

Contents

Data Tables and Associated Numerical Examples	xxi
PART I INTRODUCTION AND BASIC THEORY	
1 Introduction to Linear Models	3
1.1 Introduction and Motivation	3
1.2 Definition of the Linear Model	5
1.3 The Cell Means Model	12
1.4 Regression Models	14
1.4.1 The Single-Variable Regression Model	14
1.4.2 Regression Models with Several Inputs	19
1.5 Analysis of Variance Models	22
1.5.1 Fixed Effects Models	23
1.5.2 Constraints on the Cell Means	29
1.5.3 Mixed Models	31
1.6 Additional Topics	33
1.6.1 Designed Experiments and Observational Studies	33
1.6.2 Discrete Response Variables	34
1.6.3 Multivariate Linear Models	34
1.7 Concluding Comments	35
Exercises	36
2 The Distribution of Linear and Quadratic Forms	39
2.1 Linear and Quadratic Forms	39
2.2 Properties of the Multivariate Normal Distribution	42
2.3 The Distribution of Quadratic Forms	45
2.3.1 The Noncentral Chi-Squared Distribution	45
2.3.2 Fundamental Theorems	49
	xi

2.4	Other Noncentral Distributions	57
2.4.1	The Noncentral t Distribution	57
2.4.2	The Noncentral F Distribution	58
2.4.3	Extensions	59
	Exercises	60
3	Estimation and Inference in Simple Linear Models	65
3.1	Estimation of Parameters	65
3.1.1	Estimation for the Unconstrained Model	66
3.1.2	Estimation for the Constrained Model	70
3.1.3	Best Linear Unbiased Estimation (BLUE) of θ	77
3.1.4	A Partitioned Form of the Model	78
3.1.5	Reparameterized Models	80
3.2	Tests of Linear Hypotheses on θ	82
3.2.1	The Unconstrained Model	81
3.2.2	The Constrained Model	93
3.3	Confidence Regions and Intervals	95
3.3.1	Confidence Regions	95
3.3.2	Confidence Intervals	96
	Exercises	98
4	Simultaneous Inference: Tests and Confidence Intervals	103
4.1	Simultaneous Tests	103
4.1.1	Simultaneous Tests: General Methods	104
4.1.2	Simultaneous Tests: Cell Means Models	117
4.2	Simultaneous Confidence Intervals	124
	Exercises	126
PART II REGRESSION MODELS		
5	Regression on Functions of One Variable	131
5.1	The Simple Linear Regression Model	131
5.2	Estimation, Inference, and Prediction	136
5.2.1	Estimation	136
5.2.2	Hypothesis Testing and Confidence Intervals	141
5.2.3	Prediction and Prediction Intervals	147
5.3	Examining the Data and the Model	150
5.3.1	Residuals	151
5.3.2	Outliers, Extreme Points, and Influence	155

5.3.3	Normality, Independence, and Variance Homogeneity	161
5.4	A Test for Lack of Fit	165
5.5	Polynomial Regression Models	168
5.5.1	The Quadratic Model	169
5.5.2	Higher-Ordered Polynomial Models	174
5.5.3	Orthogonal Polynomials	176
5.5.4	Regression through the Origin	177
	Exercises	178
6	Transforming the Data and Miscellaneous Topics	183
6.1	The Need for Transformations	183
6.2	Weighted Least Squares	184
6.3	Variance Stabilizing Transformations	187
6.4	Transformations to Achieve a Linear Model	188
6.4.1	Transforming the Dependent Variable	189
6.4.2	Transforming the Predictors	194
6.5	Analysis of the Transformed Model	195
6.6	Transformations with Forbes Data	196
6.7	Other Topics	199
6.7.1	Random Inputs	200
6.7.2	Errors in the Inputs	201
6.7.3	Calibration	202
	Exercises	203
7	Regression on Functions of Several Variables	207
7.1	The Multiple Linear Regression Model	207
7.2	Preliminary Data Analysis	208
7.3	Fitting the Multiple Linear Regression Model	213
7.3.1	Fitting the Original Data	213
7.3.2	Transforming the Response	220
7.4	Partial Correlation and Added-Variable Plots	224
7.4.1	Partial Correlation	224
7.4.2	Added-Variable Plots	226
7.4.3	Simple and Partial Correlations	230
7.5	Variable Selection	232
7.5.1	The Case of Orthogonal Predictors	232
7.5.2	Nonorthogonal Predictors	236
7.5.3	Selection Methods	240
7.6	Model Specification	244
	Exercises	252

8	Collinearity in Multiple Linear Regression	257
8.1	The Collinearity Problem	257
8.1.1	Introduction	257
8.1.2	A Simple Example	258
8.1.3	The Picket Fence	262
8.1.4	Rotation of Coordinates	264
8.2	A Typical Example	265
8.2.1	Preliminary Data Analysis	266
8.2.2	Initial Regression Analysis	268
8.3	Collinearity Diagnostics	272
8.3.1	Variance Inflation Factors	272
8.3.2	Eigenvalues, Eigenvectors, and Principal Component Plots	276
8.4	Remedial Solutions: Biased Estimators	284
8.4.1	Variable Deletion	285
8.4.2	Regression on Principal Components	287
8.4.3	Ridge Regression	292
	Exercises	296
9	Influential Observations in Multiple Linear Regression	299
9.1	The Influential Data Problem	299
9.2	The Hat Matrix	301
9.2.1	The Centered and Uncentered Hat Matrices	301
9.2.2	Properties of the Hat Matrices	302
9.3	The Effects of Deleting Observations	306
9.4	Numerical Measures of Influence	310
9.4.1	The Diagonal Elements of the Hat Matrix	310
9.4.2	Residuals	310
9.4.3	The Mean Square Ratio	312
9.4.4	Cook's Distance	313
9.4.5	Other Indicators of Influential Data	314
9.5	The Dilemma Data	315
9.6	Principal Component Plots for Identifying Unusual Cases	320
9.6.1	The Projection Ellipse and Principal Component Plots	320
9.6.2	The Augmented Hat Matrix	323
9.6.3	Multiple Extremes: The Masking Problem	326
9.7	Robust/Resistant Methods in Regression Analysis	329
9.7.1	M-Estimation	329
9.7.2	Iterative, Reweighted Least Squares	330

9.7.3	Regression with Bounded Influence	332
	Exercises	333
10	Polynomial Models and Models with Qualitative Predictors	337
10.1	Polynomial Models	337
10.1.1	The Quadratic Model with Two Predictors	337
10.1.2	A Discussion of Quadratic Surfaces	339
10.2	The Analysis of Response Surfaces	343
10.2.1	Analysis with First-Order Models	343
10.2.2	Analysis with Second-Order Models	346
10.3	Models with Qualitative Predictors	348
10.3.1	Indicator Variables to Identify Groups of Data	349
10.3.2	Indicator Variables to Fit Segmented Polynomials	366
	Exercises	374
11	Related Topics	381
11.1	Introduction	381
11.2	Nonlinear Regression Models	382
11.2.1	Some Linearizeable Functions	383
11.2.2	The Modified Gauss-Newton Method	385
11.3	Nonparametric Model-Fitting Methods	388
11.3.1	Locally Weighted-Average Predictors	388
11.3.2	Projection Pursuit Regression	391
11.4	Logistic Regression	393
	Exercises	396
 PART III ANALYSIS OF VARIANCE MODELS		
12	Fixed Effects Models: I. Single-Factor Classification of Means	401
12.1	Concepts and Terminology	401
12.2	The One-Way Classification Model: Balanced Data	404
12.2.1	Parameter Estimation	405
12.2.2	The Hypothesis of Equal Means	406
12.2.3	Simultaneous Inferences about the Population Means	410
12.2.4	Simultaneous Acceptance and Confidence Ellipses	413
12.2.5	Orthogonal Contrasts	417
12.2.6	Reparameterizations of the One-Way Model	419

12.3	The One-Way Classification Model: Unbalanced Data	422
12.4	The Analysis of Covariance	426
	Exercises	428
13	Fixed Effects Models: II. Two-Way Cross-Classification	433
13.1	The Unconstrained Model: Balanced Data	433
13.1.1	Parameter Estimation	435
13.1.2	Tests of Hypotheses	436
13.1.3	Simultaneous Inference	445
13.1.4	Reparameterization of the Two-Factor Model	447
13.1.5	Experiments with One Observation per Cell: A Test for Interaction	453
13.2	The Unconstrained Model: Unbalanced Data	456
13.2.1	Discussion in Terms of the Cell Means Model	456
13.2.2	The Reparameterized Model	459
13.3	The No-Interaction Model: Balanced Data	467
13.3.1	Parameter Estimation	468
13.3.2	Tests of Hypotheses	469
13.3.3	Simultaneous Inference	472
13.3.4	Reparameterization in the No-Interaction Model	472
13.4	The No-Interaction Model: Unbalanced Data	474
13.4.1	Missing Cells: Estimation	474
13.4.2	Missing Cells: Testing Hypotheses	479
13.4.3	Connected Designs	482
13.5	Nonhomogeneous Experimental Units: The Concept of Blocking	483
13.5.1	A Model for the Randomized, Complete Block Design	484
13.5.2	Inferences on Parameters	485
	Exercises	488
14	Fixed Effects Models: III. Nested Factors and General Structure	495
14.1	The Two-Fold Nested Model	495
14.1.1	The Analysis with Balanced Data	495
14.1.2	The Analysis with Unbalanced Data	499
14.2	The Three-Factor Cross-Classified Model	502
14.3	A Three-Factor Nested-Factorial Model	507
14.3.1	The Analysis with Balanced Data	507
14.3.2	The Analysis with Unbalanced Data	510

14.4	A General Structure for Balanced, Fixed-Effects Models	511
14.4.1	The General Factorial Model	511
14.4.2	Nested-Factorial Models	515
	Exercises	519
15	Mixed Effects Models: I. The AOV with Balanced Data	525
15.1	Introduction	525
15.2	Examples of Model Formulation	527
15.3	A General Structure for Balanced Mixed Models	541
15.3.1	The Random Model	541
15.3.2	The Mixed Model	542
15.4	The Mixed Model Analysis	543
15.4.1	Estimation of Variance Components: The AOV Method	543
15.4.2	Estimation of the Fixed Effects Parameters	548
15.4.3	Properties of the Estimators and Inferential Results	549
15.5	Numerical Examples	555
15.6	Alternative Developments of Mixed Models	566
15.6.1	The Scheffé Mixed Model	570
15.6.2	A Randomization Theory	571
	Exercises	573
16	Mixed Effects Models: II. The AVE Method	577
16.1	Introduction	577
16.2	An Introduction to the AVE Method	578
16.2.1	The AVE Method for the Two-Factor Cross-Classification Model	578
16.2.2	Numerical Examples for the Two-Factor Model	586
16.3	The General AVE Method	593
16.3.1	The AVE Method: Factorial Models	593
16.3.2	The AVE Method: Nested and Nested-Factorial Models	596
16.3.3	The AVE Method: Computational Form of the VCE Table	598
16.3.4	The AVE Method: For General Mixed Effects Models	599
16.4	Numerical Examples	600
16.5	Properties of AVE Estimates	611

16.5.1	Diagnostic Analysis for the Two-Way Classification Model	612
16.5.2	Confidence Intervals	614
	Exercises	614

17 Mixed Effects Models: III. Unbalanced Data 619

17.1	Introduction	619
17.2	Parameter Estimation: Likelihood Methods	621
17.2.1	Maximum Likelihood Estimation (ML)	621
17.2.2	Restricted Maximum Likelihood Estimation (REML)	624
17.2.3	A Numerical Illustration of the Methods	628
17.3	ML and REML Estimates with Balanced Data	630
17.3.1	ML Estimation with Balanced Data	630
17.3.2	REML Estimation with Balanced Data	633
17.4	An EM Algorithm for REML Estimation	634
17.4.1	A Review of the EM Algorithm	634
17.4.2	The EM Algorithm Applied to REML Estimation	635
17.4.3	Estimation of Fixed Effects	641
17.4.4	Inferences on Variance Components and Fixed Effects	642
17.4.5	Numerical Examples that Illustrate the EM Algorithm	646
17.5	The EM Algorithm Applied to the AVE Method	649
17.5.1	The Computational Form of the EM-AVE Method	650
17.5.2	Numerical Examples of the EM-AVE Method	651
17.6	Models with Covariates	658
17.6.1	Development of the Analysis	659
17.6.2	A Numerical Example	660
17.7	Summary	663
	Exercises	663

Appendix A Mathematical Facts 667

A.I	Matrix Algebra	667
A.II	Optimization	683

Appendix B Statistical Facts 687

B.I	Estimation	687
B.II	Tests of Hypotheses and Confidence Regions	689

CONTENTS	xix
Appendix C Statistical Tables	691
Appendix D Data for Examples and Exercises	697
D.I Downloading Data Set Files from ftp Server	697
D.II Listing of Data Set Files	698
References	713
Index	727