From *SAS® Statistics by Example*. Full book available for purchase here.

Contents

List of Programs ix
Acknowledgments xv

Chapter 1 An Introduction to SAS 1

Introduction 2
What is SAS 2
Statistical Tasks Performed by SAS 3
The Structure of SAS Programs 3
SAS Data Sets 3
SAS Display Manager 4
Excel Workbooks 5
Variable Types in SAS Data Sets 11
Temporary versus Permanent SAS Data Sets 11
Creating a SAS Data Set from Raw Data 12
Data Values Separated by Delimiters 12
Reading CSV Files 14
Data Values in Fixed Columns 15
Excel Files with Invalid SAS Variable Names 16
Other Sources of Data 17
Conclusions 17

Chapter 2 Descriptive Statistics – Continuous Variables 19

Introduction 19
Computing Descriptive Statistics Using PROC MEANS 21
Descriptive Statistics Broken Down by a Classification Variable 23
Computing a 95% Confidence Interval and the Standard Error 25
Producing Descriptive Statistics, Histograms, and Probability Plots 26

Changing the Midpoint Values on the Histogram 32
Generating a Variety of Graphical Displays of Your Data 33
Displaying Multiple Box Plots for Each Value of a Categorical Variable 38
Conclusions 39

Chapter 3 Descriptive Statistics – Categorical Variables 41

Introduction 41
Computing Frequency Counts and Percentages 42
Computing Frequencies on a Continuous Variable 44
Using Formats to Group Observations 45
Histograms and Bar Charts 48
Creating a Bar Chart Using PROC SGPLOT 49
Using ODS to Send Output to Alternate Destinations 50
Creating a Cross-Tabulation Table 52
Changing the Order of Values in a Frequency Table 53
Conclusions 55

Chapter 4 Descriptive Statistics – Bivariate Associations 57

Introduction 57
Producing a Simple Scatter Plot Using PROG GPLOT 58
Producing a Scatter Plot Using PROC SGPLOT 61
Creating Multiple Scatter Plots on a Single Page Using PROC SGSCATTER 63
Conclusions 68

Chapter 5 Inferential Statistics – One-Sample Tests 69

Introduction 69
Conducting a One-Sample *t*-test Using PROC TTEST 70
Running PROC TTEST with ODS Graphics Turned On 71
Conducting a One-Sample *t*-test Using PROC UNIVARIATE 74
Testing Whether a Distribution is Normally Distributed 76

Tests for Other Distributions 78
Conclusions 78

Chapter 6 Inferential Statistics – Two-Sample Tests 79

Introduction 79
Conducting a Two-Sample *t*-test 79
Testing the Assumptions for a *t*-test 81
Customizing the Output from ODS Statistical Graphics 83
Conducting a Paired *t*-test 85
Assumption Violations 87
Conclusions 89

Chapter 7 Inferential Statistics – Comparing More than Two Means 91

Introduction 91
A Simple One-way Design 92
Conducting Multiple Comparison Tests 98
Using ODS Graphics to Produce a Diffogram 101
Two-way Factorial Designs 102
Analyzing Factorial Models with Significant Interactions 106
Analyzing a Randomized Block Design 108
Conclusions 109

Chapter 8 Correlation and Regression 111

Introduction 111
Producing Pearson Correlations 112
Generating a Correlation Matrix 115
Creating HTML Output with Data Tips 117
Generating Spearman Nonparametric Correlations 119
Running a Simple Linear Regression Model 120
Using ODS Statistical Graphics to Investigate Influential Observations 125

Using the Regression Equation to Do Prediction 129
A More Efficient Way to Compute Predicted Values 132
Conclusions 134

Chapter 9 Multiple Regression 135

Introduction 135
Fitting Multiple Regression Models 136
Running All Possible Regressions with *n* Variables 137
Producing Separate Plots Instead of a Panel 141
Choosing the Best Model (C_p and Hocking's Criteria) 142
Forward, Backward, and Stepwise Selection Methods 145
Forcing Selected Variables into a Model 152
Creating Dummy (Design) Variables for Regression 153
Detecting Collinearity 155
Influential Observations in Multiple Regression Models 157
Conclusions 161

Chapter 10 Categorical Data 163

Introduction 163
Comparing Proportions 164
Rearranging Rows and Columns in a Table 166
Tables with Expected Values Less Than 5 (Fisher's Exact Test) 169
Computing Chi-Square from Frequency Data 172
Using a Chi-Square Macro 173
A Short-Cut Method for Requesting Multiple Tables 175
Computing Coefficient Kappa—A Test of Agreement 175
Computing Tests for Trends 178
Computing Chi-Square for One-Way Tables 180
Conclusions 182

Chapter 11 Binary Logistic Regression 183

Introduction 183
Running a Logistic Regression Model with One Categorical Predictor Variable 184
Running a Logistic Regression Model with One Continuous Predictor Variable 189
Using a Format to Create a Categorical Variable from a Continuous Variable 191
Using a Combination of Categorical and Continuous Variables in a Logistic Regression Model 193
Running a Logistic Regression with Interactions 197
Conclusions 203

Chapter 12 Nonparametric Tests 205

Introduction 205
Performing a Wilcoxon Rank-Sum Test 206
Performing a Wilcoxon Signed-Rank Test (for Paired Data) 209
Performing a Kruskal-Wallis One-Way ANOVA 210
Comparing Spread: The Ansari-Bradley Test 211
Converting Data Values into Ranks 213
Using PROC RANK to Group Your Data Values 216
Conclusions 217

Chapter 13 Power and Sample Size 219

Introduction 219
Computing the Sample Size for an Unpaired *t*-Test 220
Computing the Power of an Unpaired *t*-Test 222
Computing Sample Size for an ANOVA Design 225
Computing Sample Sizes (or Power) for a Difference in Two Proportions 227
Using the SAS Power and Sample Size Interactive Application 229
Conclusions 234

Chapter 14 Selecting Random Samples 235

Introduction..... 235
Taking a Simple Random Sample 236
Taking a Random Sample with Replacement 237
Creating Replicate Samples using PROC SURVEYSELECT..... 240
Conclusions 241

References 243

Index 245

From *SAS® Statistics by Example* by Ron Cody. Copyright © 2013, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.