

Contents

<i>Preface</i>	xv
<i>Acknowledgments</i>	xviii
<i>Notation</i>	xix
Part I Classic Statistical Inference	1
1 Algorithms and Inference	3
1.1 A Regression Example	4
1.2 Hypothesis Testing	8
1.3 Notes	11
2 Frequentist Inference	12
2.1 Frequentism in Practice	14
2.2 Frequentist Optimality	18
2.3 Notes and Details	20
3 Bayesian Inference	22
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
4 Fisherian Inference and Maximum Likelihood Estimation	38
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
5 Parametric Models and Exponential Families	53

xii	<i>Contents</i>		<i>Contents</i>	xiii
16.1	Forward Stepwise Regression	299	21 Empirical Bayes Estimation Strategies	421
16.2	The Lasso	303	21.1 Bayes Deconvolution	421
16.3	Fitting Lasso Models	308	21.2 g -Modeling and Estimation	424
16.4	Least-Angle Regression	309	21.3 Likelihood, Regularization, and Accuracy	427
16.5	Fitting Generalized Lasso Models	313	21.4 Two Examples	432
16.6	Post-Selection Inference for the Lasso	317	21.5 Generalized Linear Mixed Models	437
16.7	Connections and Extensions	319	21.6 Deconvolution and f -Modeling	440
16.8	Notes and Details	321	21.7 Notes and Details	444
17	Random Forests and Boosting	324	<i>Epilogue</i>	446
17.1	Random Forests	325	<i>References</i>	453
17.2	Boosting with Squared-Error Loss	333	<i>Author Index</i>	463
17.3	Gradient Boosting	338	<i>Subject Index</i>	467
17.4	Adaboost: the Original Boosting Algorithm	341		
17.5	Connections and Extensions	345		
17.6	Notes and Details	347		
18	Neural Networks and Deep Learning	351		
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		
19	Support-Vector Machines and Kernel Methods	375		
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		
20	Inference After Model Selection	394		
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		