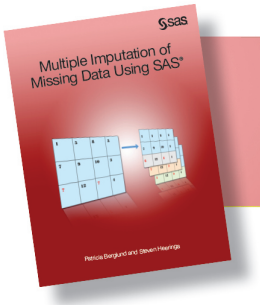

From *Multiple Imputation of Missing Data Using SAS®, Third Edition.* Full book available for purchase here.

Contents

About This Book vii
About The Authors ix
Acknowledgements xi

Chapter 1: Introduction to Missing Data and Methods for Analyzing Data with Missing Values 1
1.1 Introduction 1
1.2 Sources and Patterns of Item Missing Data 2
1.3 Item Missing Data Mechanisms 4
1.4 Review of Strategies to Address Item Missing Data 4
1.4.1 Complete Case Analysis 4
1.4.2 Complete Case Analysis with Weighting Adjustments 5
1.4.3 Full Information Maximum Likelihood 5
1.4.4 Expectation-Maximization Algorithm 5
1.4.5 Single Imputation of Missing Values 6
1.4.6 Multiple Imputation 6
1.5 Outline of Book Chapters 7
1.6 Overview of Analysis Examples 7
Chapter 2: Introduction to Multiple Imputation Theory and Methods 11
2.1 The Origins and Properties of Multiple Imputation Methods for Missing Data 11
2.1.1 A Short History of Imputation Methods 11
2.1.2 Why the Multiple Imputation Method? 12
2.1.3 Overview of Multiple Imputation Steps 14
2.2 Step 1—Defining the Imputation Model 16
2.2.1 Choosing the Variables to Include in the Imputation Model 16
2.2.2 Distributional Assumptions for the Imputation Model 17
2.3 Algorithms for the Multiple Imputation of Missing Values 17
2.3.1 General Theory for Multiple Imputation Algorithms 17
2.3.2 Methods for Monotone Missing Data Patterns 19
2.3.3 Methods for Arbitrary Missing Data Patterns 23
2.4 Step 2—Analysis of the MI Completed Data Sets 25
2.5 Step 3—Estimation and Inference for Multiply Imputed Data Sets 26
2.5.1 Multiple Imputation—Estimators and Variances for Descriptive Statistics and Model Parameters 26
2.5.2 Multiple Imputation—Confidence Intervals 27

2.6 MI Procedures for Multivariate Inference 28
2.6.1 Multiple Parameter Hypothesis Tests 28
2.6.2 Tests of Linear Hypotheses 29
2.7 How Many Multiple Imputation Repetitions Are Needed? 30
2.8 Summary 30
Chapter 3: Preparation for Multiple Imputation 31
3.1 Planning the Imputation Session 31
3.2 Choosing the Variables to Include in a Multiple Imputation 31
3.3 Amount and Pattern of Missing Data 34
3.4 Types of Variables to Be Imputed 36
3.5 Imputation Methods 39
3.6 Number of Imputations (MI Repetitions) 39
3.7 Overview of Multiple Imputation Procedures 40
3.8 Multiple Imputation Example 41
3.9 Summary 48
Chapter 4: Multiple Imputation for the Analyzsis of Complex Sample Survey Data 49
4.1 Multiple Imputation and Informative Data Collection Designs 49
4.2 Complex Sample Surveys 50
4.3 Incorporating the Complex Sample Design in the MI Imputation Step 51
4.4 Incorporating the Complex Sample Design in the MI Analysis and Inference Steps 53
4.5 MI Imputation and Analysis for Subpopulations of Complex Sample Design Data Sets 57
4.6 Summary 58
Chapter 5: Multiple Imputation of Continuous Variables 59
5.1 Introduction to Multiple Imputation of Continuous Variables 59
5.2 Imputation of Continuous Variables with Arbitrary Missing Data 60
5.3 Imputation of Continuous Variables with Mixed Covariates and a Monotone Missing Data Pattern Using the Regression and Predictive Mean Matching Methods 68
5.3.1 Imputation of Continuous Variables with Mixed Covariates and a Monotone Missing Data Pattern Using the Regression Method 68
5.3.2 Imputation of Continuous Variables with Mixed Covariates and a Monotone Missing Data Pattern Using the Predictive Mean Matching Method 80
5.4 Imputation of Continuous Variables with an Arbitrary Missing Data Pattern and Mixed Covariates Using the FCS Method 83
5.4.1 Imputation of Continuous Variables with an Arbitrary Missing Data Pattern and Mixed Covariates Using the FCS Method 83
5.5 Summary 89
Chapter 6: Multiple Imputation of Classification Variables 91
6.1 Introduction to Multiple Imputation of Classification Variables 91
6.2 Imputation of a Classification Variable with a Monotone Missing Data Pattern Using the Logistic Method 92
6.3 Imputation of Classification Variables with an Arbitrary Missing Data Pattern and Mixed Covariates Using the FCS Discriminant Function and the FCS Logistic Regression Method 97
6.4 Imputation of Classification Variables with an Arbitrary Missing Data Pattern and Mixed Covariates: A Comparison of the FCS and MCMC/Monotone Methods 103
6.4.1 Imputation of Classification Variables with Mixed Covariates and an Arbitrary Missing Data Pattern Using the FCS Method 103

6.4.2 Imputation of Classification Variables with Mixed Covariates and an Arbitrary Missing Data Pattern Using the MCMC/Monotone and Monotone Logistic Methods with a Multistep Approach 107
6.5 Summary 111
Chapter 7: Multiple Imputation Case Studies 113
7.1 Multiple Imputation Case Studies 113
7.2 Comparative Analysis of HRS 2006 Data Using Complete Case Analysis and Multiple Imputation of Missing Data 113
7.2.1 Exploration of Missing Data 114
7.2.2 Complete Case Analysis Using PROC SURVEYLOGISTIC 115
7.2.3 Multiple Imputation of Missing Data with an Arbitrary Missing Data Pattern Using the FCS Method with Diagnostic Trace Plots 116
7.2.4 Logistic Regression Analysis of Imputed Data Sets Using PROC SURVEYLOGISTIC 117
7.2.5 Use of PROC MIANALYZE with Logistic Regression Output 118
7.2.6 Comparison of Complete Case Analysis and Multiply Imputed Analysis 119
7.3 Imputation and Analysis of Longitudinal Seizure Data 120
7.3.1 Introduction to the Seizure Data 120
7.3.2 Exploratory Analysis of Seizure Data 120
7.3.3 Conversion of Multiple-Record to Single-Record Data 121
7.3.4 Multiple Imputation of Missing Data 123
7.3.5 Conversion Back to Multiple Record Data for Analysis of Imputed Data Sets 125
7.3.6 Regression Analysis of Imputed Data Sets 126
7.4 Summary 128
Chapter 8: Preparation of Data Sets for PROC MIANALYZE 129
8.1 Preparation of Data Sets for Use in PROC MIANALYZE 129
8.2 Imputation of Major League Baseball Players' Salaries 130
8.3.1 PROC GLM Output Data Set for Use in PROC MIANALYZE 130
8.3.2 PROC MIXED Output Data Set for Use in PROC MIANALYZE 133
8.4 Imputation of NCS-R Data 135
8.5 PROC SURVEYPHREG Output Data Set for Use in PROC MIANALYZE 137
8.6 Summary 138
References 139
Index 143

From *Multiple Imputation of Missing Data Using SAS®* by Patricia Berglund and Steven Heeringa. Copyright © 2014, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.