## Contents

**PART I  BASIC THEORY**

1 **Heuristics**
   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 ................................. 12

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