

# CONTENTS

---

- 1. Simple linear regression** **1**
  - 1.1 Building a simple regression model, 4
  - 1.2 Least squares estimation, 7
  - 1.3 Estimating  $\sigma^2$ , 11
  - 1.4 Properties of least squares estimates, 13
  - 1.5 Comparing models: the analysis of variance, 14
  - 1.6 The coefficient of determination,  $R^2$ , 18
  - 1.7 Confidence intervals and tests, 19
  - 1.8 The residuals, 23
    - Problems, 26
  
- 2. Multiple regression** **31**
  - 2.1 Adding a single independent variable to a simple regression model, 35
  - 2.2 Regression in matrix notation, 40
  - 2.3 The analysis of variance, 47
  - 2.4 Regression through the origin, 51
    - Problems, 52
  
- 3. Drawing conclusions** **59**
  - 3.1 Interpreting parameter estimates, 59
  - 3.2 Sampling models, 62

## x Contents

3.3	Independent variables measured with error, 67 Problems, 72	
4.	<b>Weighted least squares, testing for lack of fit, general <math>F</math>-tests, and confidence ellipsoids</b>	<b>73</b>
4.1	Generalized and weighted least squares, 73	
4.2	Testing for lack of fit, variance known, 81	
4.3	Testing for lack of fit, variance unknown, 83	
4.4	General $F$ testing, 88	
4.5	Joint confidence regions, 89 Problems, 91	
5.	<b>Case analysis I: residuals and influence</b>	<b>97</b>
5.1	The residuals, 100	
5.2	The influence function, 106	
5.3	Outliers, 113 Problems, 117	
6.	<b>Case analysis II: symptoms and remedies</b>	<b>119</b>
6.1	Diagnostics: plotting residuals, 120	
6.2	Heterogeneity of variance, 122	
6.3	Nonlinearity, 126	
6.4	Distributional assumptions, 132	
6.5	Outliers and extreme cases, 135	
6.6	Choosing a transformation, 136 Problems, 144	
7.	<b>Model building I: defining new variables</b>	<b>150</b>
7.1	Polynomial regression, 150	
7.2	Dummy variables: dichotomous, 152	
7.3	Dummy variables: polytomous, 160	

7.4	Comparing regression lines, 162	
7.5	Scaling variables, 168	
7.6	Linear transformations and principal components, 169 Problems, 172	
<b>8.</b>	<b>Model building II: collinearity and variable selection</b>	<b>174</b>
8.1	Collinearity, 174	
8.2	Assumptions and notation, 183	
8.3	Selecting subsets on substantive grounds, 184	
8.4	Finding subsets, 185	
8.5	Computational methods I: stepwise methods, 190	
8.6	Computational methods II: all possible regressions, 196 Problems, 199	
<b>9.</b>	<b>Prediction</b>	<b>203</b>
9.1	Making predictions, 205	
9.2	Interval estimates, 213	
9.3	Interpolation versus extrapolation, 215 Problems, 218	
<b>10.</b>	<b>Incomplete data</b>	<b>221</b>
10.1	Missing at random, 221	
10.2	Handling incomplete data by filling in or deleting, 224	
10.3	Maximum likelihood estimates assuming normality, 227	
10.4	Missing observation correlation, 227	
10.5	General recommendations, 228	
<b>11.</b>	<b>Nonleast squares estimation</b>	<b>230</b>
11.1	Ridge regression, 232	
11.2	Generalized ridge regression, 234	

**xii Contents**

- 11.3** Regression on principal components, 235
- 11.4** James-Stein estimators, 235
- 11.5** Summary of shrunken estimators, 236
- 11.6** Robust regression, 237

**Appendix 239**

- 1A.1** A formal development of the simple regression model, 239
- 1A.2** Means and variances of random variables, 240
- 1A.3** Least squares, 242
- 1A.4** Means and variances of least squares estimates, 242
- 1A.5** An example of round-off error, 244
- 2A.1** A brief introduction to matrices, 244
- 2A.2** Random vectors, 251
- 2A.3** Least squares, 253
- 5A.1** Relating regression equations, 255
- 8A.1** Derivation of  $C_p$ , 256
- 9A.1** Finding a minimum covering ellipsoid, 257

**Tables 259****References 269****Symbol Index 277****Index 279**