

Contents

PART I	BASIC THEORY	1
1	Heuristics	3
1	Notation and truth	3
2	Limit theory	5
2.1	Local concentration (consistency)	6
2.2	Rate of convergence	7
2.3	Limiting distribution	8
3	Efficiency heuristics	10
4	The concept of efficiency unedited	12
2	Consistency	17
1	Comparison arguments	17
1.1	Another comparison argument	20
2	Consistency: Wald's method	21
3	Uniform approximation	24
4	Global difficulties	24
5	High-level conditions	27
6	Z-estimators	30
7	A nonparametric example	30
3	Rates of convergence and limit distributions	39
1	Preliminaries on stochastic order notation	39
2	Approximation of minimizing values	41
3	Comparisons	44
4	The classical (local) regularity conditions	48
5	Estimation of a mode	50
6	Censored LAD regression needs editing	53
7	Trouble at the boundary	54
8	Likelihood ratio tests unedited junk	55
9	Strange behaviour at the boundary needs editing	55
4	Classical regularity conditions	63
1	Comparison arguments	63
2	The classical (local) regularity conditions	65
3	Asymptotics via quadratic approximation	69
4	Quadratic approximation on a grid: Le Cam's method	74

5	Likelihood ratio tests	75
6	Tangent approximations	77
7	Behaviour under alternatives	81
8	Strange behaviour at the boundary	83
5	Convexity	89
1	A convexity simplification	89
2	An approximation lemma	92
3	A constrained minimization problem	95
4	Iterative proportional fitting	96
6	Categorical data	101
1	Introduction	101
2	Goodness of fit tests	101
3	Asymptotics	105
4	Inefficient	108
5	The conditioned Poisson model	109
7	Contiguity	119
1	Definition and equivalences	119
2	Contiguity for product measures	125
3	Limit distributions under contiguous alternatives	131
8	Total variation distance between measures	137
1	Why bother with different distances?	137
2	Total variation and lattice operations	138
3	Some examples of total variation distances	143
4	Total variation and minimax rates of convergence	147
5	Hellinger and Kullback-Leibler distances	149
6	The Chen-Stein method for Poisson approximation	153
7	The L^1 norm	159
8	Second-moment bounds on total variation distance	161
9	Minimax lower bounds	169
1	Why minimax?	169
2	General techniques	170
2.1	Reduction to estimation problem	171
2.2	Bayes	171
2.3	Assouad (and Harry)	172
2.4	Fano, Tsybakov	173

3	Lower bounds for minimax risks	174
4	Achievability	179
5	Bounds on total variation distance	179
6	Pointwise estimation of densities	181
7	Quadratic functional of Gaussian means	183
8	Quadratic functionals of densities	186
10	Hellinger differentiability	193
1	Heuristics	193
2	A sufficient condition for Hellinger differentiability	196
3	Information inequality	201
4	Possible trouble at the boundary	205
5	An intrinsic characterization of Hellinger differentiability	206
6	Preservation of DQM under measurable maps	207
7	Differentiability of unit vectors	210
8	Quadratic approximation for log likelihood ratios	212
9	The vanTrees inequality	216
11	Local Asymptotic Normality	225
1	LAN and Gaussian shift families	225
2	Differentiation of unit vectors	230
3	A quadratic approximation to log likelihood ratios	231
4	Quadratic approximation on a grid: Le Cam's method	237
5	Bahadur's rescue of efficiency	238
12	Efficiency and the LAN Property	243
1	Efficiency: sunk and rescued	243
2	Local alternatives and automatic equivariance	246
3	The negligible set of points of superefficiency	248
4	Gaussian shift families	250
5	The convolution theorem	256
6	The local asymptotic minimax theorem	260
13	Case Studies	269
1	Nonlinear least squares	269
2	Estimation of a mode	274
3	Powell's LAD estimator for censored regression	280
4	Binary choice	287

PART II	COMPARISON OF EXPERIMENTS	295
14	Equivalence of experiments	297
1	Randomization	297
2	Markov equivalence	299
3	Kolmogorov equivalence	300
4	Le Cam-randomization.	303
5	Sufficiency and canonical measures	304
6	Choice of sample space	307
15	Randomization	311
1	Le Cam distance	311
2	Linear functionals and measures	312
3	Randomization: three possibilities	315
4	Conditioning	318
5	Sufficiency and canonical measures	319
6	Advantages of Le Cam randomizations	322
7	Choice of sample space	322
8	Problems	323
9	Notes	323
16	Distances between experiments	325
1	Le Cam's distance between experiments	325
2	Sufficient conditions	327
3	Risk functions	328
17	Distance between multinomial and multivariate normal models	331
1	Introduction	331
2	Conditioning	335
3	From multinomial to multivariate normal	337
4	From multivariate normal to multinomial (sketch)	341
5	Hellinger bound for smoothed Binomial	341
6	Independent normals	343
7	Variance stabilizing transformations	345
8	Intepolation of increments	347
PART III	A MINI-COURSE IN MAXIMAL INEQUALITIES	351
19	Maxima of finitely many variables	353

1	Orlicz norms	353
2	Subgaussian tails	357
3	problems	360
4	Notes Need to be checked and expanded	362
21	Symmetrization and conditioning	365
1	Precursors	365
2	Norm bounds via symmetrization	370
3	Refinements of norm bounds	372
4	Examples with norm bounds	374
5	Tail bounds by symmetrization	375
6	Massart's inequality	376
7	Exponential bounds for small variances.	379
8	Conditional chaining	385
9	Norm bounds via symmetrization	389
10	Tail bounds by symmetrization	391
11	Massart's inequality	392
12	Exponential bounds for small variances.	395
22	Combinatorics	404
1	An important process	404
2	Surround dimension	406
3	A combinatorial inequality	410
4	Packing numbers via surround dimension	412
5	An improvement of the packing bound	415
6	Fat shattering	419
7	Haussler and Long	419
23	Concentration inequalities	425
1	Introduction to concentration	425
2	Preview	427
3	Concentration inequalities for normals	429
4	Tensorization	432
5	Concentration for empirical processes	437
24	Concentration: Talagrand's method	443
1	Introduction	443
2	Control by a single point: introduction to the inductive method	447
3	Talagrand's convex hull theorem	450
4	Multi-point control unedited	454

25 Majorizing measures	465
1 Introduction	465
2 Two informative examples	467
3 Reduction to a compact or finite set	470
4 Method of Kwapień and Rosiński	470
5 Rewrite of Bednorz	470
6 Bednorz’s method for bounded sample paths	474
7 Bednorz’s method for continuous sample paths incomplete	478
8 Bounding sums by integrals	480
9 Problems	481
 26 Bracketing methods	 483
1 What is bracketing?	483
2 Independent summands	489
3 A generic bracketing bound	496
4 Phi mixing	501
5 Absolute regularity	505
6 Strong mixing	507
7 Tail probability bounds	508
 APPENDIXES	 517
 A Prerequisites	 519
1 The multivariate normal and chi-square distributions	519
2 Limit theorems for random vectors	521
3 Lebesgue decomposition	522
4 Taylor expansions, and inequalities	522
5 Convexity inequalities	523
6 Exponential inequalities	524
7 Stochastic order symbols	527
8 Exponential families	528
 B Divergence distances	 533
1 Some convexity facts	533
2 Divergence ‘distances’ between probability measures	534
3 Inequalities	536

C	Anderson's lemma	539
D	Gaussian shift families	547
1	Heuristics	547
2	A rigorous argument	548
E	Almost invariant measures	553
1	Heuristics	553
H	Minimax theorem	557
1	A general minimax theorem	557
2	Notes	560
G	Vector Lattices	561
1	Definitions	561
2	Basic facts	562
3	Order-bounded linear functionals	564
4	Lattices of bounded functions	567
5	Bands in a vector lattice	570
6	The band of countably additive measures	574
7	Projections onto bands in a vector lattice	575
I	Nets	579
1	Nets as generalized sequences	579
2	Universal subnets	580
3	Zorn's Lemma	581
J	Poisson counts	583
1	The conditioned Poisson model	583