

Contents

1	Introduction and Motivating Examples	1
1.1	Introduction	1
1.2	Model Formulation	1
1.3	Motivating Examples	5
1.3.1	Prostate Cancer	5
1.3.2	Outcome after Head Injury	9
1.3.3	Lung Cancer and Radon	10
1.3.4	Pharmacokinetic Data	11
1.3.5	Dental Growth	18
1.3.6	Spinal Bone Mineral Density	19
1.4	Nature of Randomness	21
1.5	Bayesian and Frequentist Inference	24
1.6	The Executive Summary	25
1.7	Bibliographic Notes	26

Part I Inferential Approaches

2	Frequentist Inference	29
2.1	Introduction	29
2.2	Frequentist Criteria	30
2.3	Estimating Functions	34
2.4	Likelihood	38
2.4.1	Maximum Likelihood Estimation	39
2.4.2	Variants on Likelihood	46
2.4.3	Model Misspecification	48
2.5	Quasi-Likelihood	51
2.5.1	Maximum Quasi-Likelihood Estimation	51
2.5.2	A More Complex Mean-Variance Model	55
2.6	Sandwich Estimation	58
2.7	Bootstrap Methods	65
2.7.1	The Bootstrap for a Univariate Parameter	66

2.7.2	The Bootstrap for Regression	68
2.7.3	Sandwich Estimation and the Bootstrap	69
2.8	Choice of Estimating Function	72
2.9	Hypothesis Testing	74
2.9.1	Motivation	74
2.9.2	Preliminaries	75
2.9.3	Score Tests	77
2.9.4	Wald Tests	77
2.9.5	Likelihood Ratio Tests	78
2.9.6	Quasi-Likelihood	79
2.9.7	Comparison of Test Statistics	79
2.10	Concluding Remarks	81
2.11	Bibliographic Notes	82
2.12	Exercises	83
3	Bayesian Inference	89
3.1	Introduction	89
3.2	The Posterior Distribution and its Summarization	90
3.3	Asymptotic Properties of Bayes Estimators	93
3.4	Prior Choice	94
3.4.1	Baseline Priors	94
3.4.2	Substantive Priors	97
3.4.3	Priors on Meaningful Scales	99
3.4.4	Frequentist Considerations	100
3.5	Model Misspecification	103
3.6	Bayesian Model Averaging	105
3.7	Implementation	106
3.7.1	Conjugacy	107
3.7.2	Laplace Approximation	110
3.7.3	Quadrature	112
3.7.4	Integrated Nested Laplace Approximations	114
3.7.5	Importance Sampling Monte Carlo	115
3.7.6	Direct Sampling using Conjugacy	116
3.7.7	Direct Sampling using the Rejection Algorithm	118
3.8	Markov Chain Monte Carlo	124
3.8.1	Markov Chains for Exploring Posterior Distributions	124
3.8.2	The Metropolis-Hastings Algorithm	126
3.8.3	The Metropolis Algorithm	127
3.8.4	The Gibbs Sampler	128
3.8.5	Combining Markov Kernels: Hybrid Schemes	129
3.8.6	Implementation Details	130
3.8.7	Implementation Summary	137
3.9	Exchangeability	138
3.10	Hypothesis Testing with Bayes Factors	142
3.11	Bayesian Inference Based on a Sampling Distribution	144

3.12	Concluding Remarks	147
3.13	Bibliographic Notes	149
3.14	Exercises	150
4	Hypothesis Testing and Variable Selection	157
4.1	Introduction	157
4.2	Frequentist Hypothesis Testing	157
4.2.1	Fisherian Approach	158
4.2.2	Neyman-Pearson Approach	158
4.2.3	Critique of the Fisherian Approach	158
4.2.4	Critique of the Neyman-Pearson Approach	159
4.3	Bayesian Hypothesis Testing with Bayes Factors	160
4.3.1	Overview of Approaches	160
4.3.2	Critique of the Bayes Factor Approach	162
4.3.3	A Bayesian View of Frequentist Hypothesis Testing	163
4.4	The Jeffreys-Lindley Paradox	165
4.5	Testing Multiple Hypotheses: General Considerations	169
4.6	Testing Multiple Hypotheses: Fixed Number of Tests	169
4.6.1	Frequentist Analysis	171
4.6.2	Bayesian Analysis	176
4.7	Testing Multiple Hypotheses: Variable Selection	182
4.8	Approaches to Variable Selection and Modeling	185
4.8.1	Stepwise Methods	187
4.8.2	All Possible Subsets	188
4.8.3	Bayesian Model Averaging	190
4.8.4	Shrinkage Methods	191
4.9	Model Building Uncertainty	191
4.10	A Pragmatic Compromise to Variable Selection	194
4.11	Concluding Comments	195
4.12	Bibliographic Notes	196
4.13	Exercises	196

Part II Independent Data

5	Linear Models	201
5.1	Introduction	201
5.2	Motivating Example: Prostate Cancer	201
5.3	Model Specification	202
5.4	A Justification for Linear Modeling	204
5.5	Parameter Interpretation	205
5.5.1	Causation versus Association	206
5.5.2	Multiple parameters	208
5.5.3	Data Transformations	211
5.6	Frequentist Inference	215
5.6.1	Likelihood	215

5.6.2	Least Squares Estimation	220
5.6.3	The Gauss-Markov Theorem	222
5.6.4	Sandwich Estimation	223
5.7	Bayesian Inference	225
5.8	Analysis of Variance	230
5.8.1	One-Way ANOVA	231
5.8.2	Crossed Designs	233
5.8.3	Nested Designs	236
5.8.4	Random and Mixed Effects Models	237
5.9	Bias-Variance Trade-Off	238
5.10	Robustness to Assumptions	243
5.10.1	Distribution of Errors	244
5.10.2	Non-Constant Variance	245
5.10.3	Correlated Errors	245
5.11	Assessment of Assumptions	245
5.11.1	Review of Assumptions	246
5.11.2	Residuals and Influence	247
5.11.3	Using the Residuals	249
5.12	Example: Prostate Cancer	251
5.13	Concluding Remarks	254
5.14	Bibliographic Notes	256
5.15	Exercises	256
6	General Regression Models	261
6.1	Introduction	261
6.2	Motivating Example: Pharmacokinetics of Theophylline	262
6.3	Generalized Linear Models	264
6.4	Parameter Interpretation	267
6.5	Likelihood Inference for GLMs	268
6.5.1	Estimation	268
6.5.2	Computation	271
6.5.3	Hypothesis Testing	275
6.6	Quasi-Likelihood Inference for GLMs	278
6.7	Sandwich Estimation for GLMs	279
6.8	Bayesian Inference for GLMs	280
6.8.1	Prior Specification	281
6.8.2	Computation	283
6.8.3	Hypothesis Testing	283
6.8.4	Overdispersed GLMs	283
6.9	Assessment of Assumptions for GLMs	286
6.10	Nonlinear Regression Models	290
6.11	Identifiability	292
6.12	Likelihood Inference for Nonlinear Models	293
6.12.1	Estimation	293
6.12.2	Hypothesis Testing	295

6.13	Least Squares Inference	297
6.14	Sandwich Estimation for Nonlinear Models	298
6.15	The Geometry of Least Squares	299
6.16	Bayesian Inference for Nonlinear Models	301
6.16.1	Prior Specification	301
6.16.2	Computation	303
6.16.3	Hypothesis Testing	303
6.17	Assessment of Assumptions for Nonlinear Models	306
6.18	Concluding Remarks	308
6.19	Bibliographic Notes	308
6.20	Exercises	308
7	Binary Data Models	313
7.1	Introduction	313
7.2	Motivating Examples	314
7.2.1	Outcome After Head Injury	314
7.2.2	Aircraft Fasteners	314
7.2.3	Bronchopulmonary Dysplasia	315
7.3	The Binomial Distribution	316
7.3.1	Genesis	316
7.3.2	Rare Events	317
7.4	Generalized Linear Models for Binary Data	319
7.4.1	Formulation	319
7.4.2	Link Functions	320
7.5	Overdispersion	321
7.6	Logistic Regression Models	324
7.6.1	Parameter Interpretation	324
7.6.2	Likelihood Inference for Logistic Regression Models	326
7.6.3	Quasi-Likelihood Inference for Logistic Regression Models	328
7.6.4	Bayesian Inference for Logistic Regression Models	329
7.7	Conditional Likelihood Inference	335
7.8	Assessment of Assumptions	339
7.9	Bias, Variance and Collapsibility	342
7.10	Case-Control Studies	345
7.10.1	The Epidemiological Context	345
7.10.2	Estimation for a Case-Control Study	346
7.10.3	Estimation for a Matched Case-Control Study	349
7.11	Concluding Remarks	351
7.12	Bibliographic Notes	352
7.13	Exercises	353

Part III Dependent Data

8	Linear Models	361
8.1	Introduction	361
8.2	Motivating Example: Dental Growth Curves	362
8.3	The Efficiency of Longitudinal Designs	364
8.4	Linear Mixed Models	367
8.4.1	The General Framework	367
8.4.2	Covariance Models for Clustered Data	368
8.4.3	Parameter Interpretation for Linear Mixed Models	371
8.5	Likelihood Inference for Linear Mixed Models	372
8.5.1	Inference for Fixed Effects	373
8.5.2	Inference for Variance Components via Maximum Likelihood	375
8.5.3	Inference for Variance Components via Restricted Maximum Likelihood	376
8.5.4	Inference for Random Effects	383
8.6	Bayesian Inference for Linear Mixed Models	388
8.6.1	A Three-Stage Hierarchical Model	388
8.6.2	Hyperpriors	389
8.6.3	Implementation	392
8.6.4	Extensions	396
8.7	Generalized Estimating Equations	398
8.7.1	Motivation	398
8.7.2	The GEE Algorithm	399
8.7.3	Estimation of Variance Parameters	402
8.8	Assessment of Assumptions	407
8.8.1	Review of Assumptions	407
8.8.2	Approaches to Assessment	409
8.9	Cohort and Longitudinal Effects	421
8.10	Concluding Remarks	424
8.11	Bibliographic Notes	425
8.12	Exercises	425
9	General Regression Models	433
9.1	Introduction	433
9.2	Motivating Examples	434
9.2.1	Contraception Data	434
9.2.2	Seizure Data	435
9.2.3	Pharmacokinetics of Theophylline	437
9.3	Generalized Linear Mixed Models	439
9.4	Likelihood Inference for Generalized Linear Mixed Models	440
9.5	Conditional Likelihood Inference for Generalized Linear Mixed Models	445
9.6	Bayesian Inference for Generalized Linear Mixed Models	448
9.6.1	Model Formulation	448
9.6.2	Hyperpriors	449

9.7	Generalized Linear Mixed Models with Spatial Dependence	453
9.7.1	A Markov Random Field Prior	453
9.7.2	Hyperpriors	454
9.8	Conjugate Random Effects Models	457
9.9	Generalized Estimating Equations for Generalized Linear Models	458
9.10	GEE2: Connected Estimating Equations	460
9.11	Interpretation of Marginal and Conditional Regression Coefficients	462
9.12	Introduction to Modeling Dependent Binary Data	465
9.13	Mixed Models for Binary Data	466
9.13.1	Generalized Linear Mixed Models for Binary Data	466
9.13.2	Likelihood Inference for the Binary Mixed Model	469
9.13.3	Bayesian Inference for the Binary Mixed Model	470
9.13.4	Conditional Likelihood Inference for Binary Mixed Models	473
9.14	Marginal Models for Dependent Binary Data	475
9.14.1	Generalized Estimating Equations	475
9.14.2	Loglinear Models	476
9.14.3	Further Multivariate Binary Models	478
9.15	Nonlinear Mixed Models	482
9.16	Parameterization of the Nonlinear Model	485
9.17	Likelihood Inference for the Nonlinear Mixed Effects Model	486
9.18	Bayesian Inference for the Nonlinear Mixed Model	490
9.18.1	Hyperpriors	491
9.18.2	Inference for Functions of Interest	491
9.19	Generalized Estimating Equations	495
9.20	Assessment of Assumptions for General Regression Models	496
9.21	Concluding Remarks	500
9.22	Bibliographic Notes	503
9.23	Exercises	504

Part IV Nonparametric Modeling

10	Preliminaries for Nonparametric Regression	511
10.1	Introduction	511
10.2	Motivating Examples	512
10.2.1	Light Detection and Ranging	513
10.2.2	Ethanol Data	514
10.3	The Optimal Prediction	514
10.3.1	Continuous Responses	515
10.3.2	Discrete Responses with K Categories	516
10.3.3	General Responses	518
10.3.4	In Practice	519
10.4	Measures of Predictive Accuracy	520
10.4.1	Continuous Responses	520
10.4.2	Discrete Responses with K Categories	524

10.4.3	General Responses	525
10.5	A First Look at Shrinkage Methods	525
10.5.1	Ridge Regression	526
10.5.2	The Lasso	531
10.6	Smoothing Parameter Selection	535
10.6.1	Mallows C_p	536
10.6.2	K -Fold Cross-Validation	538
10.6.3	Generalized Cross-Validation	541
10.6.4	AIC for General Models	543
10.6.5	Cross-Validation for Generalized Linear Models	547
10.7	Concluding Comments	550
10.8	Bibliographic Notes	551
10.9	Exercises	553
11	Spline and Kernel Methods	557
11.1	Introduction	557
11.2	Spline Methods	557
11.2.1	Piecewise Polynomials and Splines	557
11.2.2	Natural Cubic Splines	563
11.2.3	Cubic Smoothing Splines	564
11.2.4	B -Splines	567
11.2.5	Penalized Regression Splines	568
11.2.6	A Brief Spline Summary	571
11.2.7	Inference for Linear Smoothers	571
11.2.8	Linear Mixed Model Spline Representation: Likelihood Inference	574
11.2.9	Linear Mixed Model Spline Representation: Bayesian Inference	580
11.3	Kernel Methods	585
11.3.1	Kernels	586
11.3.2	Kernel Density Estimation	586
11.3.3	The Nadaraya-Watson Kernel Estimator	590
11.3.4	Local Polynomial Regression	592
11.4	Variance Estimation	595
11.5	Spline and Kernel Methods for Generalized Linear Models	598
11.5.1	Generalized Linear Models with Penalized Regression Splines	600
11.5.2	A Generalized Linear Mixed Model Spline Representation	603
11.5.3	Generalized Linear Models with Local Polynomials	604
11.6	Concluding Comments	605
11.7	Bibliographic Notes	605
11.8	Exercises	606

12 Nonparametric Regression with Multiple Predictors	609
12.1 Introduction	609
12.2 Generalized Additive Models	610
12.2.1 Model Formulation	610
12.2.2 Computation via Backfitting	611
12.3 Spline Methods with Multiple Predictors	612
12.3.1 Natural Thin Plate Splines	614
12.3.2 Thin Plate Regression Splines	615
12.3.3 Tensor Product Splines	616
12.4 Kernel Methods with Multiple Predictors	619
12.5 Smoothing Parameter Estimation	620
12.5.1 Conventional Approaches	620
12.5.2 Mixed Model Formulation	620
12.6 Varying-Coefficient Models	622
12.7 Regression Trees	625
12.7.1 Hierarchical Partitioning	625
12.7.2 Multiple Adaptive Regression Splines	635
12.8 Classification	637
12.8.1 Logistic Models with K Classes	638
12.8.2 Linear and Quadratic Discriminant Analysis	639
12.8.3 Kernel Density Estimation and Classification	644
12.8.4 Classification Trees	648
12.8.5 Bagging	649
12.8.6 Random Forests	653
12.9 Concluding Comments	657
12.10 Bibliographic Notes	658
12.11 Exercises	658
Part V Appendices	
A Differentiation of Matrix Expressions	663
B Matrix Results	667
C Some Linear Algebra	669
D Probability Distributions and Generating Functions	671
E Functions of Normal Random Variables	681
F Some Results from Classical Statistics	683
G Basic Large Sample Theory	687
References	689
Index	709