

For the students of Basic Statistics to understand the  
Statistical Concepts and Techniques

# FOUNDATIONS OF STATISTICS

**Dr Debdulal Dutta Roy**

Rabindrik Psychotherapy Research Institute Trust, Kolkata

**Dr E.M. Naresh Babu**

ABBS School of Management, Bengaluru

**Dr Kshitiz Sharma**

ABBS School of Management, Bengaluru



*Rita Publication*

*(Publisher and Book-seller)*

25B, Beniatola Lane, Kolkata-700 009  
[www.ritapublication.com](http://www.ritapublication.com)/[www.ritapublication.in](http://www.ritapublication.in)

## CONTENTS

|                              |     |             |
|------------------------------|-----|-------------|
| <b>1. DATA SUMMARIZATION</b> | ... | <b>1-26</b> |
|------------------------------|-----|-------------|

### 1.1. Frequency Distribution 1

- 1.1.1. Absolute Frequency 2
- 1.1.2. Relative Frequency 2
- 1.1.3. Cumulative Frequency 2
  - R-Script for Frequency Distribution
  - SPSS Output for Frequency Distribution
  - Minitab Output for Frequency Distribution
  - R-Script for Data Visualization

### 1.2. Percentage 7

- 1.2.1. Calculation of Percentage 7
- 1.2.2. Application of Percentage Distribution and Analysis 10

### 1.3. Percentile 10

- 1.3.1. Uses of Percentile in Education 11
- 1.3.2. Uses of Percentile in Business 11
  - R-Script for Percentiles
  - SPSS Out for Percentiles

### 1.4. Central Tendency 13

- 1.4.1. Mean 13
- 1.4.2. Uses of Mean in Different Fields 13
- 1.4.3. Median 13
- 1.4.4. Applications of Median in Different Fields 13
- 1.4.5. Mode 14
- 1.4.6. Applications of Mode in Different Fields 14
- 1.4.7. Differences among Mean, Median and Mode
  - R-Script for Central Tendency
  - SPSS Output for Central Tendency
  - Minitab Output for Central Tendency

### 1.5. Variability 17

- 1.5.1. Range 17
- 1.5.2. Five-Number Summary 17
- 1.5.3. Interquartile Range and Semi Interquartile Range 18
- 1.5.4. Standard Deviation 20
- 1.5.5. Application areas of Standard Deviation 20
- 1.5.6. Variance 20
  - R-Script for Measures of Variability
  - SPSS Output for Measures of Variability
  - Minitab Output Measures of Variability

- **Exercises 24**

## 2. NORMAL PROBABILITY CURVE CHARACTERISTICS

27-53

### 2.1. Normal Probability Curve Characteristics 27

- 2.1.1. History of Normal Distribution 27
- 2.1.2. The Central Limit Theorem 28
- 2.1.3. Characteristics of Normal Curve 29
- 2.1.4. Area Estimation under Normal Curve Probability Area 30
- 2.1.5. Application Areas of Normal Curve 31
  - R Output for application areas of Normal Curve
  - SPSS Output for application areas of Normal Curve
  - Minitab Output for application areas of Normal Curve
- 2.1.6. Confidence Interval 37
  - R Output for Confidence Interval

### 2.2. Skewness 38

- 2.2.1. Skewness – Meaning 38
- 2.2.2. History of Skewness 38
- 2.2.3. Characteristics of Skewness 38
- 2.2.4. Types of Skewness 38
- 2.2.5. Application areas of Skewness 40
  - R Script for Histogram and Skewness

### 2.3. Kurtosis 42

- 2.3.1. Types of Kurtosis 44
- 2.3.2. Application areas of Kurtosis 44
  - R Script for Kurtosis

### • Exercises 51

## 3. PARAMETRIC STATISTICS

...

54-108

### 3.1. Parametric Statistics 54

### 3.2. Differences between Parametric and Non-Parametric Statistics 54

### 3.3. Mean Differences 54

- 3.3.1. t-tests 55
- 3.3.2. Analysis of Variance (ANOVA) 55
  - 3.3.2.1. One-way ANOVA 56
  - 3.3.2.2. Steps in Solving the ANOVA 56
    - R-Script for Percentiles
    - Problems and applications of ANOVA
    - R Output
    - Minitab Output
  - 3.3.2.3. Factorial ANOVA 69
  - 3.3.2.4. Mixed ANOVA 70
  - 3.3.2.5. Analysis of Covariance (ANCOVA) 71
    - 3.3.2.5.1. Types of ANCOVA 71

### 3.4. Correlation 72

- 3.4.1. Assumptions related to Correlation 72
- 3.4.2. Types of Correlation 73

- 3.4.3. Co-efficient of Correlation 74
- 3.4.4. Utility of Measuring Correlation 74
- 3.4.5. Application Areas of Correlation in Business 75
  - R Output related to Correlation application in business
  - SPSS Output related to Correlation application in business

### **3.5. Regression 91**

- 3.5.1. History of Regression 92
- 3.5.2. Assumptions related to Regression 92
- 3.5.3. Application of Regression 92
- 3.5.4. Forms of Regression 93
  - R Output related to Regression
  - SPSS Output related to Regression

### **3.6. Differences between Correlation and Regression 100**

- Exercises 101

## **4. NON-PARAMETRIC STATISTICS**

...

**109-257**

### **4.1. Chi-Square Test 109**

- 4.1.1. Assumptions of Chi-Square Test 109
- 4.1.2. Application Areas of Chi Square 110
  - 4.1.2.1. Goodness of Fit 110
    - R Output for Chi Square test – Goodness of Fit
    - SPSS Output for Chi Square test – Goodness of Fit
    - Minitab Output for Chi Square test – Goodness of Fit
  - 4.1.2.2. Test of Association 122
    - R Output for Chi Square test – Test of Association
    - SPSS Output for Chi Square test – Test of Association
    - Minitab Output for Chi Square test – Test of Association
- 4.1.3. Yates Correction for Continuity 137
  - R Output for Yates Correction
  - SPSS Output for Yates Correction
  - Minitab Output for Yates Correction

### **4.2. Spearman's Rank Order Correlation 141**

- 4.2.1. Assumptions related to Spearman's Rank Order Correlation 141
- 4.2.2. Procedure of Finding Rank Order Correlation 142
  - R Output for Spearman's Rank Order Correlation
  - SPSS Output for Spearman's Rank Order Correlation
  - Minitab Output for Spearman's Rank Order Correlation

### **4.3. Wilcoxon Rank Sum test or Mann-Whitney 'U' Test 164**

- 4.3.1. History of Wilcoxon Test 164
- 4.3.2. Assumptions related to Mann-Whitney 'U' Test 164
- 4.3.3. Steps in Performing Mann-Whitney 'U' Test or Wilcoxon Rank Sum Test 165
  - R Output for Wilcoxon Rank Sum test
  - SPSS Output for Wilcoxon Rank Sum test
  - Minitab Output for Wilcoxon Rank Sum test

#### **4.4. Sign Test 190**

4.4.1. Assumptions related to Sign Test 191

4.4.2. Steps in Performing Sign Test 191

- R Output for Wilcoxon Rank Sum Test
- SPSS Output for Wilcoxon Rank Sum Test
- Minitab Output for Wilcoxon Rank Sum Test

#### **4.5. Wilcoxon Signed Rank Test 219**

4.5.1. Assumptions of Wilcoxon Signed Rank Test 219

4.5.2. Steps in Performing Wilcoxon Signed Rank Test 219

- R Output for Wilcoxon Signed Rank Test
- SPSS Output for Wilcoxon Signed Rank Test
- Minitab Output for Wilcoxon Signed Rank Test

#### **4.6. Kruskal-Wallis Test 240**

4.6.1. Assumptions related to Kruskal-Wallis Test 240

4.6.2. Steps in Performing Kruskal-Wallis Test 240

- R Output for Kruskal-Wallis Test
- SPSS Output for Kruskal-Wallis Test
- Minitab Output for Kruskal-Wallis Test
- **Exercises 255**

### **Appendix: Statistical Tables**

...

**258-272**

- Z Table 258
- F Table 263
- Chi-Square Table 264
- SPSS Installation Process 265
- Minitab Installation Process 271

