

Contents

<i>Preface</i>	<i>page</i> xiii
<i>Acknowledgments</i>	xv
1 INTRODUCTION	1
The Role of the Computer in Data Analysis	1
Statistics: Descriptive and Inferential	2
Variables and Constants	2
The Measurement of Variables	3
Discrete and Continuous Variables	8
Setting a Context with Real Data	11
Exercises	12
2 EXAMINING UNIVARIATE DISTRIBUTIONS	20
Counting the Occurrence of Data Values	20
When Variables Are Measured at the Nominal Level	20
Bar Graphs, 21 • Pie Graphs, 23	
When Variables Are Measured at the Ordinal, Interval, or Ratio Level	24
Frequency and Percent Distribution Tables, 24 •	
Stem-and-Leaf Displays, 26 • Histograms, 29 •	
Line Graphs, 31	
Describing the Shape of a Distribution	33
Accumulating Data	35
Cumulative Percent Distributions	35
Ogive Curves	35
Percentile Ranks	36
Percentiles	37
Five-Number Summaries and Boxplots	40
Exercises	45
3 MEASURES OF LOCATION, SPREAD, AND SKEWNESS	60
Characterizing the Location of a Distribution	60
The Mode	60
The Median	63
The Arithmetic Mean	65
Comparing the Mode, Median, and Mean	67
Characterizing the Spread of a Distribution	70
The Range and Interquartile Range	72
The Variance	74
The Standard Deviation	76

vi	CONTENTS	
	Characterizing the Skewness of a Distribution	77
	Selecting Measures of Location and Spread	78
	Applying What We Have Learned	79
	Exercises	82
4	RE-EXPRESSING VARIABLES	91
	Linear and Nonlinear Transformations	91
	Linear Transformations: Addition, Subtraction, Multiplication, and Division	91
	The Effect on the Shape of a Distribution	93
	The Effect on Summary Statistics of a Distribution	95
	Common Linear Transformations	95
	Standard Scores	97
	z-Scores	98
	Using z-Scores to Detect Outliers, 100 • Using z-Scores to	
	Compare Scores in Different Distributions, 101 •	
	Relating z-Scores to Percentile Ranks, 102	
	Nonlinear Transformations: Square Roots and Logarithms	103
	Nonlinear Transformations: Ranking Variables	110
	Other Transformations: Recoding and Combining Variables	111
	Recoding Variables	111
	Combining Variables	113
	Exercises	114
5	EXPLORING RELATIONSHIPS BETWEEN TWO VARIABLES	121
	When Both Variables Are at Least Interval-Leveled	121
	Scatterplots	122
	The Pearson Product Moment Correlation Coefficient	126
	Interpreting the Pearson Correlation Coefficient, 130 •	
	The Effect of Linear Transformations, 132 • Restriction	
	of Range, 132 • The Shape of the Underlying Distributions, 133 •	
	The Reliability of the Data, 133	
	When at Least One Variable Is Ordinal and the Other Is at Least Ordinal:	
	The Spearman Rank Correlation Coefficient	133
	When at Least One Variable Is Dichotomous: Other Special Cases of the Pearson	
	Correlation Coefficient	135
	The Point Biserial Correlation Coefficient: The Case of One at-Least-Interval and One	
	Dichotomous Variable	135
	The Phi Coefficient: The Case of Two Dichotomous Variables	140
	Other Visual Displays of Bivariate Relationships	144
	Selection of Appropriate Statistic/Graph to Summarize a Relationship	147
	Exercises	148
6	SIMPLE LINEAR REGRESSION	158
	The “Best-Fitting” Linear Equation	158
	The Accuracy of Prediction Using the Linear Regression Model	164
	The Standardized Regression Equation	165
	R as a Measure of the Overall Fit of the Linear Regression Model	165
	Simple Linear Regression When the Independent Variable Is Dichotomous	169
	Using <i>r</i> and <i>R</i> as Measures of Effect Size	172

CONTENTS

Emphasizing the Importance of the Scatterplot	172
Exercises	174
7 PROBABILITY FUNDAMENTALS	182
The Discrete Case	182
The Complement Rule of Probability	184
The Additive Rules of Probability	184
First Additive Rule of Probability	185
Second Additive Rule of Probability	186
The Multiplicative Rule of Probability	187
The Relationship between Independence and Mutual Exclusivity	189
Conditional Probability	190
The Law of Large Numbers	191
Exercises	192
8 THEORETICAL PROBABILITY MODELS	195
The Binomial Probability Model and Distribution	195
The Applicability of the Binomial Probability Model	200
The Normal Probability Model and Distribution	204
Using the Normal Distribution to Approximate the Binomial Distribution	210
Exercises	210
9 THE ROLE OF SAMPLING IN INFERENCE STATISTICS	217
Samples and Populations	217
Random Samples	218
Obtaining a Simple Random Sample	219
Sampling with and without Replacement	221
Sampling Distributions	223
Describing the Sampling Distribution of Means Empirically	223
Describing the Sampling Distribution of Means Theoretically: The Central Limit Theorem	226
Central Limit Theorem (CLT)	227
Estimators and BIAS	230
Exercises	231
10 INFERENCES INVOLVING THE MEAN OF A SINGLE POPULATION WHEN σ IS KNOWN	234
Estimating the Population Mean μ When the Population Standard Deviation σ Is Known	234
Interval Estimation	236
Relating the Length of a Confidence Interval, the Level of Confidence, and the Sample Size	239
Hypothesis Testing	239
The Relationship between Hypothesis Testing and Interval Estimation	247
Effect Size	248
Type II Error and the Concept of Power	249
Increasing the Level of Significance, α	253
Increasing the Effect Size, δ	253
Decreasing the Standard Error of the Mean, $\sigma_{\bar{x}}$	253

viii	CONTENTS	
	Closing Remarks	254
	Exercises	254
11	INFERENCES INVOLVING THE MEAN WHEN σ IS NOT KNOWN: ONE- AND TWO-SAMPLE DESIGNS	259
	Single Sample Designs When the Parameter of Interest Is the Mean and σ Is Not Known	259
	The t -Distribution	260
	Degrees of Freedom for the One-Sample t -Test	261
	Violating the Assumption of a Normally Distributed Parent Population in the One-Sample t -Test	262
	Confidence Intervals for the One-Sample t -Test	263
	Hypothesis Tests: The One-Sample t -Test	267
	Effect Size for the One-Sample t -Test	269
	Two-Sample Designs When the Parameter of Interest Is μ , and σ Is Not Known	273
	Independent (or Unrelated) and Dependent (or Related) Samples	274
	Independent Samples t -Test and Confidence Interval	275
	The Assumptions of the Independent Samples t -Test	277
	Effect Size for the Independent Samples t -Test, 285	
	Paired Samples t -Test and Confidence Interval	288
	The Assumptions of the Paired Samples t -Test	289
	Effect Size for the Paired Samples t -Test	293
	Summary	294
	The Standard Error of the Mean Difference for Independent Samples: A More Complete Account (Optional)	295
	Case 1: σ Known	295
	Case 2: σ Not Known	297
	Step 1: Estimating σ^2 Using the Variance Estimators $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$, 299 •	
	Step 2: Estimating the Standard Error of the Mean Difference, $\sigma_{\bar{X}_1 - \bar{X}_2}$ Using σ^2 , 299	
	Exercises	300
12	ONE-WAY ANALYSIS OF VARIANCE	315
	The Disadvantage of Multiple t -Tests	315
	The One-Way Analysis of Variance	317
	A Graphical Illustration of the Role of Variance in Tests on Means	317
	ANOVA as an Extension of the Independent Samples t -Test	318
	Developing an Index of Separation for the Analysis of Variance	319
	Carrying Out the ANOVA Computation	319
	The Between-Group Variance (MS_B), 320 • The Within-Group Variance (MS_W), 321	
	The Assumptions of the One-Way ANOVA	321
	Testing the Equality of Population Means: The F -Ratio	322
	How to Read the Tables and to Use the SPSS Compute Statement for the F -Distribution	324
	ANOVA Summary Table	327

CONTENTS

Measuring the Effect Size	328
Post-HOC Multiple Comparison Tests	330
The Bonferroni Adjustment: Testing Planned Comparisons	340
The Bonferroni Tests on Multiple Measures	343
Exercises	345
13 TWO-WAY ANALYSIS OF VARIANCE	350
The Two-Factor Design	350
The Concept of Interaction	353
The Hypotheses That Are Tested by a Two-Way Analysis of Variance	358
Assumptions of the Two-Way ANOVA	358
Balanced versus Unbalanced Factorial Designs	360
Partitioning the Total Sum of Squares	360
Using the <i>F</i> -Ratio to Test the Effects in Two-Way ANOVA	361
Carrying Out the Two-Way ANOVA Computation by Hand	362
Decomposing Score Deviations about the Grand Mean	366
Modeling Each Score as a Sum of Component Parts	367
Explaining the Interaction as a Joint (or Multiplicative) Effect	368
Measuring Effect Size	368
Fixed versus Random Factors	372
Post-Hoc Multiple Comparison Tests	373
Summary of Steps to Be Taken in a Two-Way ANOVA Procedure	379
Exercises	383
14 CORRELATION AND SIMPLE REGRESSION AS INFERENTIAL TECHNIQUES	391
The Bivariate Normal Distribution	391
Testing Whether the Population Pearson Product Moment Correlation Equals Zero	394
Using a Confidence Interval to Estimate the Size of the Population Correlation Coefficient, ρ	397
Revisiting Simple Linear Regression for Prediction	400
Estimating the Population Standard Error of Prediction, $\sigma_{Y X}$	400
Testing the <i>b</i> -Weight for Statistical Significance	401
Explaining Simple Regression Using an Analysis of Variance Framework	405
Measuring the Fit of the Overall Regression Equation: Using <i>R</i> and <i>R</i> ²	407
Relating <i>R</i> ² to $\sigma_{Y X}^2$	408
Testing <i>R</i> ² for Statistical Significance	409
Estimating the True Population <i>R</i> ² : The Adjusted <i>R</i> ²	409
Exploring the Goodness of Fit of the Regression Equation: Using Regression Diagnostics	410
Residual Plots: Evaluating the Assumptions Underlying Regression, 413 •	
Detecting Influential Observations: Discrepancy and Leverage, 415 •	
Using SPSS to Obtain Leverage, 417 • Using SPSS to Obtain Discrepancy, 417 • Using SPSS to Obtain Influence, 418	
Using the Prediction Model to Predict Ice Cream Sales	422
Simple Regression When the Predictor Is Dichotomous	422
Exercises	424

x	CONTENTS	
15	AN INTRODUCTION TO MULTIPLE REGRESSION	435
	The Basic Equation with Two Predictors	436
	Equations for b , β , and R_{Y12} When the Predictors Are Not Correlated	437
	Equations for b , β , and R_{Y12} When the Predictors Are Correlated	438
	Summarizing and Expanding on Some Important Principles of Multiple Regression	440
	Testing the b -Weights for Statistical Significance	444
	Assessing the Relative Importance of the Independent Variables in the Equation	445
	Measuring the Decrease in R^2 Directly: An Alternative to the Squared	
	Part Correlation	446
	Evaluating the Statistical Significance of the Change in R^2	446
	The b -Weight as a Partial Slope in Multiple Regression	448
	Multiple Regression When One of the Two Independent Variables Is Dichotomous	450
	The Concept of Interaction between Two Variables That Are at Least Interval-Leveled	454
	Testing the Statistical Significance of an Interaction Using SPSS	456
	Centering First-Order Effects to Achieve Meaningful Interpretations of b -Weights	460
	Understanding the Nature of a Statistically Significant Two-Way Interaction	460
	Interaction When One of the Independent Variables Is Dichotomous and the	
	Other Is Continuous	463
	Putting It All Together: A Student Project Reprinted	466
	Measuring the Variables	467
	Examining the Variables Individually and in Pairs	468
	Examining the Variables Multivariately with Mathematics Achievement as the Criterion	471
	Exercises	475
16	NONPARAMETRIC METHODS	485
	Parametric versus Nonparametric Methods	485
	Nonparametric Methods When the Dependent Variable Is at the Nominal Level	486
	The Chi-Square Distribution (χ^2)	486
	The Chi-Square Goodness-of-Fit Test	489
	The Chi-Square Test of Independence	493
	Assumptions of the Chi-Square Test of Independence, 497	
	Fisher's Exact Test	499
	Calculating the Fisher Exact Test by Hand Using the	
	Hypergeometric Distribution, 501	
	Nonparametric Methods When the Dependent Variable Is Ordinal-Leveled	505
	Wilcoxon Sign Test	505
	The Mann-Whitney U -Test	508
	The Kruskal–Wallis Analysis of Variance	512
	Exercises	514
	APPENDIX A. DATA SET DESCRIPTIONS	519
	Anscombe	519
	Basket	519
	Blood	519
	Brainsz	520
	Currency	520
	Framingham	520
	Hamburg	522

CONTENTS

Ice Cream	522
Impeach	522
Learndis	523
Mandex.sav	524
Marijuan	524
NELS	524
Skulls	528
States	529
Stepping	529
Temp	530
Wages	530

APPENDIX B. GENERATING DISTRIBUTIONS FOR CHAPTERS 8 AND 9 USING SPSS SYNTAX 531

(1) Creating a New Data Set File with ID Values for 75 Cases	531
(2) The SPSS Syntax Program (Called, in General, a Macro) to Generate the Set of 50,000 Sample Means Used to Form the Sampling Distribution of Means Graphed as the Histogram of Figure 9.2	532
(3) The SPSS Syntax Program to Generate the Set of 1,000 Normally Distributed Scores with Mean = 15 and SD = 3 as Illustrated by the Histogram of Figure 9.3	533
(4) The SPSS Syntax Program to Generate the Set of 1,000 Normally Distributed Scores with Mean = 15 and SD = 3 as Illustrated by the Histogram of Figure 9.4	534
(5) The SPSS Syntax Program to Generate the Set of 1,000 Positively Skewed Distributed Scores with Mean = 8 and SD = 4 as Illustrated by the Histogram of Figure 9.5	534
(6) The SPSS Syntax Program, Sampdisver2.Sps, to Generate the Set of 5,000 Sample Means Used to Form the Sampling Distribution of Means Graphed as the Histogram of Figure 9.6.	535

APPENDIX C. STATISTICAL TABLES 537

Table 1. Areas under the Standard Normal Curve (to the Right of the z-Score)	537
Table 2. Distribution of <i>t</i> -Values for Right-Tail Areas	538
Table 3. Distribution of <i>F</i> -Values for Right-Tail Areas	539
Table 4. Binomial Distribution Table	543
Table 5. Chi-Square Distribution Values for Right-Tailed Areas	548
Table 6. The Critical <i>q</i> -Values	549
Table 7. The Critical <i>U</i> -Values	550

APPENDIX D. REFERENCES 554

APPENDIX E. SOLUTIONS TO EXERCISES 557

Chapter 1. Solutions	557
Chapter 2. Solutions	559
Chapter 3. Solutions	579
Chapter 4. Solutions	597
Chapter 5. Solutions	607
Chapter 6. Solutions	626
Chapter 7. Solutions	640
Chapter 8. Solutions	641

Cambridge University Press
978-0-521-67637-3 - Statistics Using SPSS: An Integrative Approach, Second Edition
Sharon Lawner Weinberg and Sarah Knapp Abramowitz
Table of Contents
[More information](#)

xii	CONTENTS	
	Chapter 9. Solutions	644
	Chapter 10. Solutions	648
	Chapter 11. Solutions	649
	Chapter 12. Solutions	673
	Chapter 13. Solutions	689
	Chapter 14. Solutions	703
	Chapter 15. Solutions	715
	Chapter 16. Solutions	743
	<i>Index</i>	752