ (Normal distribution)
2. **Leptokurtic**: $\beta_2 > 3$, $\gamma_2 > 0$ (More peaked, heavier tails)
3. **Platykurtic**: $\beta_2 < 3$, $\gamma_2 < 0$ (Flatter, lighter tails)

---

### Solved Example on Moments:

Find first four central moments and $\beta_1$, $\beta_2$ for:

| x | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| f | 1 | 3 | 5 | 3 | 1 |

N = 13, $\bar{x} = \Sigma fx/N$ = (2+9+20+15+6)/13 = 52/13 = 4

$\mu_2 = \Sigma f(x-\bar{x})^2/N$ = [1(4)+3(1)+5(0)+3(1)+1(4)]/13 = 10/13 = 0.769

$\mu_3 = \Sigma f(x-\bar{x})^3/N$ = [1(-8)+3(-1)+5(0)+3(1)+1(8)]/13 = 0/13 = 0

$\mu_4 = \Sigma f(x-\bar{x})^4/N$ = [1(16)+3(1)+5(0)+3(1)+1(16)]/13 = 36/13 = 2.769

$\beta_1 = \mu_3^2/\mu_2^3 = 0/(0.769)^3 = 0$ Symmetrical
$\beta_2 = \mu_4/\mu_2^2 = 2.769/(0.769)^2 = 2.769/0.591 = 4.68$ Leptokurtic

---

# UNIT 5: CORRELATION ANALYSIS

---

## 5.1 Introduction

**Correlation** measures the **degree and direction of linear relationship** between two variables.

### Types:

1. **Positive Correlation**: Both variables move in same direction
2. **Negative Correlation**: Variables move in opposite directions
3. **Zero Correlation**: No linear relationship

### Degree:

- **Perfect**: $r = +1$ or $r = -1$
- **High**: $0.75 < |r| < 1$
- **Moderate**: $0.25 < |r| < 0.75$
- **Low**: $0 < |r| < 0.25$
- **No correlation**: $r = 0$

---

## 5.2 Scatter Diagram

A **scatter diagram** is a graph of plotted points showing the relationship between two variables.

- Points close to a line High correlation
- Points scattered Low correlation
- Upward trend Positive correlation
- Downward trend Negative correlation

---

## 5.3 Karl Pearson's Coefficient of Correlation (r)

### Formula:

$$r = \Sigma(x - \bar{x})(y - \bar{y}) / \sqrt{[\Sigma(x - \bar{x})^2 \times \Sigma(y - \bar{y})^2]}$$

### Shortcut Formula:

$$r = [n\Sigma x y - \Sigma x \Sigma y] / \sqrt{\{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]\}}$$

**Range**: $-1 \le r \le +1$

---

### Solved Example:

| x | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| y | 2 | 4 | 5 | 4 | 5 |

```
n = 5
Σx = 15,  Σy = 20
Σx² = 1+4+9+16+25 = 55
Σy² = 4+16+25+16+25 = 86
Σxy = 2+8+15+16+25 = 66

r = [5(66) - 15(20)] / √{[5(55)-(15)²][5(86)-(20)²]}
  = [330 - 300] / √{[275-225][430-400]}
  = 30 / √{50 × 30}
  = 30 / √1500
  = 30 / 38.73
  = 0.775
```

**Interpretation**: High positive correlation.

---

## 5.4 Properties of Correlation Coefficient

1. $-1 \le r \le +1$
2. r is dimensionless (unit-free)
3. r is independent of change of origin and scale
4. $r(x,y) = r(y,x)$ (symmetric)
5. If $r = 0$, variables are uncorrelated (not necessarily independent)
6. r measures only **linear** relationship

---

## 5.5 Spearman's Rank Correlation Coefficient (ρ)

Used when data is in **ranks** or **ordinal** form.

### When ranks are NOT repeated:

$$\rho = 1 - [6\Sigma D^2 / n(n^2-1)]$$

where D = Rank of x - Rank of y

### When ranks are REPEATED:

$$\rho = 1 - [6(\Sigma D^2 + CF) / n(n^2-1)]$$

where $CF = \Sigma[(m^3-m)/12]$ for each repeated rank (m = number of times repeated)

---

### Solved Example (No Repeat):

| Student | Rank in Math ($R_1$) | Rank in Stats ($R_2$) | $D = R_1-R_2$ | $D^2$ |
|---|---|---|---|---|
| A | 1 | 2 | -1 | 1 |
| B | 2 | 1 | 1 | 1 |
| C | 3 | 3 | 0 | 0 |
| D | 4 | 5 | -1 | 1 |
| E | 5 | 4 | 1 | 1 |

$n = 5, \Sigma D^2 = 4$
$\rho = 1 - [6(4) / 5(25-1)]$
$= 1 - [24/120]$
$= 1 - 0.2$
$= 0.8$

---

## 5.6 Coefficient of Determination

$$r^2 = \text{Explained Variation / Total Variation}$$

**Example**: If $r = 0.8$, then $r^2 = 0.64$ This means 64% of variation in y is explained by x.

---

# UNIT 6: REGRESSION ANALYSIS

---

## 6.1 Introduction

**Regression** is the study of the **functional relationship** between variables to **predict** one variable from the other.

### Difference from Correlation:

| Correlation | Regression |
|---|---|
| Measures degree of relationship | Gives equation of relationship |
| Symmetric ($r_{xy} = r_{yx}$) | Not symmetric |
| No prediction | Used for prediction |

---

## 6.2 Lines of Regression

### Line of Regression of y on x:

(Used to predict y from x)

$$y - \bar{y} = byx(x - \bar{x})$$

or

$$y = a + byx \times x$$

### Line of Regression of x on y:

(Used to predict x from y)

$$x - \bar{x} = bxy(y - \bar{y})$$

or

$$x = a' + bxy \times y$$

---

## 6.3 Regression Coefficients

### Regression Coefficient of y on x:

$$byx = \Sigma(x - \bar{x})(y - \bar{y}) / \Sigma(x - \bar{x})^2$$
$$= [n\Sigma xy - \Sigma x\Sigma y] / [n\Sigma x^2 - (\Sigma x)^2]$$

Also:

$$byx = r \times (\sigma y/\sigma x)$$

### Regression Coefficient of x on y:

$$bxy = \Sigma(x - \bar{x})(y - \bar{y}) / \Sigma(y - \bar{y})^2$$
$$= [n\Sigma xy - \Sigma x\Sigma y] / [n\Sigma y^2 - (\Sigma y)^2]$$

Also:

$$bxy = r \times (\sigma x/\sigma y)$$

---

## 6.4 Properties of Regression Coefficients

1. **$r^2$ = byx × bxy** $r = \pm\sqrt{(byx \times bxy)}$ (Sign of r = common sign of byx and bxy)
2. Both regression coefficients have the **same sign**
3. If one coefficient is > 1, the other must be < 1
4. **AM of byx and bxy ≥ r** $(byx + bxy)/2 \geq r$
5. Both lines pass through the point **($\bar{x}$, $\bar{y}$)**
6. $0 \leq byx \times bxy \leq 1$
7. Regression coefficients are **independent of change of origin** but **dependent on change of scale**

---

## 6.5 Angle between Regression Lines

$$\tan\theta = [(1-r^2) \times \sigma x \times \sigma y] / [r \times (\sigma x^2 + \sigma y^2)]$$

- If r = 0 lines are perpendicular (θ = 90°)
- If r = ±1 lines coincide (θ = 0°)

---

### Solved Example:

Given: n = 5

| x | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| y | 2 | 4 | 5 | 4 | 5 |

$$\Sigma x = 15, \Sigma y = 20, \Sigma x^2 = 55, \Sigma y^2 = 86, \Sigma xy = 66$$

$$byx = [5(66) - 15(20)] / [5(55) - (15)^2]$$
$$= [330-300] / [275-225]$$
$$= 30/50 = 0.6$$

$$bxy = [5(66) - 15(20)] / [5(86) - (20)^2]$$
$$= [330-300] / [430-400]$$
$$= 30/30 = 1.0$$

```
x̄ = 15/5 = 3, ȳ = 20/5 = 4
```

**Regression line of y on x**:

```
y - 4 = 0.6(x - 3)
y = 0.6x + 2.2
```

**Regression line of x on y**:

```
x - 3 = 1.0(y - 4)
x = y - 1
```

**Verification**: $r = \pm\sqrt{(0.6 \times 1.0)} = \pm\sqrt{0.6} = \pm 0.775$ (positive since both b's are positive)

---

# UNIT 7: PROBABILITY

---

## 7.1 Basic Concepts

### Experiment:

A process that produces well-defined outcomes.

### Random Experiment:

An experiment whose outcome cannot be predicted with certainty.

### Sample Space (S):

The set of all possible outcomes. **Example**: Tossing a coin S = {H, T} **Example**: Throwing a die S = {1, 2, 3, 4, 5, 6}

### Event:

A subset of the sample space.

### Types of Events:

- **Simple Event**: Contains only one outcome
- **Compound Event**: Contains more than one outcome

- **Sure Event**: S itself (probability = 1)
- **Impossible Event**:    (probability = 0)
- **Complementary Event**: A' or Ā (everything not in A)
- **Mutually Exclusive**: A    B =    (cannot occur together)
- **Exhaustive Events**: Together they cover entire sample space
- **Independent Events**: Occurrence of one doesn't affect other

---

## 7.2 Definition of Probability

### Classical Definition:

```
P(A) = Number of favorable outcomes / Total number of outcomes
     = n(A) / n(S)
```

### Properties:

1. 0 ≤ P(A) ≤ 1
2. P(S) = 1
3. P(  ) = 0
4. P(A') = 1 - P(A)

---

## 7.3 Important Counting Formulas

### Factorial:

```
n! = n × (n-1) × (n-2) × ... × 2 × 1
0! = 1
```

### Permutation (order matters):

```
ⁿP   = n! / (n-r)!
```

### Combination (order doesn't matter):

```
ⁿC   = n! / [r!(n-r)!]
```

---

## 7.4 Addition Theorem

### For two events:

```
P(A    B) = P(A) + P(B) - P(A    B)
```

### For mutually exclusive events (A    B =    ):

```
P(A    B) = P(A) + P(B)
```

### For three events:

```
P(A  B  C) = P(A)+P(B)+P(C)-P(A  B)-P(B  C)-P(A  C)+P(A  B  C)
```

---

### Solved Examples:

**Example 1**: A card is drawn from a pack of 52 cards. Find P(King or Heart).

```
P(King) = 4/52, P(Heart) = 13/52, P(King    Heart) = 1/52

P(King    Heart) = 4/52 + 13/52 - 1/52 = 16/52 = 4/13
```

**Example 2**: Two dice are thrown. Find P(sum = 7).

```
Favorable outcomes: (1,6),(2,5),(3,4),(4,3),(5,2),(6,1) = 6
Total outcomes: 6 × 6 = 36
P(sum=7) = 6/36 = 1/6
```

---

## 7.5 Multiplication Theorem & Conditional Probability

### Conditional Probability:

```
P(A|B) = P(A    B) / P(B)    [P(B) ≠ 0]
```

### Multiplication Theorem:

```
P(A    B) = P(A) × P(B|A)
        = P(B) × P(A|B)
```

### For Independent Events:

```
P(A    B) = P(A) × P(B)
```

---

### Solved Examples:

**Example 1**: A bag has 5 red and 3 blue balls. Two balls are drawn without replacement. Find P(both red).

```
P(1st red) = 5/8
P(2nd red | 1st red) = 4/7
P(both red) = 5/8 × 4/7 = 20/56 = 5/14
```

**Example 2**: P(A) = 0.4, P(B) = 0.3, A and B are independent. Find P(A    B).

```
P(A    B) = 0.4 × 0.3 = 0.12
```

---

## 7.6 Bayes' Theorem

If $B_1, B_2, \ldots, B_n$ are mutually exclusive and exhaustive events, then:

```
P(B   |A) = P(B  ) × P(A|B  ) / Σ[P(B  ) × P(A|B  )]
```

---

### Solved Example:

Three machines $M_1$, $M_2$, $M_3$ produce 30%, 45%, 25% of total items. Defective rates are 2%, 3%, 4% respectively. An item is randomly selected and found defective. Find P(from $M_2$).

$P(M_1) = 0.30$, $P(M_2) = 0.45$, $P(M_3) = 0.25$
$P(D|M_1) = 0.02$, $P(D|M_2) = 0.03$, $P(D|M_3) = 0.04$

$$
\begin{aligned}
P(D) &= 0.30(0.02) + 0.45(0.03) + 0.25(0.04) \\
&= 0.006 + 0.0135 + 0.01 \\
&= 0.0295
\end{aligned}
$$

$$
\begin{aligned}
P(M_2|D) &= P(M_2)P(D|M_2) / P(D) \\
&= 0.45(0.03) / 0.0295 \\
&= 0.0135 / 0.0295 \\
&= 0.4576 \approx 0.458
\end{aligned}
$$

---

---

# UNIT 8: RANDOM VARIABLES & PROBABILITY DISTRIBUTIONS

---

## 8.1 Random Variable

A **random variable** is a function that assigns a numerical value to each outcome of a random experiment.

### Types:

1. **Discrete Random Variable**: Takes countable values (0, 1, 2, ...)
2. **Continuous Random Variable**: Takes any value in an interval

---

## 8.2 Probability Distribution of a Discrete RV

A table showing all possible values with their probabilities.

### Conditions:

1. $P(x\ ) \geq 0$ for all i
2. $\Sigma P(x\ ) = 1$

### Expected Value (Mean):

```
E(X) = μ = Σx  P(x  )
```

### Variance:

```
Var(X) = σ² = E(X²) - [E(X)]²
```

where $E(X^2) = \Sigma x\ ^2 P(x\ )$

---

### Solved Example:

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| P(X) | 0.1 | 0.3 | 0.4 | 0.2 |

```
E(X) = 0(0.1) + 1(0.3) + 2(0.4) + 3(0.2) = 0 + 0.3 + 0.8 + 0.6 = 1.7

E(X²) = 0(0.1) + 1(0.3) + 4(0.4) + 9(0.2) = 0 + 0.3 + 1.6 + 1.8 = 3.7

Var(X) = 3.7 - (1.7)² = 3.7 - 2.89 = 0.81
SD = √0.81 = 0.9
```

---

## 8.3 Binomial Distribution

### Conditions:

1. Fixed number of trials (n)

2. Each trial has only two outcomes (success/failure)
3. Probability of success (p) is constant
4. Trials are independent

### Formula:

$P(X = r) = {}^{n}C \quad \times p \quad \times q^{n}$ (r = 0, 1, 2, ..., n)

where q = 1 - p

### Properties:

Mean = np
Variance = npq
SD = √(npq)

### Important Relations:

- Mode: If (n+1)p is integer two modes: (n+1)p and (n+1)p - 1 Otherwise Mode = floor[(n+1)p]
- Variance < Mean (always, since q < 1)

---

### Solved Example:

A coin is tossed 5 times. Find P(exactly 3 heads).

n = 5, p = 1/2, q = 1/2, r = 3

$P(X=3) = {}^{5}C_{3} \times (1/2)^{3} \times (1/2)^{2}$
= 10 × 1/8 × 1/4
= 10/32
= 5/16
= 0.3125

**Mean** = 5 × 0.5 = 2.5 **Variance** = 5 × 0.5 × 0.5 = 1.25

---

### Another Example:

20% of bolts are defective. From a sample of 10, find: (a) P(none defective) (b) P(exactly 2 defective) (c) Mean and Variance

n = 10, p = 0.2, q = 0.8

(a) $P(X=0) = {}^{10}C_{0} \times (0.2)^{0} \times (0.8)^{10} = 1 \times 1 \times 0.1074 = 0.1074$

(b) $P(X=2) = {}^{10}C_{2} \times (0.2)^{2} \times (0.8)^{8} = 45 \times 0.04 \times 0.1678 = 0.3020$

(c) Mean = 10(0.2) = 2, Variance = 10(0.2)(0.8) = 1.6

---

## 8.4 Poisson Distribution

### Conditions:

1. Events occur independently
2. Average rate (λ) is constant
3. Events occur one at a time
4. Used for **rare events** (n is large, p is small)

### Formula:

$P(X = r) = (e\ \ \lambda \times \lambda\ \ ) / r!$     (r = 0, 1, 2, ...)

where λ = np = mean

### Properties:

Mean = λ
Variance = λ
Mean = Variance (unique property!)
SD = √λ

---

### Solved Example:

Average number of phone calls per minute = 3. Find P(exactly 5 calls in a minute).

λ = 3, r = 5

$P(X=5) = (e^{\ 3} \times 3^5) / 5!$
= (0.0498 × 243) / 120
= 12.10 / 120
= 0.1008

---

### Another Example:

If 2% of items are defective. Find probability that a box of 100 items contains: (a) No defective (b) At most 2 defective

λ = np = 100 × 0.02 = 2

(a) $P(X=0) = e^{\ 2} \times 2^0/0! = 0.1353$

(b) $P(X \leq 2) = P(0) + P(1) + P(2)$
$= e^{\ 2}[2^0/0! + 2^1/1! + 2^2/2!]$
$= 0.1353[1 + 2 + 2]$
$= 0.1353 \times 5$
$= 0.6767$

---

## 8.5 Normal Distribution

### Properties:

1. Continuous, bell-shaped, symmetric about mean
2. Mean = Median = Mode
3. Defined by two parameters: $\mu$ (mean) and $\sigma$ (standard deviation)
4. Total area under curve = 1
5. Extends from $-\infty$ to $+\infty$
6. Points of inflection at $\mu \pm \sigma$

### PDF:

$$f(x) = (1/\sigma\sqrt{2\pi}) \times e^{[-(x-\mu)^2/2\sigma^2]} \quad (-\infty < x < \infty)$$

### Standard Normal Variable:

$$Z = (X - \mu) / \sigma$$

$Z \sim N(0, 1)$ [mean = 0, SD = 1]

### Area Properties:

$P(\mu - \sigma < X < \mu + \sigma) \approx 0.6827$ (68.27%)
$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.9545$ (95.45%)
$P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.9973$ (99.73%)

---

### Common Z-table Values:

| Z | Area (0 to Z) |
|---|---|
| 0.0 | 0.0000 |
| 0.5 | 0.1915 |
| 1.0 | 0.3413 |
| 1.5 | 0.4332 |
| 1.96 | 0.4750 |
| 2.0 | 0.4772 |

| | |
|---|---|
| 2.5 | 0.4938 |
| 3.0 | 0.4987 |

---

**Solved Example:**

Marks of students are normally distributed with μ = 60, σ = 10. Find: (a) P(X > 75) (b) P(45 < X < 70)

```
(a) Z = (75-60)/10 = 1.5
   P(X>75) = P(Z>1.5) = 0.5 - 0.4332 = 0.0668

(b) Z₁ = (45-60)/10 = -1.5
   Z₂ = (70-60)/10 = 1.0
   P(45<X<70) = P(-1.5<Z<1.0)
             = P(-1.5<Z<0) + P(0<Z<1.0)
             = 0.4332 + 0.3413
             = 0.7745
```

---

---

# UNIT 9: INDEX NUMBERS

---

## 9.1 Definition

An **Index Number** is a statistical measure designed to show changes in a variable or group of related variables over time, geographical location, or other characteristics.

**Base Period**: The period with which comparisons are made (index = 100) **Current Period**: The period being compared

---

## 9.2 Types of Index Numbers

1. **Price Index Numbers** – Measure changes in price
2. **Quantity Index Numbers** – Measure changes in quantity
3. **Value Index Numbers** – Measure changes in total value

---

## 9.3 Methods of Construction

### A. Simple (Unweighted) Index Numbers

**1. Simple Aggregative Method:**

$$P_{01} = (\Sigma p_1 / \Sigma p_0) \times 100$$

**2. Simple Average of Relatives Method:**

Using AM:

$$P_{01} = (1/n) \times \Sigma(p_1/p_0 \times 100)$$

Using GM:

$$P_{01} = \text{Antilog}[(1/n) \times \Sigma\log(p_1/p_0 \times 100)]$$

---

## B. Weighted Index Numbers

**1. Laspeyre's Index (Base year quantities as weights):**

$$P_{01}(L) = (\Sigma p_1q_0 / \Sigma p_0q_0) \times 100$$

**2. Paasche's Index (Current year quantities as weights):**

$$P_{01}(P) = (\Sigma p_1q_1 / \Sigma p_0q_1) \times 100$$

**3. Fisher's Index (Geometric mean of Laspeyre's and Paasche's):**

$$P_{01}(F) = \sqrt{[P_{01}(L) \times P_{01}(P)]}$$
$$= \sqrt{[(\Sigma p_1q_0/\Sigma p_0q_0) \times (\Sigma p_1q_1/\Sigma p_0q_1)]} \times 100$$

**4. Marshall-Edgeworth Index:**

$$P_{01}(ME) = [\Sigma p_1(q_0+q_1) / \Sigma p_0(q_0+q_1)] \times 100$$

**5. Bowley's Index:**

$$P_{01}(B) = [P_{01}(L) + P_{01}(P)] / 2$$

**6. Dorbish-Bowley Index:**

$$P_{01}(DB) = [L + P] / 2$$

---

### Solved Example:

| Commodity | $p_0$ | $q_0$ | $p_1$ | $q_1$ | $p_1q_0$ | $p_0q_0$ | $p_1q_1$ | $p_0q_1$ |
|---|---|---|---|---|---|---|---|---|
| A | 10 | 50 | 12 | 55 | 600 | 500 | 660 | 550 |
| B | 8 | 30 | 10 | 35 | 300 | 240 | 350 | 280 |
| C | 5 | 40 | 6 | 45 | 240 | 200 | 270 | 225 |
| **Total** | | | | | **1140** | **940** | **1280** | **1055** |

```
Laspeyre's: L = (1140/940) × 100 = 121.28

Paasche's: P = (1280/1055) × 100 = 121.33

Fisher's: F = √(121.28 × 121.33) = √14714.58 = 121.30
```

---

## 9.4 Tests of Adequacy of Index Numbers

### 1. Time Reversal Test (TRT):

$P_{01} \times P_{10} = 1$ (or 100 × 100 if in percentages)

| Method | TRT Satisfied? |
|---|---|
| Laspeyre's | No |
| Paasche's | No |
| Fisher's | Yes |
| Marshall-Edgeworth | Yes |

### 2. Factor Reversal Test (FRT):

$P_{01} \times Q_{01} = (\Sigma p_1q_1 / \Sigma p_0q_0)$

(Price index × Quantity index = Value ratio)

| Method | FRT Satisfied? |
|---|---|
| Laspeyre's | No |
| Paasche's | No |
| Fisher's | Yes |

> **Fisher's Index is called "IDEAL" index number** because it satisfies both TRT and FRT.

### 3. Circular Test:

$P_{01} \times P_{12} \times P_{20} = 1$

- Satisfied by: Simple Aggregative Method

- Not satisfied by: Laspeyre's, Paasche's, Fisher's

---

# 9.5 Cost of Living Index (CLI) / Consumer Price Index (CPI)

## Methods:

### A. Aggregate Expenditure Method:

$$CLI = (\Sigma p_1 q_0 / \Sigma p_0 q_0) \times 100$$

### B. Family Budget Method (Weighted Average of Price Relatives):

$$CLI = \Sigma RW / \Sigma W$$

where $R = (p_1/p_0) \times 100$ (Price Relative), $W = p_0 q_0$ (Weight)

---

## Solved Example (Family Budget Method):

| Item | $p_0$ | $p_1$ | $q_0$ | $R=p_1/p_0\times100$ | $W=p_0q_0$ | RW |
|---|---|---|---|---|---|---|
| Food | 20 | 25 | 10 | 125 | 200 | 25000 |
| Cloth | 15 | 18 | 5 | 120 | 75 | 9000 |
| Rent | 50 | 55 | 1 | 110 | 50 | 5500 |
| Fuel | 10 | 14 | 8 | 140 | 80 | 11200 |
| **Total** | | | | | **405** | **50700** |

$$CLI = 50700/405 = 125.19$$

**Interpretation**: Cost of living has increased by 25.19% compared to base period.

---

# 9.6 Uses of Index Numbers

1. Measure changes in price levels (inflation)
2. Measure changes in industrial/agricultural production
3. Deflating income/wage data to find real values
4. Guide for policy making
5. Dearness Allowance (DA) calculation

## Real Income:

$$\text{Real Income} = (\text{Money Income} / \text{Price Index}) \times 100$$

---

# UNIT 10: TIME SERIES ANALYSIS

---

## 10.1 Definition

A **Time Series** is a set of observations recorded at successive points of time or over successive periods of time.

**Examples**: Monthly sales, yearly population, daily stock prices

---

## 10.2 Components of Time Series

### 1. Secular Trend (T):

Long-term general tendency (increase or decrease)

### 2. Seasonal Variation (S):

Regular fluctuations within a year (due to seasons, festivals)

### 3. Cyclical Variation (C):

Oscillations over periods longer than a year (business cycles)

### 4. Irregular/Random Variation (I):

Unpredictable fluctuations (wars, earthquakes, pandemics)

---

## 10.3 Models of Time Series

### Additive Model:

```
Y = T + S + C + I
```

### Multiplicative Model:

```
Y = T × S × C × I
```

---

# 10.4 Measurement of Trend

### A. Graphic/Freehand Method:

Draw a smooth curve through the plotted data points by inspection.

### B. Semi-Average Method:

1. Divide data into two equal halves
2. Calculate average of each half
3. Plot these averages at mid-points of their periods
4. Join the two points    Trend line

---

### C. Moving Average Method:

Calculate averages of successive groups of fixed size.

**3-year Moving Average**: Average of 3 consecutive values

**4-year Moving Average**: Requires centering (average of two consecutive 4-year averages)

**Solved Example (3-year Moving Average)**:

| Year | Value | 3-yr Total | 3-yr MA |
|---|---|---|---|
| 2015 | 10 | — | — |
| 2016 | 12 | 10+12+14=36 | 12.0 |
| 2017 | 14 | 12+14+16=42 | 14.0 |
| 2018 | 16 | 14+16+18=48 | 16.0 |
| 2019 | 18 | 16+18+20=54 | 18.0 |
| 2020 | 20 | — | — |

---

### D. Method of Least Squares (Best Method)

Fit a straight line: **Y = a + bX**

### Normal Equations:

$$\Sigma Y = na + b\Sigma X$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

### If origin is shifted to make ΣX = 0:

$$a = \Sigma Y/n$$
$$b = \Sigma XY/\Sigma X^2$$

---

### Solved Example:

| Year | Y | X (coding) | XY | $X^2$ |
|---|---|---|---|---|
| 2016 | 10 | -2 | -20 | 4 |
| 2017 | 15 | -1 | -15 | 1 |
| 2018 | 20 | 0 | 0 | 0 |
| 2019 | 22 | 1 | 22 | 1 |
| 2020 | 28 | 2 | 56 | 4 |
| **Total** | **95** | **0** | **43** | **10** |

```
a = ΣY/n = 95/5 = 19
b = ΣXY/ΣX² = 43/10 = 4.3

Trend Line: Y = 19 + 4.3X (Origin: 2018)

Forecast for 2021: X = 3
Y = 19 + 4.3(3) = 19 + 12.9 = 31.9
```

---

## 10.5 Seasonal Variation — Ratio to Trend Method

1. Calculate trend values
2. Express original values as percentage of trend: (Y/T) × 100
3. Average these percentages for each season
4. Adjust so total = 400 (for quarterly) or 1200 (for monthly)

---

## 10.6 Seasonal Variation — Simple Average Method

1. Calculate average for each season (quarter/month)
2. Calculate grand average
3. Seasonal Index = (Seasonal Average / Grand Average) × 100

### Solved Example:

| Quarter | 2018 | 2019 | 2020 | Average |
|---|---|---|---|---|
| Q1 | 40 | 42 | 44 | 42 |
| Q2 | 50 | 52 | 54 | 52 |
| Q3 | 45 | 48 | 51 | 48 |
| Q4 | 35 | 38 | 41 | 38 |

```
Grand Average = (42+52+48+38)/4 = 180/4 = 45

Seasonal Index:
Q1 = (42/45) × 100 = 93.33
Q2 = (52/45) × 100 = 115.56
Q3 = (48/45) × 100 = 106.67
Q4 = (38/45) × 100 = 84.44
```

Total = 400

---

# IMPORTANT FORMULAS — QUICK REFERENCE SHEET

## Central Tendency:

AM = $\Sigma fx/\Sigma f$
GM = Antilog($\Sigma f \cdot \log x / \Sigma f$)
HM = $\Sigma f / \Sigma(f/x)$
Median = $L + [(N/2 - CF)/f] \times h$
Mode = $L + [(f_1-f_0)/(2f_1-f_0-f_2)] \times h$
Mode = 3 Median - 2 Mean

## Dispersion:

Range = Max - Min
QD = $(Q_3-Q_1)/2$
MD = $\Sigma f|x-\bar{x}| / \Sigma f$
$\sigma = \sqrt{[\Sigma f(x-\bar{x})^2/\Sigma f]}$
CV = $(\sigma/\bar{x}) \times 100\%$

## Correlation & Regression:

$r = [n\Sigma xy - \Sigma x\Sigma y] / \sqrt{\{[n\Sigma x^2-(\Sigma x)^2][n\Sigma y^2-(\Sigma y)^2]\}}$
$\rho = 1 - 6\Sigma D^2/n(n^2-1)$
$b_{yx} = r(\sigma_y/\sigma_x)$
$b_{xy} = r(\sigma_x/\sigma_y)$
$r^2 = b_{yx} \times b_{xy}$

## Probability:

P(A  B) = P(A) + P(B) - P(A  B)
P(A  B) = P(A)·P(B|A)
P(B  |A) = P(B  )P(A|B  ) / ΣP(B  )P(A|B  )

## Distributions:

Binomial: P(X=r) = $^{n}C$  p  $q^{n}$  , μ=np, $\sigma^2$=npq
Poisson: P(X=r) = e  λλ  /r!, μ=$\sigma^2$=λ
Normal: Z = (X-μ)/σ

## Index Numbers:

L = $(\Sigma p_1 q_0/\Sigma p_0 q_0)\times 100$
P = $(\Sigma p_1 q_1/\Sigma p_0 q_1)\times 100$
F = √(L×P)

---

This covers the **complete BCA Semester 2 Statistics syllabus** with:

- All topics explained with theory
- All formulas listed
- Solved numerical examples for every topic
- Properties and important notes
- Quick revision formula sheet