# Contents

Preface ix  
Acknowledgments xi  
Introduction xii  
Nomenclature xv  

1 A visual introduction to quantile regression 1  
Introduction 1  
1.1 The essential toolkit 1  
1.1.1 Unconditional mean, unconditional quantiles and surroundings 2  
1.1.2 Technical insight: Quantiles as solutions of a minimization problem 4  
1.1.3 Conditional mean, conditional quantiles and surroundings 6  
1.2 The simplest QR model: The case of the dummy regressor 8  
1.3 A slightly more complex QR model: The case of a nominal regressor 13  
1.4 A typical QR model: The case of a quantitative regressor 15  
1.5 Summary of key points 20  
References 21  

2 Quantile regression: Understanding how and why 22  
Introduction 22  
2.1 How and why quantile regression works 22  
2.1.1 The general linear programming problem 23  
2.1.2 The linear programming formulation for the QR problem 26  
2.1.3 Methods for solving the linear programming problem 31  
2.2 A set of illustrative artificial data 33  
2.2.1 Homogeneous error models 33  
2.2.2 Heterogeneous error models 35  
2.2.3 Dependent data error models 36  
2.3 How and why to work with QR 38  
2.3.1 QR for homogeneous and heterogeneous models 38  
2.3.2 QR prediction intervals 42  
2.3.3 A note on the quantile process 48  
2.4 Summary of key points 60  
References 62
3 Estimated coefficients and inference 64
   Introduction 64
   3.1 Empirical distribution of the quantile regression estimator 64
      3.1.1 The case of i.i.d. errors 66
      3.1.2 The case of i.n.i.d. errors 71
      3.1.3 The case of dependent errors 73
   3.2 Inference in QR, the i.i.d. case 76
   3.3 Wald, Lagrange multiplier, and likelihood ratio tests 84
   3.4 Summary of key points 92
   References 93

4 Additional tools for the interpretation and evaluation of the quantile regression model 94
   Introduction 94
   4.1 Data pre-processing 95
      4.1.1 Explanatory variable transformations 95
      4.1.2 Dependent variable transformations 97
   4.2 Response conditional density estimations 107
      4.2.1 The case of different scenario simulations 107
      4.2.2 The case of the response variable reconstruction 117
   4.3 Validation of the model 117
      4.3.1 Goodness of fit 117
      4.3.2 Resampling methods 120
   4.4 Summary of key points 128
   References 128

5 Models with dependent and with non-identically distributed data 131
   Introduction 131
   5.1 A closer look at the scale parameter, the independent and identically distributed case 131
      5.1.1 Estimating the variance of quantile regressions 131
      5.1.2 Confidence intervals and hypothesis testing on the estimated coefficients 134
      5.1.3 Example for the i.i.d. case 134
   5.2 The non-identically distributed case 137
      5.2.1 Example for the non-identically distributed case 141
      5.2.2 Quick ways to test equality of coefficients across quantiles in Stata 145
      5.2.3 The wage equation revisited 147
   5.3 The dependent data model 152
      5.3.1 Example with dependent data 155
   5.4 Summary of key points 158
   References 158

Appendix 5.A Heteroskedasticity tests and weighted quantile regression, Stata and R codes 159
CONTENTS vii

5.A.1 Koenker and Basset test for heteroskedasticity comparing two quantile regressions 159
5.A.2 Koenker and Basset test for heteroskedasticity comparing all quantile regressions 159
5.A.3 Quick tests for heteroskedasticity comparing quantile regressions 160
5.A.4 Compute the individual role of each explanatory variable to the dependent variable 161
5.A.5 R-codes for the Koenker and Basset test for heteroskedasticity 161

Appendix 5.B Dependent data 162

6 Additional models 163

Introduction 163
6.1 Nonparametric quantile regression 163
   6.1.1 Local polynomial regression 164
   6.1.2 Quantile smoothing splines 169
6.2 Nonlinear quantile regression 172
6.3 Censored quantile regression 175
6.4 Quantile regression with longitudinal data 183
6.5 Group effects through quantile regression 187
6.6 Binary quantile regression 195
6.7 Summary of key points 197

References 197

Appendix A Quantile regression and surroundings using R 201

Introduction 201
A.1 Loading data 202
   A.1.1 Text data 202
   A.1.2 Spreadsheet data 203
   A.1.3 Files from other statistical packages 204
A.2 Exploring data 205
   A.2.1 Graphical tools 205
   A.2.2 Summary statistics 209
A.3 Modeling data 211
   A.3.1 Ordinary least squares regression analysis 211
   A.3.2 Quantile regression analysis 212
A.4 Exporting figures and tables 217
   A.4.1 Exporting figures 217
   A.4.2 Exporting tables 218

References 218

Appendix B Quantile regression and surroundings using SAS 220

Introduction 220
B.1 Loading data 221
CONTENTS

B.1.1 Text data 221
B.1.2 Spreadsheet data 222
B.1.3 Files from other statistical packages 222
B.2 Exploring data 223
B.2.1 Graphical tools 223
B.2.2 Summary statistics 227
B.3 Modeling data 229
B.3.1 Ordinary least squares regression analysis 229
B.3.2 Quantile regression analysis 233
B.4 Exporting figures and tables 239
References 241

Appendix C Quantile regression and surroundings using Stata 242
Introduction 242
C.1 Loading data 243
C.1.1 Text data 243
C.1.2 Spreadsheet data 244
C.1.3 Files from other statistical packages 245
C.2 Exploring data 245
C.2.1 Graphical tools 245
C.2.2 Summary statistics 248
C.3 Modeling data 249
C.3.1 Ordinary least squares regression analysis 249
C.3.2 Quantile regression analysis 251
C.4 Exporting figures and tables 255
C.4.1 Exporting figures 255
C.4.2 Exporting tables 255
References 256

Index 257