Contents

List of examples xiii

1 Introduction 1

1.1 Strange observations 1
1.2 Residual plots 3
1.3 Leverage 7
1.4 Computing 11
1.5 Further reading 11

2 The algebra of deletion 13

2.1 Least squares 13
2.2 Deletion of observations 19
2.2.1 General 19
2.2.2 The parameter estimate $\hat{\beta}_{(I)}$ 20
2.2.3 Effect on the residual sum of squares 21
2.2.4 Deletion of one observation 21

3 Diagnostic quantities 22

3.1 Deletion residuals 22
3.2 Cook's statistic 24

4 The presentation of results 29

4.1 Lists, tests and plots 29
4.2 Half-normal plots with envelopes 34
4.3 Two examples 48
4.4 Simulation 60

5 Diagnostic plots for explanatory variables 62

5.1 Residual plots 62
5.2 Added variable plots 67
5.3 Partial regression leverage plots 72
5.4 Partial residual plots 73
5.5 Deletion of observations and addition of carriers 78

6 Transformations and constructed variables **80**

6.1 The need for transformations 80
6.2 A parametric family of power transformations 85
6.3 Asymptotically equivalent tests of hypotheses 92
6.4 Score tests and constructed variables 97
6.5 Examples 105

7 Transformations for percentages and proportions **137**

7.1 A general class of transformations 137
7.2 The folded power transformation 138
7.3 The Guerrero and Johnson transformation 144
7.4 The Aranda-Ordaz transformation 148
7.5 Discussion 151

8 Other constructed variables **155**

8.1 Constructed variables for power transformations 155
8.2 Quick estimates of the transformation parameter 173
8.3 Transformation of the explanatory variables 177
8.4 Simultaneous transformation of all variables 180

9 The power transformation with shifted location **184**

9.1 Constructed variables for a shift in location 184
9.2 Detailed structure of plots of constructed variables 192
9.3 Simultaneous estimation of shift in location and of the power transformation 195
9.4 Several shift parameters 208

10 Further diagnostic quantities **220**

10.1 Multiple deletion diagnostics 220
10.2 Other measures of influence 223
10.2.1 A subset of parameters 223
10.2.2 Confidence regions 226
10.3 Extensions and generalizations 228
10.3.1 Nonlinear least squares 228
10.3.2 Generalized linear models 231

11 Goodness of link tests **234**

11.1 Constructed variables for tests of the link in a generalized linear model 234
11.2 A Monte-Carlo test for separate families of hypotheses 240

12 Further problems, other analyses **246**

12.1 Robust transformations 246
12.2 Constructed variables for constancy of variance 249
12.3 Transformations and leverage 252

12.4 Other analyses 264
12.5 Beyond diagnostic analyses 268

Appendix 272

References 273

Author Index 279

Subject Index 281