

Contents

	<i>Preface</i>	<i>page</i> xvii
	<i>List of Acronyms</i>	xx
1	Introduction	1
	1.1 Background	1
	1.2 Notation	1
	1.2.1 Probability Distributions	2
	1.2.2 Conditional Probability Distributions	2
	1.2.3 Expectations and Conditional Expectations	3
	1.2.4 Unified Notation	3
	1.2.5 General Random Variables	3
	1.3 Statistical Inference	4
	1.3.1 Statistical Model	5
	1.3.2 Some Generic Estimation Problems	6
	1.3.3 Some Generic Detection Problems	6
	1.4 Performance Analysis	7
	1.5 Statistical Decision Theory	7
	1.5.1 Conditional Risk and Optimal Decision Rules	8
	1.5.2 Bayesian Approach	9
	1.5.3 Minimax Approach	10
	1.5.4 Other Non-Bayesian Rules	11
	1.6 Derivation of Bayes Rule	12
	1.7 Link Between Minimax and Bayesian Decision Theory	14
	1.7.1 Dual Concept	14
	1.7.2 Game Theory	15
	1.7.3 Saddlepoint	15
	1.7.4 Randomized Decision Rules	16
	Exercises	18
	References	21
	Part I Hypothesis Testing	23
2	Binary Hypothesis Testing	25
	2.1 General Framework	25
	2.2 Bayesian Binary Hypothesis Testing	26

x	Contents	
	2.2.1 Likelihood Ratio Test	27
	2.2.2 Uniform Costs	28
	2.2.3 Examples	28
2.3	Binary Minimax Hypothesis Testing	32
	2.3.1 Equalizer Rules	33
	2.3.2 Bayes Risk Line and Minimum Risk Curve	34
	2.3.3 Differentiable $V(\pi_0)$	35
	2.3.4 Nondifferentiable $V(\pi_0)$	35
	2.3.5 Randomized LRTs	37
	2.3.6 Examples	38
2.4	Neyman–Pearson Hypothesis Testing	40
	2.4.1 Solution to the NP Optimization Problem	41
	2.4.2 NP Rule	42
	2.4.3 Receiver Operating Characteristic	43
	2.4.4 Examples	44
	2.4.5 Convex Optimization	46
	Exercises	47
3	Multiple Hypothesis Testing	54
	3.1 General Framework	54
	3.2 Bayesian Hypothesis Testing	55
	3.2.1 Optimal Decision Regions	56
	3.2.2 Gaussian Ternary Hypothesis Testing	58
	3.3 Minimax Hypothesis Testing	58
	3.4 Generalized Neyman–Pearson Detection	62
	3.5 Multiple Binary Tests	62
	3.5.1 Bonferroni Correction	63
	3.5.2 False Discovery Rate	64
	3.5.3 Benjamini–Hochberg Procedure	65
	3.5.4 Connection to Bayesian Decision Theory	66
	Exercises	67
	References	70
4	Composite Hypothesis Testing	71
	4.1 Introduction	71
	4.2 Random Parameter Θ	72
	4.2.1 Uniform Costs Over Each Hypothesis	73
	4.2.2 Nonuniform Costs Over Hypotheses	76
	4.3 Uniformly Most Powerful Test	77
	4.3.1 Examples	77
	4.3.2 Monotone Likelihood Ratio Theorem	79
	4.3.3 Both Composite Hypotheses	80
	4.4 Locally Most Powerful Test	82
	4.5 Generalized Likelihood Ratio Test	84

	4.5.1 GLRT for Gaussian Hypothesis Testing	84
	4.5.2 GLRT for Cauchy Hypothesis Testing	86
4.6	Random versus Nonrandom θ	87
4.7	Non-Dominated Tests	88
4.8	Composite m -ary Hypothesis Testing	90
	4.8.1 Random Parameter Θ	90
	4.8.2 Non-Dominated Tests	91
	4.8.3 m -GLRT	92
4.9	Robust Hypothesis Testing	92
	4.9.1 Robust Detection with Conditionally Independent Observations	96
	4.9.2 Epsilon-Contamination Class	97
	Exercises	99
	References	103
5	Signal Detection	105
	5.1 Introduction	105
	5.2 Problem Formulation	106
	5.3 Detection of Known Signal in Independent Noise	107
	5.3.1 Signal in i.i.d. Gaussian Noise	107
	5.3.2 Signal in i.i.d. Laplacian Noise	108
	5.3.3 Signal in i.i.d. Cauchy Noise	110
	5.3.4 Approximate NP Test	111
	5.4 Detection of Known Signal in Correlated Gaussian Noise	112
	5.4.1 Reduction to i.i.d. Noise Case	113
	5.4.2 Performance Analysis	114
	5.5 m -ary Signal Detection	115
	5.5.1 Bayes Classification Rule	116
	5.5.2 Performance Analysis	116
	5.6 Signal Selection	117
	5.6.1 i.i.d. Noise	118
	5.6.2 Correlated Noise	118
	5.7 Detection of Gaussian Signals in Gaussian Noise	120
	5.7.1 Detection of a Gaussian Signal in White Gaussian Noise	121
	5.7.2 Detection of i.i.d. Zero-Mean Gaussian Signal	122
	5.7.3 Diagonalization of Signal Covariance	123
	5.7.4 Performance Analysis	125
	5.7.5 Gaussian Signals With Nonzero Mean	126
	5.8 Detection of Weak Signals	127
	5.9 Detection of Signal with Unknown Parameters in White Gaussian Noise	128
	5.9.1 General Approach	129
	5.9.2 Linear Gaussian Model	130
	5.9.3 Nonlinear Gaussian Model	130
	5.9.4 Discrete Parameter Set	132
	5.10 Deflection-Based Detection of Non-Gaussian Signal in Gaussian Noise	135

	Exercises	139
	References	143
6	Convex Statistical Distances	145
	6.1 Kullback–Leibler Divergence	145
	6.2 Entropy and Mutual Information	147
	6.3 Chernoff Divergence, Chernoff Information, and Bhattacharyya Distance	149
	6.4 Ali–Silvey Distances	151
	6.5 Some Useful Inequalities	155
	Exercises	156
	References	158
7	Performance Bounds for Hypothesis Testing	160
	7.1 Simple Lower Bounds on Conditional Error Probabilities	160
	7.2 Simple Lower Bounds on Error Probability	162
	7.3 Chernoff Bound	163
	7.3.1 Moment-Generating and Cumulant-Generating Functions	163
	7.3.2 Chernoff Bound	164
	7.4 Application of Chernoff Bound to Binary Hypothesis Testing	167
	7.4.1 Exponential Upper Bounds on P_F and P_M	168
	7.4.2 Bayesian Error Probability	170
	7.4.3 Lower Bound on ROC	172
	7.4.4 Example	172
	7.5 Bounds on Classification Error Probability	173
	7.5.1 Upper and Lower Bounds in Terms of Pairwise Error Probabilities	173
	7.5.2 Bonferroni’s Inequalities	176
	7.5.3 Generalized Fano’s Inequality	176
	7.6 Appendix: Proof of Theorem 7.4	178
	Exercises	181
	References	183
8	Large Deviations and Error Exponents for Hypothesis Testing	184
	8.1 Introduction	184
	8.2 Chernoff Bound for Sum of i.i.d. Random Variables	185
	8.2.1 Cramér’s Theorem	185
	8.2.2 Why is the Central Limit Theorem Inapplicable Here?	186
	8.3 Hypothesis Testing with i.i.d. Observations	187
	8.3.1 Bayesian Hypothesis Testing with i.i.d. Observations	188
	8.3.2 Neyman–Pearson Hypothesis Testing with i.i.d. Observations	189
	8.3.3 Hoeffding Problem	189
	8.3.4 Example	191
	8.4 Refined Large Deviations	194
	8.4.1 The Method of Exponential Tilting	194

	8.4.2 Sum of i.i.d. Random Variables	195
	8.4.3 Lower Bounds on Large-Deviations Probabilities	198
	8.4.4 Refined Asymptotics for Binary Hypothesis Testing	199
	8.4.5 Non-i.i.d. Components	200
	8.5 Appendix: Proof of Lemma 8.1	202
	Exercises	203
	References	206
9	Sequential and Quickest Change Detection	208
	9.1 Sequential Detection	208
	9.1.1 Problem Formulation	208
	9.1.2 Stopping Times and Decision Rules	209
	9.1.3 Two Formulations of the Sequential Hypothesis Testing Problem	209
	9.1.4 Sequential Probability Ratio Test	210
	9.1.5 SPRT Performance Evaluation	212
	9.2 Quickest Change Detection	217
	9.2.1 Minimax Quickest Change Detection	219
	9.2.2 Bayesian Quickest Change Detection	223
	Exercises	227
	References	229
10	Detection of Random Processes	231
	10.1 Discrete-Time Random Processes	231
	10.1.1 Periodic Stationary Gaussian Processes	232
	10.1.2 Stationary Gaussian Processes	234
	10.1.3 Markov Processes	235
	10.2 Continuous-Time Processes	238
	10.2.1 Covariance Kernel	239
	10.2.2 Karhunen–Loève Transform	240
	10.2.3 Detection of Known Signals in Gaussian Noise	244
	10.2.4 Detection of Gaussian Signals in Gaussian Noise	246
	10.3 Poisson Processes	248
	10.4 General Processes	250
	10.4.1 Likelihood Ratio	250
	10.4.2 Ali–Silvey Distances	252
	10.5 Appendix: Proof of Proposition 10.1	253
	Exercises	254
	References	256
Part II Estimation		257
11	Bayesian Parameter Estimation	259
	11.1 Introduction	259

11.2	Bayesian Parameter Estimation	259
11.3	MMSE Estimation	260
11.4	MMAE Estimation	262
11.5	MAP Estimation	263
11.6	Parameter Estimation for Linear Gaussian Models	265
11.7	Estimation of Vector Parameters	266
11.7.1	Vector MMSE Estimation	267
11.7.2	Vector MMAE Estimation	267
11.7.3	Vector MAP Estimation	267
11.7.4	Linear MMSE Estimation	268
11.7.5	Vector Parameter Estimation in Linear Gaussian Models	269
11.7.6	Other Cost Functions for Bayesian Estimation	270
11.8	Exponential Families	270
11.8.1	Basic Properties	271
11.8.2	Conjugate Priors	273
	Exercises	276
	References	279
12	Minimum Variance Unbiased Estimation	280
12.1	Nonrandom Parameter Estimation	280
12.2	Sufficient Statistics	281
12.3	Factorization Theorem	283
12.4	Rao–Blackwell Theorem	284
12.5	Complete Families of Distributions	286
12.5.1	Link Between Completeness and Sufficiency	288
12.5.2	Link Between Completeness and MVUE	289
12.5.3	Link Between Completeness and Exponential Families	289
12.6	Discussion	291
12.7	Examples: Gaussian Families	291
	Exercises	294
	References	296
13	Information Inequality and Cramér–Rao Lower Bound	297
13.1	Fisher Information and the Information Inequality	297
13.2	Cramér–Rao Lower Bound	300
13.3	Properties of Fisher Information	302
13.4	Conditions for Equality in Information Inequality	305
13.5	Vector Parameters	306
13.6	Information Inequality for Random Parameters	311
13.7	Biased Estimators	312
13.8	Appendix: Derivation of (13.16)	314
	Exercises	315
	References	318

14	Maximum Likelihood Estimation	319
	14.1 Introduction	319
	14.2 Computation of ML Estimates	320
	14.3 Invariance to Reparameterization	322
	14.4 MLE in Exponential Families	323
	14.4.1 Mean-Value Parameterization	324
	14.4.2 Relation to MVUEs	324
	14.4.3 Asymptotics	325
	14.5 Estimation of Parameters on Boundary	327
	14.6 Asymptotic Properties for General Families	329
	14.6.1 Consistency	329
	14.6.2 Asymptotic Efficiency and Normality	331
	14.7 Nonregular ML Estimation Problems	334
	14.8 Nonexistence of MLE	335
	14.9 Non-i.i.d. Observations	338
	14.10 M-Estimators and Least-Squares Estimators	338
	14.11 Expectation-Maximization (EM) Algorithm	339
	14.11.1 General Structure of the EM Algorithm	340
	14.11.2 Convergence of EM Algorithm	341
	14.11.3 Examples	341
	14.12 Recursive Estimation	347
	14.12.1 Recursive MLE	347
	14.12.2 Recursive Approximations to Least-Squares Solution	349
	14.13 Appendix: Proof of Theorem 14.2	350
	14.14 Appendix: Proof of Theorem 14.4	351
	Exercises	352
	References	356
15	Signal Estimation	358
	15.1 Linear Innovations	358
	15.2 Discrete-Time Kalman Filter	360
	15.2.1 Time-Invariant Case	365
	15.3 Extended Kalman Filter	367
	15.4 Nonlinear Filtering for General Hidden Markov Models	369
	15.5 Estimation in Finite Alphabet Hidden Markov Models	372
	15.5.1 Viterbi Algorithm	373
	15.5.2 Forward-Backward Algorithm	375
	15.5.3 Baum–Welch Algorithm for HMM Learning	378
	Exercises	381
	References	383
	Appendix A Matrix Analysis	384
	Appendix B Random Vectors and Covariance Matrices	390

Appendix C	Probability Distributions	391
Appendix D	Convergence of Random Sequences	393
	<i>Index</i>	395