# Asymptotic Theory of Statistics and Probability

Anirban DasGupta

To my mother, and to the loving memories of my father

# Contents


**1 Basic Convergence Concepts and Theorems** **10**

1.1 Some Basic Notation and Convergence Theorems . . . . . 10
1.2 Three Series Theorem and Kolmogorov's Zero-One Law . . . . . 15
1.3 Central Limit Theorem and Law of the Iterated Logarithm . . . . . 16
1.4 Further Illustrative Examples . . . . . 18
1.5 Exercises . . . . . 21
1.6 References . . . . . 25

**2 Metrics, Information Theory, Convergence, and Poisson Approximations** **26**

2.1 Some Common Metrics and Their Usefulness . . . . . 27
2.2 Convergence in Total Variation and Further Useful Formulas . . . . 29
2.3 Information Theoretic Distances, de Bruijn's Identity and Relations to Convergence . . . . . 31
2.4 Poisson Approximations . . . . . 36
2.5 Exercises . . . . . 40
2.6 References . . . . . 41

**3 More General Weak and Strong Laws and the Delta Theorem** **44**

3.1 General LLN and Uniform Strong Law . . . . . 44
3.2 Median Centering and Kesten's Theorem . . . . . 46
3.3 The Ergodic Theorem . . . . . 47
3.4 Delta Theorem and Examples . . . . . 49
3.5 Approximation of Moments . . . . . 52
3.6 Exercises . . . . . 54
3.7 References . . . . . 55

**4 Transformations** **57**

4.1 Variance Stabilizing Transformations . . . . . 58
4.2 Examples . . . . . 59
4.3 Bias Correction of the VST . . . . . 61
4.4 Symmetrizing Transformations . . . . . 64
4.5 VST or Symmetrizing Transform ? . . . . . 66
4.6 Exercises . . . . . 68
4.7 References . . . . . 69

**5 More General CLTs** **71**
5.1 The Independent Not IID Case and a Key Example . . . . . 71
5.2 CLT without a Variance . . . . . 73
5.3 Combinatorial CLT . . . . . 74
5.4 CLT for Exchangeable Sequences . . . . . 75
5.5 CLT for a Random Number of Summands . . . . . 77
5.6 Infinite Divisibility and Stable Laws . . . . . 78
5.7 Exercises . . . . . 85
5.8 References . . . . . 87

**6 Moment Convergence and Uniform Integrability** **89**
6.1 Basic Results . . . . . 89
6.2 The Moment Problem . . . . . 91
6.3 Exercises . . . . . 94
6.4 References . . . . . 95

**7 Sample Percentiles and Order Statistics** **96**
7.1 Asymptotic Distribution of One Order Statistic . . . . . 96
7.2 Joint Asymptotic Distribution of Several Order Statistics . . . . . 98
7.3 Bahadur Representations . . . . . 99
7.4 Confidence Intervals for Quantiles . . . . . 100
7.5 Regression Quantiles . . . . . 101
7.6 Exercises . . . . . 103
7.7 References . . . . . 104

**8 Sample Extremes** **106**
8.1 Sufficient Conditions . . . . . 106
8.2 Characterizations . . . . . 109
8.3 Limiting Distribution of the Sample Range . . . . . 111
8.4 Multiplicative Strong Law . . . . . 112
8.5 Additive Strong Law . . . . . 113
8.6 Dependent Sequences . . . . . 114
8.7 Exercises . . . . . 117
8.8 References . . . . . 120

**9 Central Limit Theorems for Dependent Sequences 122**
9.1 Stationary m-dependence . . . 122
9.2 Sampling Without Replacement . . . 123
9.3 Martingales and Examples . . . 125
9.4 The Martingale and Reverse Martingale CLT . . . 128
9.5 Exercises . . . 130
9.6 References . . . 131

**10 Central Limit Theorem for Markov Chains 133**
10.1 Notation and Basic Definitions . . . 133
10.2 Normal Limits . . . 134
10.3 Nonnormal Limits . . . 137
10.4 Convergence to Stationarity: Diaconis-Stroock-Fill Bound . . . 137
10.5 Exercises . . . 140
10.6 References . . . 142

**11 Accuracy of CLTs 144**
11.1 Uniform Bounds: Berry-Esseen Inequality . . . 144
11.2 Local Bounds . . . 147
11.3 The Multidimensional Berry-Esseen Theorems . . . 147
11.4 Other Statistics . . . 149
11.5 Exercises . . . 151
11.6 References . . . 152

**12 Invariance Principles 153**
12.1 Motivating Examples . . . 154
12.2 Two Relevant Gaussian Processes . . . 155
12.3 The Erdös-Kac Invariance Principle . . . 157
12.4 Invariance Principles, Donsker's Theorem and the KMT Construction 159
12.5 Invariance Principle for Empirical Processes . . . 162
12.6 Extensions of Donsker's Principle and Vapnik-Chervonenkis Classes 164
12.7 Glivenko-Cantelli Theorem for VC Classes . . . 165
12.8 CLTs for Empirical Measures and Applications . . . 168
12.8.1 Notation and Formulation . . . 169
12.8.2 Entropy Bounds and Specific CLTs . . . 170

12.9 Dependent Sequences: Martingales, Mixing and Short Range Dependence . . . 173
12.10 Weighted Empirical Processes and Approximations . . . 177
12.11 Exercises . . . 180
12.12 References . . . 182

**13 Edgeworth Expansions and Cumulants** **187**
13.1 Expansion for Means . . . 187
13.2 Using the Edgeworth Expansion . . . 190
13.3 Edgeworth Expansion for Sample Percentiles . . . 190
13.4 Edgeworth Expansion for the t-statistic . . . 192
13.5 Cornish-Fisher Expansions . . . 194
13.6 Cumulants and Fisher's $k$-Statistics . . . 195
13.7 Exercises . . . 198
13.8 References . . . 201

**14 Saddlepoint Approximations** **203**
14.1 Approximate Evaluation of Integrals . . . 203
14.2 Density of Means and Exponential Tilting . . . 207
14.2.1 Derivation by Edgeworth Expansion and Exponential Tilting . . . 209
14.3 Some Examples . . . 210
14.4 Application to Exponential Family and the Magic Formula . . . 212
14.5 Tail Area Approximation and the Lugannani-Rice Formula . . . 213
14.6 Edgeworth vs Saddlepoint vs Chisquare Approximation . . . 216
14.7 Tail Areas for Sample Percentiles . . . 217
14.8 Quantile Approximation and Inverting the Lugannani-Rice Formula 218
14.9 The Multidimensional Case . . . 220
14.10 Exercises . . . 222
14.11 References . . . 223

**15 $U$-Statistics** **225**
15.1 Examples . . . 225
15.2 Asymptotic Distribution of U-statistics . . . 226
15.3 Moments of U-statistics and the Martingale Structure . . . 228
15.4 Edgeworth Expansions . . . 229

15.5 Nonnormal Limits . . . 231
15.6 Exercises . . . 232
15.7 References . . . 233

**16 Maximum Likelihood Estimates 235**
16.1 Some Examples . . . 235
16.2 Inconsistent MLEs . . . 238
16.3 MLEs in Exponential Family . . . 239
16.4 More General Cases and Asymptotic Normality . . . 241
16.5 Observed and Expected Fisher Information . . . 243
16.6 Edgeworth Expansions for MLEs . . . 244
16.7 Asymptotic Optimality of the MLE and Superefficiency . . . 246
16.8 Hajek-Le Cam Convolution Theorem . . . 248
16.9 Loss of Information and Efron's Curvature . . . 249
16.10 Exercises . . . 253
16.11 References . . . 257

**17 M Estimates 259**
17.1 Examples . . . 260
17.2 Consistency and Asymptotic Normality . . . 262
17.3 Bahadur Expansion of M Estimates . . . 265
17.4 Exercises . . . 267
17.5 References . . . 268

**18 The Trimmed Mean 269**
18.1 Asymptotic Distribution and the Bahadur Representation . . . 269
18.2 Lower Bounds on Efficiencies . . . 270
18.3 Multivariate Trimmed Mean . . . 271
18.4 The $10 - 20 - 30 - 40$ Rule . . . 273
18.5 Exercises . . . 276
18.6 References . . . 276

**19 Multivariate Location Parameter and Multivariate Medians 278**
19.1 Notions of Symmetry of Multivariate Data . . . 278
19.2 Multivariate Medians . . . 279
19.3 Asymptotic Theory for Multivariate Medians . . . 280
19.4 The Asymptotic Covariance Matrix . . . 281

19.5 Asymptotic Covariance Matrix of the $L_1$ median . . . . . 283
19.6 Exercises . . . . . 286
19.7 References . . . . . 286

**20 Bayes Procedures and Posterior Distributions 287**
20.1 Motivating Examples . . . . . 287
20.2 Bernstein-von Mises Theorem . . . . . 289
20.3 Posterior Expansions . . . . . 292
20.4 Expansions for Posterior Mean, Variance, and Percentiles . . . . . 296
20.5 The Tierney-Kadane Approximation . . . . . 298
20.6 Frequentist Approximation of Posterior Summaries . . . . . 299
20.7 Consistency of Posteriors . . . . . 301
20.8 The Difference Between Bayes Estimates and the MLE . . . . . 303
20.9 Using the Brown Identity to Obtain Bayesian Asymptotics . . . . . 303
20.10 Testing . . . . . 307
20.11 Interval and Set Estimation . . . . . 308
20.12 Infinite Dimensional Problems and the Diaconis-Freedman Results 309
20.13 Exercises . . . . . 314
20.14 References . . . . . 317

**21 Testing Problems 319**
21.1 Likelihood Ratio Tests . . . . . 319
21.2 Examples . . . . . 320
21.3 Asymptotic Theory of Likelihood Ratio Test Statistics . . . . . 329
21.4 Distribution under Alternatives . . . . . 330
21.5 Bartlett Correction . . . . . 332
21.6 The Wald and Rao Score Tests . . . . . 332
21.7 Likelihood Ratio Confidence Intervals . . . . . 334
21.8 Exercises . . . . . 336
21.9 References . . . . . 338

**22 Asymptotic Efficiency in Testing 340**
22.1 Pitman Efficiencies . . . . . 341
22.2 Bahadur Slopes and Bahadur Efficiency . . . . . 346
22.3 Bahadur slopes of $U$ statistics . . . . . 353
22.4 Exercises . . . . . 355

22.5 References . . . . . 356

**23 Some General Large Deviation Results** **358**
23.1 Generalization of the Cramér-Chernoff Theorem . . . . . 358
23.2 The Gärtner-Ellis Theorem . . . . . 360
23.3 Large Deviation for Local Limit Theorems. . . . . 363
23.4 Exercises . . . . . 367
23.5 References . . . . . 368

**24 Classical Nonparametrics** **370**
24.1 Some Early Illustrative Examples . . . . . 371
24.2 Sign Test . . . . . 372
24.3 Consistency of the Sign Test . . . . . 374
24.4 Wilcoxon Signed-Rank Test . . . . . 376
24.5 Robustness of the $t$-Confidence Interval . . . . . 380
24.6 The Bahadur-Savage Theorem . . . . . 385
24.7 Kolmogorov-Smirnov & Anderson Confidence Intervals . . . . . 386
24.8 Hodges-Lehmann Confidence Interval . . . . . 388
24.9 Power of the Wilcoxon Test . . . . . 389
24.10 Exercises . . . . . 389
24.11 References . . . . . 391

**25 Two-Sample Problems** **392**
25.1 Behrens-Fisher Problem . . . . . 392
25.2 Wilcoxon Rank-Sum and Mann-Whitney Test . . . . . 396
25.3 Two-Sample U-Statistics & Power Approximations . . . . . 398
25.4 Hettmansperger's Generalization . . . . . 400
25.5 The Nonparametric Behrens-Fisher Problem . . . . . 402
25.6 Robustness of the Mann-Whitney Test . . . . . 405
25.7 Exercises . . . . . 407
25.8 References . . . . . 408

**26 Goodness of Fit** **411**
26.1 Kolmogorov-Smirnov and Other Tests Based on $F_n$ . . . . . 411
26.2 Computational Formulas . . . . . 412
26.3 Some Heuristics . . . . . 413
26.4 Asymptotic Null Distributions of $D_n, C_n, A_n$ and $V_n$ . . . . . 413

26.5 Consistency and Distributions under Alternative . . . . . 415
26.6 Finite Sample Distributions and Other EDF Based Tests . . . . . 416
26.7 The Berk-Jones Procedure . . . . . 417
26.8 $\varphi$-Divergences and the Jager-Wellner Tests . . . . . 419
26.9 The Two Sample Case . . . . . 421
26.10 Tests for Normality . . . . . 423
26.11 Exercises . . . . . 425
26.12 References . . . . . 427

**27 Chi-square Tests for Goodness of Fit** **430**
27.1 The Pearson $\chi^2$ Test . . . . . 430
27.2 Asymptotic Distribution of Pearson's Chi-square . . . . . 430
27.3 Asymptotic Distribution Under Alternative and Consistency . . . . 431
27.4 Choice of $k$ . . . . . 432
27.5 Recommendation of Mann and Wald . . . . . 433
27.6 Power at Local Alternatives and Choice of $k$ . . . . . 434
27.7 Exercises . . . . . 438
27.8 References . . . . . 439

**28 Goodness of Fit with Estimated Parameters** **440**
28.1 Preliminary Analysis by Stochastic Expansions . . . . . 440
28.2 Asymptotic Distribution of EDF Based Statistics for Composite Nulls . . . . . 442
28.3 Chisquare Tests with Estimated Parameters and the Chernoff-Lehmann Result . . . . . 444
28.4 Chisquare Tests With Random Cells . . . . . 446
28.5 Exercises . . . . . 447
28.6 References . . . . . 448

**29 The Bootstrap** **450**
29.1 Bootstrap Distribution and Meaning of Consistency . . . . . 451
29.2 Consistency in the Kolmogorov and Wasserstein Metric . . . . . 453
29.3 Delta Theorem for the Bootstrap . . . . . 456
29.4 Second Order Accuracy of Bootstrap . . . . . 457
29.5 Other Statistics . . . . . 460
29.6 Some Numerical Examples . . . . . 461

29.7 Failure of Bootstrap . . . 463
29.8 $m$ out of $n$ Bootstrap . . . 465
29.9 Bootstrap Confidence Intervals . . . 466
29.10 Some Numerical Examples . . . 470
29.11 Bootstrap Confidence Intervals for Quantiles . . . 471
29.12 Bootstrap in Regression . . . 472
29.13 Residual Bootstrap . . . 472
29.14 Confidence Intervals . . . 473
29.15 Distribution Estimates in Regression . . . 474
29.16 Bootstrap for Dependent Data . . . 475
29.17 Consistent Bootstrap for Stationary Autoregression . . . 477
29.18 Block Bootstrap Methods . . . 478
29.19 Optimal Block Length . . . 480
29.20 Exercises . . . 481
29.21 References . . . 484

**30 Jackknife** **488**
30.1 Notation and Motivating Examples . . . 488
30.2 Bias Correction by Jackknife . . . 490
30.3 Variance Estimation . . . 491
30.4 Delete-d Jackknife and von Mises Functionals . . . 493
30.5 A Numerical Example . . . 495
30.6 Jackknife Histogram . . . 496
30.7 Exercises . . . 499
30.8 References . . . 500

**31 Permutation Tests** **501**
31.1 General Permutation Tests and Basic Group Theory . . . 502
31.2 Exact Similarity of Permutation Tests . . . 504
31.3 Power of Permutation Tests . . . 506
31.4 Exercises . . . 508
31.5 References . . . 508

**32 Density Estimation** **510**
32.1 Basic Terminology and Some Popular Methods . . . 510
32.2 Measures of Quality of Density Estimates . . . 512

32.3 Certain Negative Results . . . 513
32.4 Minimaxity Criterion . . . 515
32.5 Performance of Some Popular Methods: A Preview . . . 516
32.6 Rate of Convergence of Histograms . . . 518
32.7 Consistency of Kernel Estimates . . . 519
32.8 Order of Optimal Bandwidth and Superkernels . . . 522
32.9 The Epanechnikov Kernel . . . 525
32.10 Choice of Bandwidth by Cross Validation . . . 525
32.10.1 Maximum Likelihood CV . . . 526
32.10.2 Least Squares CV . . . 528
32.10.3 Stone's Result . . . 530
32.11 Comparison of Bandwidth Selectors and Recommendations . . . 532
32.12 $\mathcal{L}^1$ Optimal Bandwidths . . . 533
32.13 Variable Bandwidths . . . 535
32.14 Strong Uniform Consistency and Confidence Bands . . . 536
32.15 Multivariate Density Estimation and Curse of Dimensionality . . . 538
32.15.1 Kernel Estimates and Optimal Bandwidths . . . 542
32.16 Estimating a Unimodal Density and the Grenander Estimate . . . 543
32.16.1 The Grenander Estimate . . . 544
32.17 Mode Estimation and Chernoff's Distribution . . . 547
32.18 Exercises . . . 550
32.19 References . . . 553

**33 Mixture Models and Nonparametric Deconvolution 558**
33.1 Mixtures as Dense Families . . . 558
33.2 $z$ Distributions and Other Gaussian Mixtures as Useful Models . . . 560
33.3 Estimation Methods and Their Properties: Finite Mixtures . . . 563
33.3.1 Maximum Likelihood . . . 564
33.3.2 Minimum Distance Method . . . 565
33.3.3 Moment Estimates . . . 566
33.4 Estimation in General Mixtures . . . 567
33.5 Strong Consistency and Weak Convergence of the MLE . . . 569
33.6 Convergence Rates for Finite Mixtures and Nonparametric Deconvolution . . . 571
33.6.1 Nonparametric Deconvolution . . . 572
33.7 Exercises . . . 575

33.8 References . . . . . 576

**34 High Dimensional Inference and False Discovery 581**

34.1 Chisquare Tests with Many Cells and Sparse Multinomials . . . . . 582
34.2 Regression Models with Many Parameters: The Portnoy Paradigm . 585
34.3 Multiple Testing and False Discovery: Early Developments . . . . . 587
34.4 False Discovery : Definitions, Control and the Benjamini-Hochberg Rule . . . . . 589
34.5 Distribution Theory for False Discoveries and Poisson and First Passage Asymptotics . . . . . 592
34.6 Newer FDR Controlling Procedures . . . . . 594
34.6.1 Storey-Taylor-Siegmund Rule . . . . . 594
34.7 Higher Criticism and the Donoho-Jin Developments . . . . . 596
34.8 False Nondiscovery and Decision Theory Formulation . . . . . 599
34.8.1 Genovese-Wasserman Procedure . . . . . 600
34.9 Asymptotic Expansions . . . . . 602
34.10 Lower Bounds on Number of False Hypotheses . . . . . 604
34.10.1 B*ü*hlmann-Meinshausen-Rice Method . . . . . 605
34.11 The Dependent Case and the Hall-Jin Results . . . . . 608
34.11.1 Increasing and Multivariate Totally Positive Distributions 608
34.11.2 Higher Criticism Under Dependence : Hall-Jin Results . . 611
34.12 Exercises . . . . . 613
34.13 References . . . . . 616

**35 A Collection of Inequalities in Probability, Linear Algebra and Analysis 621**

35.1 Probability Inequalities . . . . . 621
35.1.1 **Improved Bonferroni Inequalities** . . . . . 621
35.1.2 **Concentration Inequalities** . . . . . 622
35.1.3 **Tail Inequalities for Specific Distributions** . . . . . 626
35.1.4 **Inequalities under Unimodality** . . . . . 627
35.1.5 **Moment and Monotonicity Inequalities** . . . . . 629
35.1.6 **Inequalities on Order Statistics** . . . . . 637
35.1.7 **Inequalities for Normal Distributions** . . . . . 640
35.1.8 **Inequalities for Binomial and Poisson** . . . . . 641
35.1.9 **Inequalities on the Central Limit Theorem** . . . . . 643

35.1.10 **Martingale Inequalities** . . . . . 645
35.2 Matrix Inequalities . . . . . 647
35.2.1 **Rank, Determinant and Trace Inequalities** . . . . . 647
35.2.2 **Eigenvalue and Quadratic Form Inequalities** . . . . 650
35.3 Series and Polynomial Inequalities . . . . . 654
35.4 Integral and Derivative Inequalities . . . . . 658

**36 Glossary of Symbols** **667**

# Recommended Chapter Selections

| Course Type | Chapters |
|---|---|
| Semester I, Classical asymptotics | 1,2,3,4,7,8,11,13,15,17,21,26,27 |
| Semester II, Classical asymptotics | 9,14,16,22,24,25,28,29,30,31,32 |
| Semester I, Inference | 1,2,3,4,7,14,16,17,19,20,21,26,27 |
| Semester II, Inference | 8,11,12,13,22,24,25,29,30,32,33,34 |
| Semester I, Emphasis on Probability | 1,2,3,4,5,6,8,9,10,11,12,23 |
| Semester I, Contemporary topics | 1,2,3,8,10,12,14,29,30,32,33,34 |
| Semester I, Nonparametrics | 1,3,5,7,11,13,15,18,24,26, 29,30,32 |
| Semester I, Modelling and data analysis | 1,3,4,8,9,10,16,19,26,27,29,32,33 |
| My Favorite Course, Semester I | 1,2,3,4,6,7,8,11,13,14,15,16,20 |
| My Favorite Course, Semester II | 5,9,12,17,21,22,24,26,28,29,30,32,34 |