

---

## Contents

<i>Preface</i>	xv
<i>Acknowledgments</i>	xviii
<i>Notation</i>	xix
<b>Part I Classic Statistical Inference</b>	<b>1</b>
<b>1 Algorithms and Inference</b>	<b>3</b>
1.1 A Regression Example	4
1.2 Hypothesis Testing	8
1.3 Notes	11
<b>2 Frequentist Inference</b>	<b>12</b>
2.1 Frequentism in Practice	14
2.2 Frequentist Optimality	18
2.3 Notes and Details	20
<b>3 Bayesian Inference</b>	<b>22</b>
3.1 Two Examples	24
3.2 Uninformative Prior Distributions	28
3.3 Flaws in Frequentist Inference	30
3.4 A Bayesian/Frequentist Comparison List	33
3.5 Notes and Details	36
<b>4 Fisherian Inference and Maximum Likelihood Estimation</b>	<b>38</b>
4.1 Likelihood and Maximum Likelihood	38
4.2 Fisher Information and the MLE	41
4.3 Conditional Inference	45
4.4 Permutation and Randomization	49
4.5 Notes and Details	51
<b>5 Parametric Models and Exponential Families</b>	<b>53</b>

x	<i>Contents</i>	
5.1	Univariate Families	54
5.2	The Multivariate Normal Distribution	55
5.3	Fisher's Information Bound for Multiparameter Families	59
5.4	The Multinomial Distribution	61
5.5	Exponential Families	64
5.6	Notes and Details	69
	<b>Part II Early Computer-Age Methods</b>	<b>73</b>
<b>6</b>	<b>Empirical Bayes</b>	<b>75</b>
6.1	Robbins' Formula	75
6.2	The Missing-Species Problem	78
6.3	A Medical Example	84
6.4	Indirect Evidence 1	88
6.5	Notes and Details	88
<b>7</b>	<b>James–Stein Estimation and Ridge Regression</b>	<b>91</b>
7.1	The James–Stein Estimator	91
7.2	The Baseball Players	94
7.3	Ridge Regression	97
7.4	Indirect Evidence 2	102
7.5	Notes and Details	104
<b>8</b>	<b>Generalized Linear Models and Regression Trees</b>	<b>108</b>
8.1	Logistic Regression	109
8.2	Generalized Linear Models	116
8.3	Poisson Regression	120
8.4	Regression Trees	124
8.5	Notes and Details	128
<b>9</b>	<b>Survival Analysis and the EM Algorithm</b>	<b>131</b>
9.1	Life Tables and Hazard Rates	131
9.2	Censored Data and the Kaplan–Meier Estimate	134
9.3	The Log-Rank Test	139
9.4	The Proportional Hazards Model	143
9.5	Missing Data and the EM Algorithm	146
9.6	Notes and Details	150
<b>10</b>	<b>The Jackknife and the Bootstrap</b>	<b>155</b>
10.1	The Jackknife Estimate of Standard Error	156
10.2	The Nonparametric Bootstrap	159
10.3	Resampling Plans	162

<i>Contents</i>		xi
10.4	The Parametric Bootstrap	169
10.5	Influence Functions and Robust Estimation	174
10.6	Notes and Details	177
<b>11</b>	<b>Bootstrap Confidence Intervals</b>	181
11.1	Neyman's Construction for One-Parameter Problems	181
11.2	The Percentile Method	185
11.3	Bias-Corrected Confidence Intervals	190
11.4	Second-Order Accuracy	192
11.5	Bootstrap- $t$ Intervals	195
11.6	Objective Bayes Intervals and the Confidence Distribution	198
11.7	Notes and Details	204
<b>12</b>	<b>Cross-Validation and <math>C_p</math> Estimates of Prediction Error</b>	208
12.1	Prediction Rules	208
12.2	Cross-Validation	213
12.3	Covariance Penalties	218
12.4	Training, Validation, and Ephemeral Predictors	227
12.5	Notes and Details	230
<b>13</b>	<b>Objective Bayes Inference and MCMC</b>	233
13.1	Objective Prior Distributions	234
13.2	Conjugate Prior Distributions	237
13.3	Model Selection and the Bayesian Information Criterion	243
13.4	Gibbs Sampling and MCMC	251
13.5	Example: Modeling Population Admixture	256
13.6	Notes and Details	261
<b>14</b>	<b>Postwar Statistical Inference and Methodology</b>	264
<b>Part III Twenty-First-Century Topics</b>		269
<b>15</b>	<b>Large-Scale Hypothesis Testing and FDRs</b>	271
15.1	Large-Scale Testing	272
15.2	False-Discovery Rates	275
15.3	Empirical Bayes Large-Scale Testing	278
15.4	Local False-Discovery Rates	282
15.5	Choice of the Null Distribution	286
15.6	Relevance	290
15.7	Notes and Details	294
<b>16</b>	<b>Sparse Modeling and the Lasso</b>	298

xii	<i>Contents</i>	
16.1	Forward Stepwise Regression	299
16.2	The Lasso	303
16.3	Fitting Lasso Models	308
16.4	Least-Angle Regression	309
16.5	Fitting Generalized Lasso Models	313
16.6	Post-Selection Inference for the Lasso	317
16.7	Connections and Extensions	319
16.8	Notes and Details	321
<b>17</b>	<b>Random Forests and Boosting</b>	<b>324</b>
17.1	Random Forests	325
17.2	Boosting with Squared-Error Loss	333
17.3	Gradient Boosting	338
17.4	Adaboost: the Original Boosting Algorithm	341
17.5	Connections and Extensions	345
17.6	Notes and Details	347
<b>18</b>	<b>Neural Networks and Deep Learning</b>	<b>351</b>
18.1	Neural Networks and the Handwritten Digit Problem	353
18.2	Fitting a Neural Network	356
18.3	Autoencoders	362
18.4	Deep Learning	364
18.5	Learning a Deep Network	368
18.6	Notes and Details	371
<b>19</b>	<b>Support-Vector Machines and Kernel Methods</b>	<b>375</b>
19.1	Optimal Separating Hyperplane	376
19.2	Soft-Margin Classifier	378
19.3	SVM Criterion as Loss Plus Penalty	379
19.4	Computations and the Kernel Trick	381
19.5	Function Fitting Using Kernels	384
19.6	Example: String Kernels for Protein Classification	385
19.7	SVMs: Concluding Remarks	387
19.8	Kernel Smoothing and Local Regression	387
19.9	Notes and Details	390
<b>20</b>	<b>Inference After Model Selection</b>	<b>394</b>
20.1	Simultaneous Confidence Intervals	395
20.2	Accuracy After Model Selection	402
20.3	Selection Bias	408
20.4	Combined Bayes–Frequentist Estimation	412
20.5	Notes and Details	417

<i>Contents</i>		xiii
<b>21</b>	<b>Empirical Bayes Estimation Strategies</b>	421
21.1	Bayes Deconvolution	421
21.2	$g$ -Modeling and Estimation	424
21.3	Likelihood, Regularization, and Accuracy	427
21.4	Two Examples	432
21.5	Generalized Linear Mixed Models	437
21.6	Deconvolution and $f$ -Modeling	440
21.7	Notes and Details	444
	<i>Epilogue</i>	446
	<i>References</i>	453
	<i>Author Index</i>	463
	<i>Subject Index</i>	467