
Contents

<i>Preface</i>	xi
<i>Nomenclature</i>	xiii
1 Introduction	1
1.1 Why Probabilities?	2
1.2 Events and Probabilities	3
1.2.1 Events	3
1.2.2 Assigning Probabilities	5
1.2.3 Joint and Conditional Probabilities	5
1.2.4 Independence	8
1.3 Random Variables	11
1.4 Probability Densities	12
1.5 Expectations	15
1.5.1 Indicator Functions	16
1.5.2 Statistics	17
1.6 Probabilistic Inference	19
1.A Measure for Measure	23
2 Survey of Distributions	25
2.1 Discrete Distributions	26
2.1.1 Binomial Distribution	26
2.1.2 Hypergeometric Distribution	27
2.1.3 Poisson Distribution	28
2.2 Continuous Distributions	30
2.2.1 Normal Distribution	30
2.2.2 Gamma Distribution	31
2.2.3 Beta Distribution	34
2.2.4 Cauchy Distribution	35
2.3 Multivariate Distributions	35
2.3.1 Multinomial Distribution	35
2.3.2 Multivariate Normal Distribution	37
2.3.3 Dirichlet Distribution	37
2.4 Exponential Family	38
	v

vi		<i>Contents</i>
2.A	The Gamma Function	41
2.B	The Beta Function	43
3	Monte Carlo	45
3.1	Random Deviates	46
	3.1.1 Estimating Expectations	46
	3.1.2 Monte Carlo Integration	47
3.2	Uniform Random Deviates	49
3.3	Non-Uniform Random Deviates	51
	3.3.1 Transformation Method	51
	3.3.2 Rejection Sampling	54
	3.3.3 Multivariate Deviates	56
4	Discrete Random Variables	59
4.1	Bernoulli Trials	60
4.2	Binomial Distribution	61
	4.2.1 Statistical Properties of the Binomial Distribution	62
	4.2.2 Cumulants	64
4.3	Beyond the Binomial Distribution	66
	4.3.1 Large n Limit	66
	4.3.2 Poisson Distribution	67
	4.3.3 Multinomial Distribution	68
5	The Normal Distribution	74
5.1	Ubiquitous Normals	75
5.2	Basic Properties	76
	5.2.1 Integrating Gaussians	76
	5.2.2 Moments and Cumulants	78
5.3	Central Limit Theorem	81
5.4	Cumulative Distribution Function of a Normal Distribution	90
5.5	Best of n	92
5.6	Multivariate Normal Distributions	96
5.A	Dirac Delta	103
5.B	Characteristic Function for the Cauchy Distribution	107
6	Handling Experimental Data	110
6.1	Estimating the Error in the Mean	111
	6.1.1 Computing Means with Errors	114
	6.1.2 Bernoulli Trials	114
6.2	Histogram	115
6.3	Significance Tests	116
6.4	Maximum Likelihood Estimate	123
6.A	Deriving the T-Distribution	126
7	Mathematics of Random Variables	132
7.1	Convergence	133

<i>Contents</i>	vii
7.1.1 Laws of Large Numbers	138
7.1.2 Martingales	139
7.2 Inequalities	142
7.2.1 Cauchy–Schwarz Inequality	142
7.2.2 Markov’s Inequality	143
7.2.3 Chebyshev’s Inequality	149
7.2.4 Tail Bounds and Concentration Theorems	152
7.2.5 Chernoff Bounds for Independent Bernoulli Trials	157
7.2.6 Jensen’s Inequality	159
7.2.7 Union Bound	165
7.3 Comparing Distributions	166
7.3.1 Kullback–Leibler Divergence	166
7.3.2 Wasserstein Distance	170
7.A Taylor Expansion	177
7.B Hoeffding’s Bound for Negatively Correlated Random Boolean Variables	179
7.C Constrained Optimisation	182
8 Bayes	187
8.1 Bayesian Statistics	188
8.2 Performing Bayesian Inference	197
8.2.1 Conjugate Priors	198
8.2.2 Uninformative Priors	205
8.2.3 Model Selection	211
8.3 Bayes with Complex Likelihoods	214
8.3.1 Hierarchical Models	214
8.3.2 MAP Solution	215
8.3.3 Variational Approximation	220
8.4 Latent Variable Models	227
8.4.1 Hidden Markov Models	235
8.4.2 Variational Auto-Encoders	240
8.5 Machine Learning	243
8.5.1 Naive Bayes Classifier	244
8.5.2 Graphical Models	245
8.A Bertrand’s Paradox	256
9 Entropy	258
9.1 Shannon Entropy	259
9.1.1 Information Theory	259
9.1.2 Properties of Entropy	265
9.2 Applications	267
9.2.1 Mutual Information	267
9.2.2 Maximum Entropy	267
9.2.3 The Second Law of Thermodynamics	272
9.3 Beyond Information Theory	276
9.3.1 Kolmogorov Complexity	276
9.3.2 Minimum Description Length	277
9.3.3 Fisher Information	285
9.A Functionals	290

viii		<i>Contents</i>
10	Collective Behaviour	293
10.1	Random Walk	294
10.2	Branching Processes	295
10.3	Percolation	297
10.4	Ising Model	299
10.5	Self-Organised Criticality	302
10.6	Disordered Systems	304
11	Markov Chains	308
11.1	Markov Chains	309
	11.1.1 Markov Chains and Stochastic Matrices	310
	11.1.2 Properties of Stochastic Matrices	313
	11.1.3 Ergodic Stochastic Matrices	316
11.2	Markov Chain Monte Carlo	317
	11.2.1 Detailed Balance	318
	11.2.2 Bayes and MCMC	323
	11.2.3 Convergence of MCMC	326
	11.2.4 Hybrid Monte Carlo	329
11.3	Kalman and Particle Filtering	330
	11.3.1 The Kalman Filter	332
	11.3.2 Particle Filtering	338
	11.3.3 Approximate Bayesian Computation	344
11.A	Eigenvalues and Eigenvectors of General Square Matrices	347
12	Stochastic Processes	349
12.1	Stochastic Processes	350
	12.1.1 What Are Stochastic Processes?	350
	12.1.2 Gaussian Processes	351
	12.1.3 Markov Processes	356
12.2	Diffusion Processes	359
	12.2.1 Brownian Motion	359
	12.2.2 Stochastic Differential Equations	361
	12.2.3 Fokker–Planck Equation	370
	12.2.4 Stationary Distribution of Stochastic Processes	374
	12.2.5 Pricing Options	377
12.3	Point Processes	380
	12.3.1 Poisson Processes	380
	12.3.2 Poisson Processes in One Dimension	383
	12.3.3 Chemical Reactions	385
A	Answers to Exercises	391
A.1	Answers to Chapter 1. Introduction	392
A.2	Answers to Chapter 2. Survey of Distributions	396
A.3	Answers to Chapter 3. Monte Carlo	399
A.4	Answers to Chapter 4. Discrete Random Variables	401
A.5	Answers to Chapter 5. The Normal Distribution	407
A.6	Answers to Chapter 6. Handling Experimental Data	411

<i>Contents</i>	ix
A.7 Answers to Chapter 7. Mathematics of Random Variables	414
A.8 Answers to Chapter 8. Bayes	421
A.9 Answers to Chapter 9. Entropy	427
A.10 Answers to Chapter 10. Collective Behaviour	430
A.11 Answers to Chapter 11. Markov Chains	432
A.12 Answers to Chapter 12. Stochastic Processes	441
B Probability Distributions	445
Table B1. Discrete Univariate Distributions	446
Table B2. Continuous Univariate Distributions	447
Table B3. Continuous and Discrete Multivariate Distributions	448
<i>Bibliography</i>	449
<i>Index</i>	454