Understanding Statistics for the Social Sciences with IBM SPSS

Robert Ho

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Contents

Preface xiii
Author xv

1. Introduction to the Scientific Methodology of Research 1
1.1 Introduction 1
1.2 The Scientific Approach versus the Layperson's Approach to Knowledge 1
1.2.1 Sampling 2
1.2.2 Research Designs 2
1.2.3 Between-Groups Design 2
1.3 The Univariate Approach 3
1.3.1 The Multivariate Approach 3
1.4 Correlational Design 5
1.5 Hypothesis Testing and Probability Theory 5
1.5.1 Probability 6
1.5.2 Statistics and Scientific Research 6
1.6 Definition of Statistics 6
1.6.1 Descriptive Statistics 6
1.6.2 Inferential Statistics 7

2. Introduction to SPSS 9
2.1 Learning How to Use the SPSS Software Program 9
2.2 Introduction to SPSS 10
2.2.1 Setting Up a Data File 11
2.2.2 Preparing a Codebook 11
2.2.3 Data Set 11
2.2.4 Creating SPSS Data File 12
2.2.5 Data Entry 16
2.2.6 Saving and Editing Data File 17
2.3 SPSS Analysis: Windows Method versus Syntax Method 17
2.4 SPSS Analysis: Windows Method 17
2.5 SPSS Analysis: Syntax Method 19
2.5.1 SPSS Output 21
2.5.2 Results and Interpretation 22

Section I Descriptive Statistics

3. Basic Mathematical Concepts and Measurement 27
3.1 Basic Mathematical Concepts 27

3.1.1 Mathematical Notations ... 27
3.2 Measurement Scales (Levels of Measurement) ... 29
3.2.1 Nominal Scales ... 29
3.2.2 Ordinal Scales ... 30
3.2.3 Interval Scales ... 30
3.2.4 Ratio Scales ... 31
3.3 Types of Variables ... 31
3.3.1 IV and DV ... 31
3.3.2 Continuous and Discrete Variables ... 32
3.3.3 Real Limits of Continuous Variables ... 32
3.3.4 Rounding ... 33

4. Frequency Distributions ... 35
4.1 Ungrouped Frequency Distributions ... 35
4.1.1 SPSS: Data Entry Format ... 36
4.1.2 SPSS Windows Method ... 36
4.1.3 SPSS Syntax Method ... 38
4.1.4 SPSS Output ... 39
4.1.5 Results and Interpretation ... 40
4.2 Grouped Frequency Distributions ... 41
4.2.1 Grouping Scores into Class Intervals ... 41
4.2.2 Computing a Frequency Distribution of Grouped Scores ... 42
4.2.3 SPSS Method ... 44
4.2.4 SPSS Windows Method ... 46
4.2.5 SPSS Syntax Method ... 47
4.2.6 SPSS Output ... 48
4.3 Percentiles and Percentile Ranks ... 49
4.3.1 Percentiles ... 49
4.3.2 Computation of Percentiles (Finding the Score below which a Specified Percentage of Scores will Fall) ... 49
4.3.3 SPSS Syntax Method ... 50
4.3.4 Data Entry Format ... 50
4.3.5 SPSS Syntax Method ... 51
4.3.6 SPSS Output ... 52
4.3.7 Another Example ... 52
4.3.8 Data Entry Format ... 53
4.3.9 SPSS Syntax Method ... 53
4.3.10 SPSS Output ... 54
4.3.11 Percentile Rank ... 55
4.3.12 Computation of Percentile Ranks (Finding the Percentage of Scores that Fall below a Given Score) ... 55
4.3.13 Data Entry Format ... 56
4.3.14 SPSS Syntax Method ... 56
4.3.15 SPSS Output ... 58

4.3.16 Another Example 58
4.3.17 Data Entry Format 59
4.3.18 SPSS Syntax Method 59
4.3.19 SPSS Output 60

5. Graphing 61
5.1 Graphing Frequency Distributions 61
5.2 Bar Graph 61
5.2.1 An Example 61
5.2.2 Data Entry Format 62
5.2.3 SPSS Windows Method 63
5.2.4 SPSS Syntax Method 64
5.2.5 SPSS Bar Graph Output 66
5.3 Histogram 67
5.3.1 An Example 68
5.3.2 SPSS Windows Method 68
5.3.3 SPSS Syntax Method 69
5.3.4 SPSS Histogram Output 71
5.4 Frequency Polygon 71
5.4.1 An Example 72
5.4.2 SPSS Windows Method 73
5.4.3 SPSS Syntax Method 74
5.4.4 SPSS Frequency Polygon Output 76
5.5 Cumulative Percentage Curve 77
5.5.1 An Example 78
5.5.2 SPSS Windows Method 79
5.5.3 SPSS Syntax Method 80
5.5.4 SPSS Cumulative Percentage Output 82

6. Measures of Central Tendency 85
6.1 Why Is Central Tendency Important? 85
6.2 Measures of Central Tendency 86
6.3 The Arithmetic Mean 86
6.3.1 How to Calculate the Arithmetic Mean 87
6.3.2 SPSS Window Method 87
6.3.3 SPSS Syntax Method 89
6.3.4 SPSS Output 90
6.3.5 How to Calculate the Mean from a Grouped Frequency Distribution 91
6.3.6 An Example 91
6.3.7 Calculating the Mean from Grouped Frequency Distribution Using SPSS 92
6.3.8 Data Entry Format 92
6.3.9 SPSS Syntax Method 92
6.3.10 SPSS Output 93

6.3.11 The Overall Mean 94
6.3.12 An Example 96
6.3.13 How to Calculate the Overall Mean Using SPSS 96
6.3.14 Data Entry Format 96
6.3.15 SPSS Syntax Method 97
6.3.16 SPSS Output 98
6.3.17 Properties of the Mean 98
6.4 The Median 101
6.4.1 Calculating the Median for Ungrouped Scores 101
6.4.2 Calculating the Median for Grouped Scores 102
6.4.3 Properties of the Median 102
6.5 The Mode 104
6.5.1 SPSS Windows Method 105
6.5.2 SPSS Syntax Method 106
6.5.3 SPSS Histogram Output 107
6.5.4 The Mode for Grouped Scores 107
6.6 Comparison of the Mean, Median, and Mode 107
6.7 Measures of Central Tendency: Symmetry and Skewness 109

7. Measures of Variability/Dispersion 111
7.1 What Is Variability? 111
7.2 Range 112
7.3 Standard Deviation 113
7.3.1 Calculating the Standard Deviation Using the Deviation Scores Method 113
7.3.2 Calculating the Standard Deviation Using the Raw Scores Method 116
7.4 Variance 117
7.5 Using SPSS to Calculate the Range, the Standard Deviation, and the Variance 117
7.5.1 SPSS Windows Method 117
7.5.2 SPSS Syntax Method 120
7.5.3 SPSS Output 121

8. The Normal Distribution and Standard Scores 123
8.1 The Normal Distribution 123
8.2 Areas Contained under the Standard Normal Distribution 123
8.3 Standard Scores (z Scores) and the Normal Curve 124
8.3.1 Calculating the Percentile Rank with z Scores 126
8.3.2 SPSS Windows Method 126
8.3.3 SPSS Syntax Method 127
8.3.4 SPSS Data File Containing the First 10 Computed z Scores 128
8.3.5 Calculating the Percentage of Scores that Fall between Two Known Scores 129

8.3.6 Calculating the Percentile Point with *z* Scores 130
8.3.7 SPSS Windows Method 132
8.3.8 SPSS Syntax Method 134
8.3.9 Table Showing the 90th Percentile for the Set of 50 Exam Scores 135
8.3.10 Calculating the Scores that Bound a Specified Area of the Distribution 135
8.3.11 SPSS Windows Method 138
8.3.12 SPSS Syntax Method 140
8.3.13 Table from either Window or Syntax Methods for Displaying Lower and Upper Bound Scores Binding the Middle 70% of the EX11.SAV data set 141
8.3.14 Using *z* Scores to Compare Performance between Different Distributions 141

9. Correlation 145
9.1 The Concept of Correlation 145
9.2 Linear and Nonlinear Relationships 145
9.3 Characteristics of Correlation 147
9.3.1 Magnitude (Strength) of Relationships 148
9.3.2 Direction of Relationships 148
9.4 Correlation Coefficient and *z* Scores 149
9.4.1 Scatter Plot (SPSS Windows Method) 150
9.4.2 Scatter Plot (SPSS Syntax Method) 152
9.4.3 Scatter Plot 153
9.4.4 Converting Raw Scores into *z* Scores (SPSS Windows Method) 153
9.4.5 Converting Raw Scores into *z* Scores (SPSS Syntax Method) 154
9.4.6 SPSS Data File Containing the 6 Pairs of Computed *z* Scores 155
9.5 Pearson *r* and the Linear Correlation Coefficient 155
9.5.1 Example of the Pearson *r* Calculation 156
9.5.2 SPSS Windows Method 157
9.5.3 SPSS Syntax Method 158
9.5.4 The Calculated Pearson *r* 159
9.6 Some Issues with Correlation 160
9.6.1 Can Correlation Show Causality? 160
9.6.2 Spurious Correlation 161

10. Linear Regression 163
10.1 What Is Linear Regression? 163
10.2 Linear Regression and Imperfect Relationships 164
10.2.1 Scatter Plot and the Line of Best Fit 164

10.2.2 SPSS Windows Method (Scatter Plot and Line of Best Fit) 164
10.2.3 SPSS Syntax Method (Scatter Plot) 167
10.2.4 Scatter Plot with Line of Best Fit 169
10.2.5 Least-Squares Regression (Line of Best Fit): Predicting *Y* from *X* 169
10.2.6 How to Construct the Least-Squares Regression Line: Predicting *Y* from *X* 170
10.2.7 SPSS Windows Method (Constructing the Least-Squares Regression Line Equation) 172
10.2.8 SPSS Syntax Method (Constructing the Least-Squares Regression Line Equation) 175
10.2.9 SPSS Output 175
10.2.10 Results and Interpretation 176

Section II Inferential Statistics

11. Statistical Inference and Probability 179
11.1 Introduction to Inferential Statistics 179
11.2 Probability 180
11.2.1 The Classical Approach to Probability 180
11.2.2 The Empirical Approach to Probability 181
11.2.3 Expressing Probability Values 181
11.2.4 Computing Probability: The Addition Rule and the Multiplication Rule 182
11.2.5 The Addition Rule 182
11.2.6 The Multiplication Rule 185
11.2.7 Using the Multiplication and Addition Rules Together 188
11.2.8 Computing Probability for Continuous Variables 190
11.3 Sampling 192
11.3.1 Simple Random Sampling 192
11.3.2 Stratified Proportionate Random Sampling 193
11.3.3 Cluster Sampling 194
11.3.4 Nonrandom Sampling Techniques: Systematic Sampling; Quota Sampling 195
11.3.5 Sampling with or without Replacement 195
11.4 Confidence Interval and Confidence Level 196
11.4.1 How to Calculate the Confidence Interval 198
11.4.2 SPSS Windows Method 198
11.4.3 SPSS Syntax Method 200
11.4.4 SPSS Output 202

12. Introduction to Hypothesis Testing 203
12.1 Introduction to Hypothesis Testing 203
12.2 Types of Hypotheses 205
12.2.1 Research/Alternative Hypothesis 205
12.2.2 Null Hypothesis 205
12.2.3 Hypotheses: Nondirectional or Directional 206
12.3 Testing Hypotheses 206
12.3.1 Level of Significance 207
12.3.2 Two-Tailed and One-Tailed Test of Significance 209
12.3.3 Type I and Type II Errors 210

13. Hypothesis Testing: *t* test for Independent and Correlated Groups 215
13.1 Introduction to the *t* test 215
13.2 Independent *t* test 217
13.2.1 SPSS Windows Method: Independent *t* test 220
13.2.2 SPSS Syntax Method 222
13.2.3 SPSS Output 222
13.2.4 Results and Interpretation 222
13.3 Dependent/Correlated *t* test 224
13.3.1 SPSS Windows Method: Dependent *t* test 227
13.3.2 SPSS Syntax Method 228
13.3.3 SPSS Output 228
13.3.4 Results and Interpretation 228

14. Hypothesis Testing: One-Way Analysis of Variance 229
14.1 One-Way Analysis of Variance 229
14.1.1 An Example 229
14.1.2 Scheffé *Post Hoc* Test 232
14.1.3 SPSS Windows Method: One-Way ANOVA 233
14.1.4 SPSS Syntax Method 235
14.1.5 SPSS Output 236
14.1.6 Results and Interpretation 237
14.1.7 Post Hoc Comparisons 237

15. Hypothesis Testing: Chi-Square Test 239
15.1 Nonparametric Tests 239
15.2 Chi-Square (χ^2) Test 239
15.2.1 Chi-Square Goodness-of-Fit Test 240
15.2.2 SPSS Windows Method 242
15.2.3 SPSS Syntax Method 243
15.2.4 SPSS Output 243
15.2.5 Results and Interpretation 244

15.3 Chi-Square (χ^2) Test of Independence between Two Variables 244
15.3.1 SPSS Windows Method 248
15.3.2 SPSS Syntax Method 251
15.3.3 SPSS Output 251
15.3.4 Results and Interpretation 252

Appendix 253

Bibliography 267

Index 269