

Contents

Suggested Courses with Different Themes	xix
1 Review of Univariate Probability	1
1.1 Experiments and Sample Spaces	1
1.2 Conditional Probability and Independence.....	5
1.3 Integer-Valued and Discrete Random Variables	8
1.3.1 CDF and Independence.....	9
1.3.2 Expectation and Moments.....	13
1.4 Inequalities	19
1.5 Generating and Moment-Generating Functions	22
1.6 * Applications of Generating Functions to a Pattern Problem	26
1.7 Standard Discrete Distributions	28
1.8 Poisson Approximation to Binomial	34
1.9 Continuous Random Variables.....	36
1.10 Functions of a Continuous Random Variable	42
1.10.1 Expectation and Moments.....	45
1.10.2 Moments and the Tail of a CDF	49
1.11 Moment-Generating Function and Fundamental Inequalities.....	51
1.11.1 * Inversion of an MGF and Post's Formula	53
1.12 Some Special Continuous Distributions.....	54
1.13 Normal Distribution and Confidence Interval for a Mean	61
1.14 Stein's Lemma.....	66
1.15 * Chernoff's Variance Inequality	68
1.16 * Various Characterizations of Normal Distributions	69
1.17 Normal Approximations and Central Limit Theorem	71
1.17.1 Binomial Confidence Interval	74
1.17.2 Error of the CLT	76
1.18 Normal Approximation to Poisson and Gamma	79
1.18.1 Confidence Intervals.....	80
1.19 * Convergence of Densities and Edgeworth Expansions.....	82
References.....	92

- 2 Multivariate Discrete Distributions** 95
 - 2.1 Bivariate Joint Distributions and Expectations of Functions 95
 - 2.2 Conditional Distributions and Conditional Expectations 100
 - 2.2.1 Examples on Conditional Distributions
and Expectations 101
 - 2.3 Using Conditioning to Evaluate Mean and Variance 104
 - 2.4 Covariance and Correlation 107
 - 2.5 Multivariate Case 111
 - 2.5.1 Joint MGF 112
 - 2.5.2 Multinomial Distribution 114
 - 2.6 * The Poissonization Technique 116

- 3 Multidimensional Densities** 123
 - 3.1 Joint Density Function and Its Role 123
 - 3.2 Expectation of Functions 132
 - 3.3 Bivariate Normal 136
 - 3.4 Conditional Densities and Expectations 140
 - 3.4.1 Examples on Conditional Densities and Expectations 142
 - 3.5 Posterior Densities, Likelihood Functions, and Bayes Estimates 147
 - 3.6 Maximum Likelihood Estimates 152
 - 3.7 Bivariate Normal Conditional Distributions 154
 - 3.8 * Useful Formulas and Characterizations for Bivariate Normal 155
 - 3.8.1 Computing Bivariate Normal Probabilities 157
 - 3.9 * Conditional Expectation Given a Set and Borel’s Paradox 158
- References 165

- 4 Advanced Distribution Theory** 167
 - 4.1 Convolutions and Examples 167
 - 4.2 Products and Quotients and the t - and F -Distribution 172
 - 4.3 Transformations 177
 - 4.4 Applications of Jacobian Formula 178
 - 4.5 Polar Coordinates in Two Dimensions 180
 - 4.6 * n -Dimensional Polar and Helmert’s Transformation 182
 - 4.6.1 Efficient Spherical Calculations with Polar
Coordinates 182
 - 4.6.2 Independence of Mean and Variance
in Normal Case 185
 - 4.6.3 The t Confidence Interval 187
 - 4.7 The Dirichlet Distribution 188
 - 4.7.1 * Picking a Point from the Surface of a Sphere 191
 - 4.7.2 * Poincaré’s Lemma 191
 - 4.8 * Ten Important High-Dimensional Formulas
for Easy Reference 191
- References 197

5 Multivariate Normal and Related Distributions199

5.1 Definition and Some Basic Properties199

5.2 Conditional Distributions202

5.3 Exchangeable Normal Variables205

5.4 Sampling Distributions Useful in Statistics207

5.4.1 * Wishart Expectation Identities208

5.4.2 * Hotelling’s T^2 and Distribution of Quadratic Forms209

5.4.3 * Distribution of Correlation Coefficient212

5.5 Noncentral Distributions213

5.6 Some Important Inequalities for Easy Reference214

References218

6 Finite Sample Theory of Order Statistics and Extremes221

6.1 Basic Distribution Theory221

6.2 More Advanced Distribution Theory225

6.3 Quantile Transformation and Existence of Moments229

6.4 Spacings233

6.4.1 Exponential Spacings and Rényi’s Representation233

6.4.2 Uniform Spacings234

6.5 Conditional Distributions and Markov Property235

6.6 Some Applications238

6.6.1 * Records238

6.6.2 The Empirical CDF241

6.7 * Distribution of the Multinomial Maximum243

References247

7 Essential Asymptotics and Applications249

7.1 Some Basic Notation and Convergence Concepts250

7.2 Laws of Large Numbers254

7.3 Convergence Preservation259

7.4 Convergence in Distribution262

7.5 Preservation of Convergence and Statistical Applications267

7.5.1 Slutsky’s Theorem268

7.5.2 Delta Theorem269

7.5.3 Variance Stabilizing Transformations272

7.6 Convergence of Moments274

7.6.1 Uniform Integrability275

7.6.2 The Moment Problem and Convergence in Distribution...277

7.6.3 Approximation of Moments278

7.7 Convergence of Densities and Scheffé’s Theorem282

References292

8	Characteristic Functions and Applications	293
8.1	Characteristic Functions of Standard Distributions	294
8.2	Inversion and Uniqueness	298
8.3	Taylor Expansions, Differentiability, and Moments	302
8.4	Continuity Theorems	303
8.5	Proof of the CLT and the WLLN	305
8.6	* Producing Characteristic Functions	306
8.7	Error of the Central Limit Theorem	308
8.8	Lindeberg–Feller Theorem for General Independent Case	311
8.9	* Infinite Divisibility and Stable Laws	315
8.10	* Some Useful Inequalities	317
	References	322
9	Asymptotics of Extremes and Order Statistics	323
9.1	Central-Order Statistics	323
9.1.1	Single-Order Statistic	323
9.1.2	Two Statistical Applications	325
9.1.3	Several Order Statistics	326
9.2	Extremes	328
9.2.1	Easily Applicable Limit Theorems	328
9.2.2	The Convergence of Types Theorem	332
9.3	* Fisher–Tippett Family and Putting it Together	333
	References	338
10	Markov Chains and Applications	339
10.1	Notation and Basic Definitions	340
10.2	Examples and Various Applications as a Model	340
10.3	Chapman–Kolmogorov Equation	345
10.4	Communicating Classes	349
10.5	Gambler’s Ruin	352
10.6	First Passage, Recurrence, and Transience	354
10.7	Long Run Evolution and Stationary Distributions	359
	References	374
11	Random Walks	375
11.1	Random Walk on the Cubic Lattice	375
11.1.1	Some Distribution Theory	378
11.1.2	Recurrence and Transience	379
11.1.3	* Pólya’s Formula for the Return Probability	382
11.2	First Passage Time and Arc Sine Law	383
11.3	The Local Time	387
11.4	Practically Useful Generalizations	389
11.5	Wald’s Identity	390
11.6	Fate of a Random Walk	392

- 11.7 Chung–Fuchs Theorem394
- 11.8 Six Important Inequalities396
- References.....400

- 12 Brownian Motion and Gaussian Processes401**
 - 12.1 Preview of Connections to the Random Walk402
 - 12.2 Basic Definitions403
 - 12.2.1 Condition for a Gaussian Process to be Markov406
 - 12.2.2 * Explicit Construction of Brownian Motion407
 - 12.3 Basic Distributional Properties408
 - 12.3.1 Reflection Principle and Extremes410
 - 12.3.2 Path Properties and Behavior Near Zero and Infinity412
 - 12.3.3 * Fractal Nature of Level Sets415
 - 12.4 The Dirichlet Problem and Boundary Crossing Probabilities416
 - 12.4.1 Recurrence and Transience.....418
 - 12.5 The Local Time of Brownian Motion419
 - 12.6 Invariance Principle and Statistical Applications421
 - 12.7 Strong Invariance Principle and the KMT Theorem425
 - 12.8 Brownian Motion with Drift and Ornstein–Uhlenbeck Process.....427
 - 12.8.1 Negative Drift and Density of Maximum.....427
 - 12.8.2 * Transition Density and the Heat Equation428
 - 12.8.3 * The Ornstein–Uhlenbeck Process429
 - References.....435

- 13 Poisson Processes and Applications.....437**
 - 13.1 Notation.....438
 - 13.2 Defining a Homogeneous Poisson Process.....439
 - 13.3 Important Properties and Uses as a Statistical Model440
 - 13.4 * Linear Poisson Process and Brownian Motion: A Connection ...448
 - 13.5 Higher-Dimensional Poisson Point Processes450
 - 13.5.1 The Mapping Theorem452
 - 13.6 One-Dimensional Nonhomogeneous Processes453
 - 13.7 * Campbell’s Theorem and Shot Noise456
 - 13.7.1 Poisson Process and Stable Laws458
 - References.....462

- 14 Discrete Time Martingales and Concentration Inequalities463**
 - 14.1 Illustrative Examples and Applications in Statistics463
 - 14.2 Stopping Times and Optional Stopping468
 - 14.2.1 Stopping Times469
 - 14.2.2 Optional Stopping470
 - 14.2.3 Sufficient Conditions for Optional Stopping Theorem472
 - 14.2.4 Applications of Optional Stopping474

14.3	Martingale and Concentration Inequalities.....	477
14.3.1	Maximal Inequality.....	477
14.3.2	* Inequalities of Burkholder, Davis, and Gundy.....	480
14.3.3	Inequalities of Hoeffding and Azuma.....	483
14.3.4	* Inequalities of McDiarmid and Devroye.....	485
14.3.5	The Upcrossing Inequality.....	488
14.4	Convergence of Martingales.....	490
14.4.1	The Basic Convergence Theorem.....	490
14.4.2	Convergence in L_1 and L_2	493
14.5	* Reverse Martingales and Proof of SLLN.....	494
14.6	Martingale Central Limit Theorem.....	497
	References.....	503
15	Probability Metrics.....	505
15.1	Standard Probability Metrics Useful in Statistics.....	505
15.2	Basic Properties of the Metrics.....	508
15.3	Metric Inequalities.....	515
15.4	Differential Metrics for Parametric Families.....	519
15.4.1	* Fisher Information and Differential Metrics.....	520
15.4.2	* Rao's Geodesic Distances on Distributions.....	522
	References.....	525
16	Empirical Processes and VC Theory.....	527
16.1	Basic Notation and Definitions.....	527
16.2	Classic Asymptotic Properties of the Empirical Process.....	529
16.2.1	Invariance Principle and Statistical Applications.....	531
16.2.2	* Weighted Empirical Process.....	534
16.2.3	The Quantile Process.....	536
16.2.4	Strong Approximations of the Empirical Process.....	537
16.3	Vapnik–Chervonenkis Theory.....	538
16.3.1	Basic Theory.....	538
16.3.2	Concrete Examples.....	540
16.4	CLTs for Empirical Measures and Applications.....	543
16.4.1	Notation and Formulation.....	543
16.4.2	Entropy Bounds and Specific CLTs.....	544
16.4.3	Concrete Examples.....	547
16.5	Maximal Inequalities and Symmetrization.....	547
16.6	* Connection to the Poisson Process.....	551
	References.....	557
17	Large Deviations.....	559
17.1	Large Deviations for Sample Means.....	560
17.1.1	The Cramér–Chernoff Theorem in \mathcal{R}	560
17.1.2	Properties of the Rate Function.....	564
17.1.3	Cramér's Theorem for General Sets.....	566

- 17.2 The Gärtner–Ellis Theorem and Markov Chain Large Deviations567
- 17.3 The t -Statistic570
- 17.4 Lipschitz Functions and Talagrand’s Inequality572
- 17.5 Large Deviations in Continuous Time.....574
 - 17.5.1 * Continuity of a Gaussian Process.....576
 - 17.5.2 * Metric Entropy of T and Tail of the Supremum577
- References.....582

- 18 The Exponential Family and Statistical Applications583**
 - 18.1 One-Parameter Exponential Family583
 - 18.1.1 Definition and First Examples584
 - 18.2 The Canonical Form and Basic Properties589
 - 18.2.1 Convexity Properties590
 - 18.2.2 Moments and Moment Generating Function591
 - 18.2.3 Closure Properties594
 - 18.3 Multiparameter Exponential Family.....596
 - 18.4 Sufficiency and Completeness600
 - 18.4.1 * Neyman–Fisher Factorization and Basu’s Theorem602
 - 18.4.2 * Applications of Basu’s Theorem to Probability604
 - 18.5 Curved Exponential Family607
 - References.....612

- 19 Simulation and Markov Chain Monte Carlo613**
 - 19.1 The Ordinary Monte Carlo.....615
 - 19.1.1 Basic Theory and Examples.....615
 - 19.1.2 Monte Carlo P -Values622
 - 19.1.3 Rao–Blackwellization623
 - 19.2 Textbook Simulation Techniques624
 - 19.2.1 Quantile Transformation and Accept–Reject.....624
 - 19.2.2 Importance Sampling and Its Asymptotic Properties629
 - 19.2.3 Optimal Importance Sampling Distribution633
 - 19.2.4 Algorithms for Simulating from Common Distributions634
 - 19.3 Markov Chain Monte Carlo.....637
 - 19.3.1 Reversible Markov Chains639
 - 19.3.2 Metropolis Algorithms642
 - 19.4 The Gibbs Sampler.....645
 - 19.5 Convergence of MCMC and Bounds on Errors.....651
 - 19.5.1 Spectral Bounds653
 - 19.5.2 * Dobrushin’s Inequality and Diaconis–Fill–Stroock Bound657
 - 19.5.3 * Drift and Minorization Methods.....659

- 19.6 MCMC on General Spaces662
 - 19.6.1 General Theory and Metropolis Schemes662
 - 19.6.2 Convergence666
 - 19.6.3 Convergence of the Gibbs Sampler670
- 19.7 Practical Convergence Diagnostics673
- References.....686

- 20 Useful Tools for Statistics and Machine Learning689**
 - 20.1 The Bootstrap.....689
 - 20.1.1 Consistency of the Bootstrap.....692
 - 20.1.2 Further Examples696
 - 20.1.3 * Higher-Order Accuracy of the Bootstrap.....699
 - 20.1.4 Bootstrap for Dependent Data701
 - 20.2 The EM Algorithm704
 - 20.2.1 The Algorithm and Examples706
 - 20.2.2 Monotone Ascent and Convergence of EM711
 - 20.2.3 * Modifications of EM714
 - 20.3 Kernels and Classification715
 - 20.3.1 Smoothing by Kernels715
 - 20.3.2 Some Common Kernels in Use717
 - 20.3.3 Kernel Density Estimation719
 - 20.3.4 Kernels for Statistical Classification724
 - 20.3.5 Mercer’s Theorem and Feature Maps.....732
 - References.....744

- A Symbols, Useful Formulas, and Normal Table747**
 - A.1 Glossary of Symbols747
 - A.2 Moments and MGFs of Common Distributions750
 - A.3 Normal Table755

- Author Index.....757**

- Subject Index763**



<http://www.springer.com/978-1-4419-9633-6>

Probability for Statistics and Machine Learning
Fundamentals and Advanced Topics

DasGupta, A.

2011, XX, 784 p., Hardcover

ISBN: 978-1-4419-9633-6