

TABLE OF CONTENTS – Applications of Regression Models in Epidemiology

Preface

Acknowledgments

About the Authors

1 Basic Concepts for Statistical Modeling 1

1.1 Introduction 1

1.2 Parameter Versus Statistic 2

1.3 Probability Definition 3

1.4 Conditional Probability 3

1.5 Concepts of Prevalence and Incidence 4

1.6 Random Variables 4

1.7 Probability Distributions 4

1.8 Centrality and Dispersion Parameters of a Random Variable 6

1.9 Independence and Dependence of Random Variables 7

1.10 Special Probability Distributions 7

1.11 Hypothesis Testing 11

1.12 Confidence Intervals 14

1.13 Clinical Significance Versus Statistical Significance 14

1.14 Data Management 15

1.15 Concept of Causality 21

References 22

2 Introduction to Simple Linear Regression Models 25

2.1 Introduction 25

2.2 Specific Objectives 26

2.3 Model Definition 26

2.4 Model Assumptions 28

2.5 Graphic Representation 29

2.6 Geometry of the Simple Regression Model 29

2.7 Estimation of Parameters 30

2.8 Variance of Estimators 31

2.9 Hypothesis Testing About the Slope of the Regression Line 32

2.10 Coefficient of Determination R^2 34

2.11 Pearson Correlation Coefficient 34

2.12 Estimation of Regression Line Values and Prediction 35

2.13 Example 36

2.14 Predictions 39

2.15 Conclusions 46

Practice Exercise 47

References 48

3 Matrix Representation of the Linear Regression Model 49

3.1 Introduction 49

3.2 Specific Objectives 49

3.3 Definition 50

3.3.1 Matrix 50

3.4 Matrix Representation of a SLRM 50

3.5 Matrix Arithmetic 51

3.6 Matrix Multiplication	52
3.7 Special Matrices	53
3.8 Linear Dependence	54
3.9 Rank of a Matrix	54
3.10 Inverse Matrix $[A]$	54
3.11 Application of an Inverse Matrix in a SLRM	56
3.12 Estimation of β Parameters in a SLRM	56
3.13 Multiple Linear Regression Model (MLRM)	57
3.14 Interpretation of the Coefficients in a MLRM	58
3.15 ANOVA in a MLRM	58
3.16 Using Indicator Variables (Dummy Variables)	60
3.17 Polynomial Regression Models	63
3.18 Centering	64
3.19 Multicollinearity	65
3.20 Interaction Terms	65
3.21 Conclusion	66
Practice Exercise	66
References	67
4 Evaluation of Partial Tests of Hypotheses in a MLRM	69
4.1 Introduction	69
4.2 Specific Objectives	69
4.3 Definition of Partial Hypothesis	70
4.4 Evaluation Process of Partial Hypotheses	71
4.5 Special Cases	71
4.6 Examples	72

4.7 Conclusion 75

Practice Exercise 75

References 75

5 Selection of Variables in a Multiple Linear Regression Model 77

5.1 Introduction 77

5.2 Specific Objectives 77

5.3 Selection of Variables According to the Study Objectives 77

5.4 Criteria for Selecting the Best Regression Model 78

5.5 Stepwise Method in Regression 80

5.6 Limitations of Stepwise Methods 83

5.7 Conclusion 83

Practice Exercise 84

References 85

6 Correlation Analysis 87

6.1 Introduction 87

6.2 Specific Objectives 87

6.3 Main Correlation Coefficients Based on SLRM 87

6.4 Major Correlation Coefficients Based on MLRM 89

6.5 Partial Correlation Coefficient 90

6.6 Significance Tests 92

6.7 Suggested Correlations 92

6.8 Example 92

6.9 Conclusion 94

Practice Exercise 95

References 95

7 Strategies for Assessing the Adequacy of the Linear Regression Model 97

7.1 Introduction 97

7.2 Specific Objectives 98

7.3 Residual Definition 98

7.4 Initial Exploration 98

7.5 Initial Considerations 102

7.6 Standardized Residual 102

7.7 Jackknife Residuals (R-Student Residuals) 104

7.8 Normality of the Errors 105

7.9 Correlation of Errors 106

7.10 Criteria for Detecting Outliers, Leverage, and Influential Points 107

7.11 Leverage Values 108

7.12 Cook's Distance 108

7.13 COV RATIO 109

7.14 DFBETAS 110

7.15 DFFITS 110

7.16 Summary of the Results 111

7.17 Multicollinearity 111

7.18 Transformation of Variables 114

7.19 Conclusion 114

Practice Exercise 115

References 116

8 Weighted Least-Squares Linear Regression 117

8.1 Introduction 117

8.2 Specific Objectives 117

8.3 Regression Model with Transformation into the Original Scale of Y 117

8.4 Matrix Notation of the Weighted Linear Regression Model 119

8.5 Application of the WLS Model with Unequal Number of Subjects 120

8.6 Applications of the WLS Model When Variance Increases 123

8.7 Conclusions 125

Practice Exercise 126

References 127

9 Generalized Linear Models 129

9.1 Introduction 129

9.2 Specific Objectives 129

9.3 Exponential Family of Probability Distributions 130

9.4 Exponential Family of Probability Distributions with Dispersion 131

9.5 Mean and Variance in EF and EDF 132

9.6 Definition of a Generalized Linear Model 133

9.7 Estimation Methods 134

9.8 Deviance Calculation 135

9.9 Hypothesis Evaluation 136

9.10 Analysis of Residuals 138

9.11 Model Selection 139

9.12 Bayesian Models 139

9.13 Conclusions 140

References 140

10 Poisson Regression Models for Cohort Studies 141

10.1 Introduction 141

10.2 Specific Objectives 142

10.3 Incidence Measures	142
10.4 Confounding Variable	146
10.5 Stratified Analysis	147
10.6 Poisson Regression Model	148
10.7 Definition of Adjusted Relative Risk	149
10.8 Interaction Assessment	150
10.9 Relative Risk Estimation	151
10.10 Implementation of the Poisson Regression Model	152
10.11 Conclusion	161
Practice Exercise	162
References	162
11 Logistic Regression in Case–Control Studies	165
11.1 Introduction	165
11.2 Specific Objectives	166
11.3 Graphical Representation	166
11.4 Definition of the Odds Ratio	167
11.5 Confounding Assessment	168
11.6 Effect Modification	168
11.7 Stratified Analysis	169
11.8 Unconditional Logistic Regression Model	170
11.9 Types of Logistic Regression Models	171
11.10 Computing the OR _{crude}	173
11.11 Computing the Adjusted OR	173
11.12 Inference on OR	174
11.13 Example of the Application of ULR Model: Binomial Case	175

11.14 Conditional Logistic Regression Model 178

11.15 Conclusions 183

Practice Exercise 183

References 188

12 Regression Models in a Cross-Sectional Study 191

12.1 Introduction 191

12.2 Specific Objectives 192

12.3 Prevalence Estimation Using the Normal Approach 192

12.4 Definition of the Magnitude of the Association 198

12.5 POR Estimation 200

12.6 Prevalence Ratio 204

12.7 Stratified Analysis 204

12.8 Logistic Regression Model 207

12.9 Conclusions 210

Practice Exercise 210

References 211

13 Solutions to Practice Exercises 213

Chapter 2 Practice Exercise 213

Chapter 3 Practice Exercise 216

Chapter 4 Practice Exercise 220

Chapter 5 Practice Exercise 221

Chapter 6 Practice Exercise 223

Chapter 7 Practice Exercise 225

Chapter 8 Practice Exercise 228

Chapter 10 Practice Exercise 230

Chapter 11 Practice Exercise 233

Chapter 12 Practice Exercise 240

Index 245.