Second Edition

Statistics Made Simple

Do It Yourself on PC

K.V.S. Sarma
To
my mother—
the real motherboard in me
Contents

Preface xi
Preface to the First Edition xiii

1. The Role of Statistics in Research 1–27
   1.1 Statistics in Research 1
   1.2 Research in Statistics 2
   1.3 Common Statistical Issues in Research 3
   1.4 Data Collection 4
      1.4.1 Survey Method 5
      1.4.2 Experimental Method 7
   1.5 Coding of Data 8
   1.6 Tabulation and Presentation of Data 9
   1.7 Some Case Studies in Statistical Analysis 10
   1.8 The Statistics Toolkit 1 11
      1.8.1 Tabulation of Data—One-way Frequencies 11
      1.8.2 Cross Tabulations—Two-way Frequencies 13
      1.8.3 Histogram—The Graphic Way of Describing a Variable 15
      1.8.4 Bars, Lines and Pie Diagrams as Visual Aids 16
   1.9 The Statistics Toolkit 2 19
      1.9.1 Summary Statistics 19
      1.9.2 The Population and the Sample 23
      1.9.3 Sample Statistics and Their Formulae 24
   1.10 Role of Computer in Statistical Analysis 26

References 27
Suggested Readings 27
Do It Yourself 27

2. Basics of a Computer 28–37
   2.1 The Evolution of Computer 28
   2.2 The Personal Computer (PC) and Its Components 29
   2.3 Peripheral Devices for the Computer 32
   2.4 Getting Started with the PC 35
   2.5 Hard Disk Partitions, Files and Folders 36

References 37
Do It Yourself 37
3. Windows and MS-Office for Research Studies  38–62
   3.1 About Windows XP/Vista  38
   3.2 Windows Desktop and Start Menu  39
   3.3 My Computer and Windows Explorer  43
   3.4 About MS-Office 2007  49
   3.5 About MS-Word 2007  51
   3.6 Opening MS-Word  51
   3.7 Creating a Document  53
   3.8 Saving and Printing the Document  56
   3.9 Editing Features  58
   3.10 Review and Proofing  60
   References  61
   Do It Yourself  61

4. Data Handling and Statistics through MS-Access  63–91
   4.1 Introduction  63
   4.2 Creating MS-Access Database  63
   4.3 Creating a Table  67
   4.4 Creating a Form for Data Entry  70
   4.5 Creating Queries  75
   4.6 Creating Reports  79
   4.7 Data File for Blood Bank Operations—A Case of Hospital Management  81
   4.8 Data File for Plant Growth Problem—A Case of Experimental Data  85
   4.9 Data File for the Tribal Food Problem  87
   References  90
   Suggested Readings  91
   Do It Yourself  91

5. Data Handling in Excel  92–117
   5.1 Getting Started with Excel 2007  92
   5.2 The Excel Worksheet  93
   5.3 Data Entry on the Worksheet  98
   5.4 Some Important Editing Features of Excel  100
   5.5 Calculations on the Worksheet  103
   5.6 Built-in Functions for Quick Use  106
   5.7 Editing Data Using Excel Form  110
   5.8 Array Functions  111
   5.9 Printing the Data and Results  115
   References  116
   Do It Yourself  116
6. **Graphs and Charts in Excel** 118–137
   6.1 Excel Charts and Application Areas 118
   6.2 Construction of a Column/Bar Chart 120
   6.3 Improving the Chart 122
   6.4 Construction of a Histogram 129
   6.5 Construction of a PIE Chart 130
   6.6 Construction of a Line Chart 133
   6.7 Construction of a Scatter Chart 135
   References 136
   Do It Yourself 136

7. **Descriptive Statistics Using Excel** 138–153
   7.1 Constructing a Histogram 138
   7.2 Data Analysis Pak in Excel 140
   7.3 Descriptive Statistics 141
      7.3.1 Summary Statistics 142
      7.3.2 Frequency Distribution and Histogram 144
   7.4 Cross-Tabulations and Pivot Tables 146
   7.5 The Concept of Probability 150
   References 151
   Do It Yourself 151

8. **Inferential Statistics Using Excel** 154–185
   8.1 Estimation of Unknown Parameter Values 154
   8.2 Testing of Hypotheses 155
   8.3 Statistical Tests Concerning Means 159
      8.3.1 The One-sample Z-Test for Mean 159
      8.3.2 The One-sample t-Test for Mean 160
      8.3.3 The Two-sample Z-Test for Means 162
      8.3.4 The Two-sample t-Test for Means 164
      8.3.5 The Paired t-Test 166
   8.4 The F-Test for Variance 168
   8.5 Analysis of Variance 169
      8.5.1 One-way Anova 170
      8.5.2 Two-way Anova with Replication 174
      8.5.3 Two-way Anova without Replication 177
   8.6 The Chi-Square Test 178
      8.6.1 Chi-square Test for Goodness of Fit 178
      8.6.2 Chi-square Test for Independence 181
   References 182
   Do It Yourself 183

9. **Correlation and Regression Analysis** 186–206
   9.1 Correlation Analysis 186
   9.2 Simple Regression Analysis 190
9.3 Multiple Linear Regression  197
9.4 Diagnostic Analysis of Regression  201
References  204
Do It Yourself  204

10. Data Analysis Using Crystal Reports  207–230

10.1 Introduction  207
10.2 Salient Features of Crystal Reports  207
10.3 The Standard Report Wizard  209
   10.3.1 Connecting Data File to the Report  211
   10.3.2 Report Design  213
10.4 Editing the Report  214
10.5 Formula Fields  219
   10.5.1 Formula to Concatenate Fields  219
   10.5.2 Formula to Use a Condition  220
10.6 The Select Expert  221
10.7 Other Important Features of a Report  224
10.8 Cross Tab Expert  226
10.9 Chart Expert  228
References  230
Do It Yourself  230


11.1 Fitting a Trend Line to the Observed Data  231
11.2 Polynomial Trends  235
11.3 Logarithmic, Power and Exponential Trends  237
11.4 Moving Averages  239
11.5 Exponential Smoothing  242
11.6 Linear and Compound Growth Values  244
   11.6.1 The Forecast and Trend Functions  244
   11.6.2 The Growth Function  246
11.7 Financial Functions and Related Tools in Excel  248
   11.7.1 Simple Financial Functions  248
   11.7.2 Ranking of Data  249
   11.7.3 Random Number Generation  250
   11.7.4 Creation of Statistical Tables  251
References  254
Do It Yourself  254


12.1 The SPSS Software  256
12.2 Creating a Data File in SPSS  258
12.3 Editing Features of SPSS  262
12.4 Creating Frequency Tables  267
12.5 Cross Tabulations and Chi-square Tests  270
## CONTENTS

12.6 Tests of Significance Using SPSS 273  
12.7 Analysis of Variance 277  
12.8 Multiple Comparison Tests 280  

References 285  
Do It Yourself 285  

## Appendices 287–299

A. Data of the File C:\STATMAN\TRIBAL.DBF 287  
B. Data of the File C:\STATMAN\BLOOD.DBF 293  
C. Data of the File C:\STATMAN\PLANT.DBF 295  
D. Values of Student’s t-distribution for Different Degrees of Freedom (df) at Selected Level of Significance (α) 297  
E. Cumulative Standard Normal Distribution Values from $-\infty$ to $z$ 298  

## Index 301–303
Preface

The first edition of this book was based on the facilities available in Microsoft Office 2000. Nowadays the user is familiar with Microsoft Office 2007 and would probably shift to new versions that are likely to be available in the near future. The Operating Systems used for discussion in the first edition were DOS and Windows (98/2000), while most of the computers now operate with Windows XP/Vista.

Since one of the objectives of this book is to make the reader familiar with the software for statistics, it became necessary to revise the contents to fit into the Office 2007 environment. The revision has been made by effecting the following changes in the contents of the book.

1. The earlier Chapter 4 Windows and MS-Office for research studies has been changed to Chapter 3 with modified contents.
2. The earlier Chapter 3 Data handling and statistics through FoxPro has been removed and in its place the new Chapter 4 Data handling and statistics through MS-Access is introduced. Since FoxPro is not much in use, MS-Access has been introduced in this book. A detailed discussion of activities such as creating data files, manipulation through queries and making reports has been provided with the help of practical and live data sets. New examples have been introduced to enable understanding of Access data.
3. The earlier Chapter 10 FoxPro programming for quick statistics has been replaced with Data analysis using crystal reports. New software called Crystal Reports (CRP), which is currently used by software developers for making reports from databases, has been introduced in this edition. A large portion of data analysis in social science research can be carried out with CRP. It is a highly useful tool for producing reports from Access and Excel files.
4. The discussion in Chapter 12 is totally devoted to SPSS (with the title suitably changed). The formats used in the earlier versions of SPSS have been replaced with the contemporary versions, and the statistical aspects are thoroughly discussed with examples.
Most statistical books are focused on statistical theory versus application, but Statistics Made Simple for School Leaders presents statistics in a simple, practical, conceptual, and immediately applicable manner. It enables administrators to take their data and manage it into strategic information so the results can be used for action plans that benefit the school system.