

# Contents

<b>1</b>	<b>Basic Convergence Concepts and Theorems</b> . . . . .	1
1.1	Some Basic Notation and Convergence Theorems . . . . .	1
1.2	Three Series Theorem and Kolmogorov's Zero-One Law . . . . .	6
1.3	Central Limit Theorem and Law of the Iterated Logarithm . . . . .	7
1.4	Further Illustrative Examples . . . . .	10
1.5	Exercises . . . . .	12
	References . . . . .	16
<b>2</b>	<b>Metrics, Information Theory, Convergence, and Poisson Approximations</b> . . . . .	19
2.1	Some Common Metrics and Their Usefulness . . . . .	20
2.2	Convergence in Total Variation and Further Useful Formulas . . . . .	22
2.3	Information-Theoretic Distances, de Bruijn's Identity, and Relations to Convergence . . . . .	24
2.4	Poisson Approximations . . . . .	28
2.5	Exercises . . . . .	31
	References . . . . .	33
<b>3</b>	<b>More General Weak and Strong Laws and the Delta Theorem</b> . . . . .	35
3.1	General LLNs and Uniform Strong Law . . . . .	35
3.2	Median Centering and Kesten's Theorem . . . . .	38
3.3	The Ergodic Theorem . . . . .	39
3.4	Delta Theorem and Examples . . . . .	40
3.5	Approximation of Moments . . . . .	44
3.6	Exercises . . . . .	45
	References . . . . .	47

<b>4</b>	<b>Transformations</b>	49
4.1	Variance-Stabilizing Transformations	50
4.2	Examples	51
4.3	Bias Correction of the VST	54
4.4	Symmetrizing Transformations	57
4.5	VST or Symmetrizing Transform?	59
4.6	Exercises	59
	References	61
<b>5</b>	<b>More General Central Limit Theorems</b>	63
5.1	The Independent Not IID Case and a Key Example	63
5.2	CLT without a Variance	66
5.3	Combinatorial CLT	67
5.4	CLT for Exchangeable Sequences	68
5.5	CLT for a Random Number of Summands	70
5.6	Infinite Divisibility and Stable Laws	71
5.7	Exercises	77
	References	80
<b>6</b>	<b>Moment Convergence and Uniform Integrability</b>	83
6.1	Basic Results	83
6.2	The Moment Problem	85
6.3	Exercises	88
	References	89
<b>7</b>	<b>Sample Percentiles and Order Statistics</b>	91
7.1	Asymptotic Distribution of One Order Statistic	92
7.2	Joint Asymptotic Distribution of Several Order Statistics	93
7.3	Bahadur Representations	94
7.4	Confidence Intervals for Quantiles	96
7.5	Regression Quantiles	97
7.6	Exercises	98
	References	100
<b>8</b>	<b>Sample Extremes</b>	101
8.1	Sufficient Conditions	101
8.2	Characterizations	105
8.3	Limiting Distribution of the Sample Range	107
8.4	Multiplicative Strong Law	108
8.5	Additive Strong Law	109

8.6	Dependent Sequences .....	111
8.7	Exercises .....	114
	References .....	116
<b>9</b>	<b>Central Limit Theorems for Dependent Sequences .....</b>	<b>119</b>
9.1	Stationary $m$ -dependence .....	119
9.2	Sampling without Replacement .....	121
9.3	Martingales and Examples .....	123
9.4	The Martingale and Reverse Martingale CLTs .....	126
9.5	Exercises .....	127
	References .....	129
<b>10</b>	<b>Central Limit Theorem for Markov Chains .....</b>	<b>131</b>
10.1	Notation and Basic Definitions .....	131
10.2	Normal Limits .....	132
10.3	Nonnormal Limits .....	135
10.4	Convergence to Stationarity: Diaconis-Stroock-Fill Bound .....	135
10.5	Exercises .....	137
	References .....	139
<b>11</b>	<b>Accuracy of Central Limit Theorems .....</b>	<b>141</b>
11.1	Uniform Bounds: Berry-Esseen Inequality .....	142
11.2	Local Bounds .....	144
11.3	The Multidimensional Berry-Esseen Theorems .....	145
11.4	Other Statistics .....	146
11.5	Exercises .....	147
	References .....	149
<b>12</b>	<b>Invariance Principles .....</b>	<b>151</b>
12.1	Motivating Examples .....	152
12.2	Two Relevant Gaussian Processes .....	153
12.3	The Erdős-Kac Invariance Principle .....	156
12.4	Invariance Principles, Donsker's Theorem, and the KMT Construction .....	157
12.5	Invariance Principle for Empirical Processes .....	161
12.6	Extensions of Donsker's Principle and Vapnik-Chervonenkis Classes .....	163
12.7	Glivenko-Cantelli Theorem for VC Classes .....	164
12.8	CLTs for Empirical Measures and Applications .....	167
	12.8.1 Notation and Formulation .....	168
	12.8.2 Entropy Bounds and Specific CLTs .....	169

12.9	Dependent Sequences: Martingales, Mixing, and Short-Range Dependence	172
12.10	Weighted Empirical Processes and Approximations	175
12.11	Exercises	178
	References	180
<b>13</b>	<b>Edgeworth Expansions and Cumulants</b>	<b>185</b>
13.1	Expansion for Means	186
13.2	Using the Edgeworth Expansion	188
13.3	Edgeworth Expansion for Sample Percentiles	189
13.4	Edgeworth Expansion for the $t$ -statistic	190
13.5	Cornish-Fisher Expansions	192
13.6	Cumulants and Fisher's $k$ -statistics	194
13.7	Exercises	198
	References	200
<b>14</b>	<b>Saddlepoint Approximations</b>	<b>203</b>
14.1	Approximate Evaluation of Integrals	204
14.2	Density of Means and Exponential Tilting	208
14.2.1	Derivation by Edgeworth Expansion and Exponential Tilting	210
14.3	Some Examples	211
14.4	Application to Exponential Family and the Magic Formula	213
14.5	Tail Area Approximation and the Lugannani-Rice Formula	213
14.6	Edgeworth vs. Saddlepoint vs. Chi-square Approximation	217
14.7	Tail Areas for Sample Percentiles	218
14.8	Quantile Approximation and Inverting the Lugannani-Rice Formula	219
14.9	The Multidimensional Case	221
14.10	Exercises	222
	References	223
<b>15</b>	<b><math>U</math>-statistics</b>	<b>225</b>
15.1	Examples	226
15.2	Asymptotic Distribution of $U$ -statistics	227
15.3	Moments of $U$ -statistics and the Martingale Structure	229
15.4	Edgeworth Expansions	230
15.5	Nonnormal Limits	232

15.6	Exercises	232
	References	234
<b>16</b>	<b>Maximum Likelihood Estimates</b>	<b>235</b>
16.1	Some Examples	235
16.2	Inconsistent MLEs	239
16.3	MLEs in the Exponential Family	240
16.4	More General Cases and Asymptotic Normality	242
16.5	Observed and Expected Fisher Information	244
16.6	Edgeworth Expansions for MLEs	245
16.7	Asymptotic Optimality of the MLE and Superefficiency	247
16.8	Hajek-LeCam Convolution Theorem	249
16.9	Loss of Information and Efron's Curvature	251
16.10	Exercises	253
	References	258
<b>17</b>	<b>M Estimates</b>	<b>259</b>
17.1	Examples	260
17.2	Consistency and Asymptotic Normality	262
17.3	Bahadur Expansion of M Estimates	265
17.4	Exercises	267
	References	268
<b>18</b>	<b>The Trimmed Mean</b>	<b>271</b>
18.1	Asymptotic Distribution and the Bahadur Representation	271
18.2	Lower Bounds on Efficiencies	273
18.3	Multivariate Trimmed Mean	273
18.4	The 10-20-30-40 Rule	275
18.5	Exercises	277
	References	278
<b>19</b>	<b>Multivariate Location Parameter and Multivariate Medians</b>	<b>279</b>
19.1	Notions of Symmetry of Multivariate Data	279
19.2	Multivariate Medians	280
19.3	Asymptotic Theory for Multivariate Medians	282
19.4	The Asymptotic Covariance Matrix	283
19.5	Asymptotic Covariance Matrix of the $L_1$ Median	284
19.6	Exercises	287
	References	288

<b>20</b>	<b>Bayes Procedures and Posterior Distributions</b>	289
20.1	Motivating Examples	290
20.2	Bernstein-von Mises Theorem	291
20.3	Posterior Expansions	294
20.4	Expansions for Posterior Mean, Variance, and Percentiles	298
20.5	The Tierney-Kadane Approximations	300
20.6	Frequentist Approximation of Posterior Summaries	302
20.7	Consistency of Posteriors	304
20.8	The Difference between Bayes Estimates and the MLE	305
20.9	Using the Brown Identity to Obtain Bayesian Asymptotics	306
20.10	Testing	311
20.11	Interval and Set Estimation	312
20.12	Infinite-Dimensional Problems and the Diaconis-Freedman Results	314
20.13	Exercises	317
	References	320
<b>21</b>	<b>Testing Problems</b>	323
21.1	Likelihood Ratio Tests	323
21.2	Examples	324
21.3	Asymptotic Theory of Likelihood Ratio Test Statistics	334
21.4	Distribution under Alternatives	336
21.5	Bartlett Correction	338
21.6	The Wald and Rao Score Tests	339
21.7	Likelihood Ratio Confidence Intervals	340
21.8	Exercises	342
	References	344
<b>22</b>	<b>Asymptotic Efficiency in Testing</b>	347
22.1	Pitman Efficiencies	348
22.2	Bahadur Slopes and Bahadur Efficiency	353
22.3	Bahadur Slopes of $U$ -statistics	361
22.4	Exercises	362
	References	363
<b>23</b>	<b>Some General Large-Deviation Results</b>	365
23.1	Generalization of the Cramér-Chernoff Theorem	365
23.2	The Gärtner-Ellis Theorem	367

23.3	Large Deviation for Local Limit Theorems . . . . .	370
23.4	Exercises . . . . .	374
	References . . . . .	375
<b>24</b>	<b>Classical Nonparametrics . . . . .</b>	<b>377</b>
24.1	Some Early Illustrative Examples . . . . .	378
24.2	Sign Test . . . . .	380
24.3	Consistency of the Sign Test . . . . .	381
24.4	Wilcoxon Signed-Rank Test . . . . .	383
24.5	Robustness of the $t$ Confidence Interval . . . . .	388
24.6	The Bahadur-Savage Theorem . . . . .	393
24.7	Kolmogorov-Smirnov and Anderson Confidence Intervals . . . . .	394
24.8	Hodges-Lehmann Confidence Interval . . . . .	396
24.9	Power of the Wilcoxon Test . . . . .	397
24.10	Exercises . . . . .	398
	References . . . . .	399
<b>25</b>	<b>Two-Sample Problems . . . . .</b>	<b>401</b>
25.1	Behrens-Fisher Problem . . . . .	402
25.2	Wilcoxon Rank Sum and Mann-Whitney Test . . . . .	405
25.3	Two-Sample $U$ -statistics and Power Approximations . . . . .	408
25.4	Hettmansperger's Generalization . . . . .	410
25.5	The Nonparametric Behrens-Fisher Problem . . . . .	412
25.6	Robustness of the Mann-Whitney Test . . . . .	415
25.7	Exercises . . . . .	417
	References . . . . .	418
<b>26</b>	<b>Goodness of Fit . . . . .</b>	<b>421</b>
26.1	Kolmogorov-Smirnov and Other Tests Based on $F_n$ . . . . .	422
26.2	Computational Formulas . . . . .	422
26.3	Some Heuristics . . . . .	423
26.4	Asymptotic Null Distributions of $D_n, C_n, A_n,$ and $V_n$ . . . . .	424
26.5	Consistency and Distributions under Alternatives . . . . .	425
26.6	Finite Sample Distributions and Other EDF-Based Tests . . . . .	426
26.7	The Berk-Jones Procedure . . . . .	428
26.8	$\varphi$ -Divergences and the Jager-Wellner Tests . . . . .	429
26.9	The Two-Sample Case . . . . .	431
26.10	Tests for Normality . . . . .	434
26.11	Exercises . . . . .	436
	References . . . . .	438

<b>27</b>	<b>Chi-square Tests for Goodness of Fit</b>	441
27.1	The Pearson $\chi^2$ Test	441
27.2	Asymptotic Distribution of Pearson's Chi-square	442
27.3	Asymptotic Distribution under Alternatives and Consistency	442
27.4	Choice of $k$	443
27.5	Recommendation of Mann and Wald	445
27.6	Power at Local Alternatives and Choice of $k$	445
27.7	Exercises	448
	References	449
<b>28</b>	<b>Goodness of Fit with Estimated Parameters</b>	451
28.1	Preliminary Analysis by Stochastic Expansion	452
28.2	Asymptotic Distribution of EDF-Based Statistics for Composite Nulls	453
28.3	Chi-square Tests with Estimated Parameters and the Chernoff-Lehmann Result	455
28.4	Chi-square Tests with Random Cells	457
28.5	Exercises	457
	References	458
<b>29</b>	<b>The Bootstrap</b>	461
29.1	Bootstrap Distribution and the Meaning of Consistency	462
29.2	Consistency in the Kolmogorov and Wasserstein Metrics	464
29.3	Delta Theorem for the Bootstrap	468
29.4	Second-Order Accuracy of the Bootstrap	468
29.5	Other Statistics	471
29.6	Some Numerical Examples	473
29.7	Failure of the Bootstrap	475
29.8	$m$ out of $n$ Bootstrap	476
29.9	Bootstrap Confidence Intervals	478
29.10	Some Numerical Examples	482
29.11	Bootstrap Confidence Intervals for Quantiles	483
29.12	Bootstrap in Regression	483
29.13	Residual Bootstrap	484
29.14	Confidence Intervals	485
29.15	Distribution Estimates in Regression	486
29.16	Bootstrap for Dependent Data	487
29.17	Consistent Bootstrap for Stationary Autoregression	488



29.18	Block Bootstrap Methods	489
29.19	Optimal Block Length	491
29.20	Exercises	492
	References	495
<b>30</b>	<b>Jackknife</b>	499
30.1	Notation and Motivating Examples	499
30.2	Bias Correction by the Jackknife	502
30.3	Variance Estimation	503
30.4	Delete- $d$ Jackknife and von Mises Functionals	504
30.5	A Numerical Example	507
30.6	Jackknife Histogram	508
30.7	Exercises	511
	References	512
<b>31</b>	<b>Permutation Tests</b>	513
31.1	General Permutation Tests and Basic Group Theory	514
31.2	Exact Similarity of Permutation Tests	516
31.3	Power of Permutation Tests	519
31.4	Exercises	520
	References	521
<b>32</b>	<b>Density Estimation</b>	523
32.1	Basic Terminology and Some Popular Methods	523
32.2	Measures of the Quality of Density Estimates	526
32.3	Certain Negative Results	526
32.4	Minimaxity Criterion	529
32.5	Performance of Some Popular Methods: A Preview	530
32.6	Rate of Convergence of Histograms	531
32.7	Consistency of Kernel Estimates	533
32.8	Order of Optimal Bandwidth and Superkernels	535
32.9	The Epanechnikov Kernel	538
32.10	Choice of Bandwidth by Cross Validation	539
	32.10.1 Maximum Likelihood CV	540
	32.10.2 Least Squares CV	542
	32.10.3 Stone's Result	544
32.11	Comparison of Bandwidth Selectors and Recommendations	545
32.12	$\mathcal{L}^1$ Optimal Bandwidths	547
32.13	Variable Bandwidths	548
32.14	Strong Uniform Consistency and Confidence Bands	550

32.15	Multivariate Density Estimation and Curse of Dimensionality . . . . .	552
32.15.1	Kernel Estimates and Optimal Bandwidths . . . . .	556
32.16	Estimating a Unimodal Density and the Grenander Estimate . . . . .	558
32.16.1	The Grenander Estimate . . . . .	558
32.17	Mode Estimation and Chernoff's Distribution . . . . .	561
32.18	Exercises . . . . .	564
	References . . . . .	568
<b>33</b>	<b>Mixture Models and Nonparametric Deconvolution . . . . .</b>	<b>571</b>
33.1	Mixtures as Dense Families . . . . .	572
33.2	$z$ Distributions and Other Gaussian Mixtures as Useful Models . . . . .	573
33.3	Estimation Methods and Their Properties: Finite Mixtures . . . . .	577
33.3.1	Maximum Likelihood . . . . .	577
33.3.2	Minimum Distance Method . . . . .	578
33.3.3	Moment Estimates . . . . .	579
33.4	Estimation in General Mixtures . . . . .	580
33.5	Strong Consistency and Weak Convergence of the MLE . . . . .	582
33.6	Convergence Rates for Finite Mixtures and Nonparametric Deconvolution . . . . .	584
33.6.1	Nonparametric Deconvolution . . . . .	585
33.7	Exercises . . . . .	587
	References . . . . .	589
<b>34</b>	<b>High-Dimensional Inference and False Discovery . . . . .</b>	<b>593</b>
34.1	Chi-square Tests with Many Cells and Sparse Multinomials . . . . .	594
34.2	Regression Models with Many Parameters: The Portnoy Paradigm . . . . .	597
34.3	Multiple Testing and False Discovery: Early Developments . . . . .	599
34.4	False Discovery: Definitions, Control, and the Benjamini-Hochberg Rule . . . . .	601
34.5	Distribution Theory for False Discoveries and Poisson and First-Passage Asymptotics . . . . .	604
34.6	Newer FDR Controlling Procedures . . . . .	606
34.6.1	Storey-Taylor-Siegmund Rule . . . . .	606
34.7	Higher Criticism and the Donoho-Jin Developments . . . . .	608

- 34.8 False Nondiscovery and Decision Theory Formulation . . . 611
  - 34.8.1 Genovese-Wasserman Procedure . . . . . 612
- 34.9 Asymptotic Expansions . . . . . 614
- 34.10 Lower Bounds on the Number of False Hypotheses . . . . 616
  - 34.10.1 Bühlmann-Meinshausen-Rice Method . . . . . 617
- 34.11 The Dependent Case and the Hall-Jin Results . . . . . 620
  - 34.11.1 Increasing and Multivariate Totally  
Positive Distributions . . . . . 620
  - 34.11.2 Higher Criticism under Dependence:  
Hall-Jin Results . . . . . 623
- 34.12 Exercises . . . . . 625
- References . . . . . 628
  
- 35 A Collection of Inequalities in Probability, Linear Algebra,  
and Analysis . . . . . 633**
  - 35.1 Probability Inequalities . . . . . 633
    - 35.1.1 Improved Bonferroni Inequalities . . . . . 633
    - 35.1.2 Concentration Inequalities . . . . . 634
    - 35.1.3 Tail Inequalities for Specific Distributions . . . . 639
    - 35.1.4 Inequalities under Unimodality . . . . . 641
    - 35.1.5 Moment and Monotonicity Inequalities . . . . . 643
    - 35.1.6 Inequalities in Order Statistics . . . . . 652
    - 35.1.7 Inequalities for Normal Distributions . . . . . 655
    - 35.1.8 Inequalities for Binomial and Poisson  
Distributions . . . . . 656
    - 35.1.9 Inequalities in the Central Limit Theorem . . . . 658
    - 35.1.10 Martingale Inequalities . . . . . 661
  - 35.2 Matrix Inequalities . . . . . 663
    - 35.2.1 Rank, Determinant, and Trace Inequalities . . . 663
    - 35.2.2 Eigenvalue and Quadratic Form Inequalities . . . 667
  - 35.3 Series and Polynomial Inequalities . . . . . 671
  - 35.4 Integral and Derivative Inequalities . . . . . 675
  
- Glossary of Symbols . . . . . 689**
  
- Index . . . . . 693**