Contents

Preface *page* xv
Suggestions for instructors and self study xix
Acknowledgements xxi

1 Introduction and review of probability 1
1.1 Probability models 1
1.1.1 The sample space of a probability model 3
1.1.2 Assigning probabilities for finite sample spaces 4
1.2 The axioms of probability theory 5
1.2.1 Axioms for events 7
1.2.2 Axioms of probability 8
1.3 Probability review 9
1.3.1 Conditional probabilities and statistical independence 9
1.3.2 Repeated idealized experiments 11
1.3.3 Random variables 12
1.3.4 Multiple random variables and conditional probabilities 14
1.4 Stochastic processes 16
1.4.1 The Bernoulli process 17
1.5 Expectations and more probability review 19
1.5.1 Random variables as functions of other random variables 23
1.5.2 Conditional expectations 25
1.5.3 Typical values of random variables; mean and median 28
1.5.4 Indicator random variables 29
1.5.5 Moment generating functions and other transforms 29
1.6 Basic inequalities 31
1.6.1 The Markov inequality 32
1.6.2 The Chebyshev inequality 32
1.6.3 Chernoff bounds 33
1.7 The laws of large numbers 36
1.7.1 Weak law of large numbers with a finite variance 36
1.7.2 Relative frequency 39
1.7.3 The central limit theorem (CLT) 39
1.7.4 Weak law with an infinite variance 44
1.7.5 Convergence of random variables 45
1.7.6 Convergence with probability 1 48

1.8 Relation of probability models to the real world 51
1.8.1 Relative frequencies in a probability model 52
1.8.2 Relative frequencies in the real world 52
1.8.3 Statistical independence of real-world experiments 55
1.8.4 Limitations of relative frequencies 56
1.8.5 Subjective probability 57
1.9 Summary 57
1.10 Exercises 58

2 Poisson processes 72
2.1 Introduction 72
2.1.1 Arrival processes 72
2.2 Definition and properties of a Poisson process 74
2.2.1 Memoryless property 75
2.2.2 Probability density of S_n and joint density of $S_1, \ldots, S_n$ 78
2.2.3 The probability mass function (PMF) for $N(t)$ 79
2.2.4 Alternative definitions of Poisson processes 80
2.2.5 The Poisson process as a limit of shrinking Bernoulli processes 82
2.3 Combining and splitting Poisson processes 84
2.3.1 Subdividing a Poisson process 86
2.3.2 Examples using independent Poisson processes 87
2.4 Non-homogeneous Poisson processes 89
2.5 Conditional arrival densities and order statistics 92
2.6 Summary 96
2.7 Exercises 97

3 Gaussian random vectors and processes 105
3.1 Introduction 105
3.2 Gaussian random variables 105
3.3 Gaussian random vectors 107
3.3.1 Generating functions of Gaussian random vectors 108
3.3.2 IID normalized Gaussian random vectors 108
3.3.3 Jointly-Gaussian random vectors 109
3.3.4 Joint probability density for Gaussian n-rv s (special case) 112
3.4 Properties of covariance matrices 114
3.4.1 Symmetric matrices 114
3.4.2 Positive definite matrices and covariance matrices 115
3.4.3 Joint probability density for Gaussian n-rv s (general case) 117
3.4.4 Geometry and principal axes for Gaussian densities 118
3.5 Conditional PDFs for Gaussian random vectors 120
3.6 Gaussian processes 124
3.6.1 Stationarity and related concepts 126
3.6.2 Orthonormal expansions 128
3.6.3 Continuous-time Gaussian processes 130
3.6.4 Gaussian sinc processes 132

3.6.5 Filtered Gaussian sinc processes 134
3.6.6 Filtered continuous-time stochastic processes 136
3.6.7 Interpretation of spectral density and covariance 138
3.6.8 White Gaussian noise 139
3.6.9 The Wiener process/Brownian motion 142
3.7 Circularly-symmetric complex random vectors 144
3.7.1 Circular symmetry and complex Gaussian random variables 145
3.7.2 Covariance and pseudo-covariance of complex n-dimensional random vectors 146
3.7.3 Covariance matrices of complex n-dimensional random vectors 148
3.7.4 Linear transformations of $\boldsymbol{W} \sim \mathcal{CN}(0, [I_\ell])$ 149
3.7.5 Linear transformations of $\boldsymbol{Z} \sim \mathcal{CN}(0, [K])$ 150
3.7.6 The PDF of circularly-symmetric Gaussian n-dimensional random vectors 150
3.7.7 Conditional PDFs for circularly-symmetric Gaussian random vectors 153
3.7.8 Circularly-symmetric Gaussian processes 154
3.8 Summary 155
3.9 Exercises 156

4 Finite-state Markov chains 161
4.1 Introduction 161
4.2 Classification of states 163
4.3 The matrix representation 168
4.3.1 Steady state and $[P^n]$ for large n 168
4.3.2 Steady state assuming $[P] > 0$ 171
4.3.3 Ergodic Markov chains 172
4.3.4 Ergodic unichains 173
4.3.5 Arbitrary finite-state Markov chains 175
4.4 The eigenvalues and eigenvectors of stochastic matrices 176
4.4.1 Eigenvalues and eigenvectors for $\mathsf{M} = 2$ states 176
4.4.2 Eigenvalues and eigenvectors for $\mathsf{M} > 2$ states 177
4.5 Markov chains with rewards 180
4.5.1 Expected first-passage times 181
4.5.2 The expected aggregate reward over multiple transitions 185
4.5.3 The expected aggregate reward with an additional final reward 188
4.6 Markov decision theory and dynamic programming 189
4.6.1 Dynamic programming algorithm 190
4.6.2 Optimal stationary policies 194
4.6.3 Policy improvement and the search for optimal stationary policies 197
4.7 Summary 201
4.8 Exercises 202

5 Renewal processes 214
5.1 Introduction 214
5.2 The strong law of large numbers and convergence with probability 1 217
5.2.1 Convergence with probability 1 (WP1) 217
5.2.2 Strong law of large numbers 219
5.3 Strong law for renewal processes 221
5.4 Renewal–reward processes; time averages 226
5.4.1 General renewal–reward processes 230
5.5 Stopping times for repeated experiments 233
5.5.1 Wald's equality 236
5.5.2 Applying Wald's equality to $\mathsf{E}[N(t)]$ 238
5.5.3 Generalized stopping trials, embedded renewals, and G/G/1 queues 239
5.5.4 Little's theorem 242
5.5.5 M/G/1 queues 245
5.6 Expected number of renewals; ensemble averages 249
5.6.1 Laplace transform approach 250
5.6.2 The elementary renewal theorem 251
5.7 Renewal–reward processes; ensemble averages 254
5.7.1 Age and duration for arithmetic processes 255
5.7.2 Joint age and duration: non-arithmetic case 258
5.7.3 Age $Z(t)$ for finite t: non-arithmetic case 259
5.7.4 Age $Z(t)$ as $t \to \infty$: non-arithmetic case 262
5.7.5 Arbitrary renewal–reward functions: non-arithmetic case 264
5.8 Delayed renewal processes 266
5.8.1 Delayed renewal–reward processes 268
5.8.2 Transient behavior of delayed renewal processes 269
5.8.3 The equilibrium process 270
5.9 Summary 270
5.10 Exercises 271

6 Countable-state Markov chains 287
6.1 Introductory examples 287
6.2 First-passage times and recurrent states 289
6.3 Renewal theory applied to Markov chains 294
6.3.1 Renewal theory and positive recurrence 294
6.3.2 Steady state 296
6.3.3 Blackwell's theorem applied to Markov chains 299
6.3.4 Age of an arithmetic renewal process 300
6.4 Birth–death Markov chains 302
6.5 Reversible Markov chains 303
6.6 The M/M/1 sampled-time Markov chain 307
6.7 Branching processes 309
6.8 Round-robin service and processor sharing 312

6.9 Summary 317
6.10 Exercises 319

7 Markov processes with countable-state spaces 324
7.1 Introduction 324
7.1.1 The sampled-time approximation to a Markov process 328
7.2 Steady-state behavior of irreducible Markov processes 329
7.2.1 Renewals on successive entries to a given state 330
7.2.2 The limiting fraction of time in each state 331
7.2.3 Finding $\{p_j(i);\ j \geq 0\}$ in terms of $\{\pi_j;\ j \geq 0\}$ 332
7.2.4 Solving for the steady-state process probabilities directly 335
7.2.5 The sampled-time approximation again 336
7.2.6 Pathological cases 336
7.3 The Kolmogorov differential equations 337
7.4 Uniformization 341
7.5 Birth–death processes 342
7.5.1 The M/M/1 queue again 342
7.5.2 Other birth–death systems 343
7.6 Reversibility for Markov processes 344
7.7 Jackson networks 350
7.7.1 Closed Jackson networks 355
7.8 Semi-Markov processes 357
7.8.1 Example – the M/G/1 queue 360
7.9 Summary 361
7.10 Exercises 363

8 Detection, decisions, and hypothesis testing 375
8.1 Decision criteria and the maximum a posteriori probability (MAP) criterion 376
8.2 Binary MAP detection 379
8.2.1 Sufficient statistics I 381
8.2.2 Binary detection with a one-dimensional observation 381
8.2.3 Binary MAP detection with vector observations 385
8.2.4 Sufficient statistics II 391
8.3 Binary detection with a minimum-cost criterion 395
8.4 The error curve and the Neyman–Pearson rule 396
8.4.1 The Neyman–Pearson detection rule 402
8.4.2 The min–max detection rule 403
8.5 Finitely many hypotheses 403
8.5.1 Sufficient statistics with $\mathsf{M} \geq 2$ hypotheses 407
8.5.2 More general minimum-cost tests 409
8.6 Summary 409
8.7 Exercises 410

9 Random walks, large deviations, and martingales 417
9.1 Introduction 417
9.1.1 Simple random walks 418
9.1.2 Integer-valued random walks 419
9.1.3 Renewal processes as special cases of random walks 419
9.2 The queueing delay in a G/G/1 queue 420
9.3 Threshold crossing probabilities in random walks 423
9.3.1 The Chernoff bound 423
9.3.2 Tilted probabilities 425
9.3.3 Large deviations and compositions 428
9.3.4 Back to threshold crossings 431
9.4 Thresholds, stopping rules, and Wald's identity 433
9.4.1 Wald's identity for two thresholds 434
9.4.2 The relationship of Wald's identity to Wald's equality 435
9.4.3 Zero-mean random walks 435
9.4.4 Exponential bounds on the probability of threshold crossing 436
9.5 Binary hypotheses with IID observations 438
9.5.1 Binary hypotheses with a fixed number of observations 438
9.5.2 Sequential decisions for binary hypotheses 442
9.6 Martingales 444
9.6.1 Simple examples of martingales 444
9.6.2 Scaled branching processes 446
9.6.3 Partial isolation of past and future in martingales 446
9.7 Submartingales and supermartingales 447
9.8 Stopped processes and stopping trials 450
9.8.1 The Wald identity 452
9.9 The Kolmogorov inequalities 453
9.9.1 The SLLN 455
9.9.2 The martingale convergence theorem 456
9.10 A simple model for investments 458
9.10.1 Portfolios with constant fractional allocations 461
9.10.2 Portfolios with time-varying allocations 465
9.11 Markov modulated random walks 468
9.11.1 Generating functions for Markov random walks 469
9.11.2 Stopping trials for martingales relative to a process 471
9.11.3 Markov modulated random walks with thresholds 471
9.12 Summary 472
9.13 Exercises 473

10 Estimation 488
10.1 Introduction 488
10.1.1 The squared-cost function 489
10.1.2 Other cost functions 490
10.2 MMSE estimation for Gaussian random vectors 491

10.2.1 Scalar iterative estimation 494
10.2.2 Scalar Kalman filter 496
10.3 LLSE estimation 498
10.4 Filtered vector signal plus noise 500
10.4.1 Estimate of a single random variable in IID vector noise 502
10.4.2 Estimate of a single random variable in arbitrary vector noise 502
10.4.3 Vector iterative estimation 503
10.4.4 Vector Kalman filter 504
10.5 Estimation for circularly-symmetric Gaussian rv s 505
10.6 The vector space of random variables; orthogonality 507
10.7 MAP estimation and sufficient statistics 512
10.8 Parameter estimation 513
10.8.1 Fisher information and the Cramer–Rao bound 516
10.8.2 Vector observations 518
10.8.3 Information 519
10.9 Summary 521
10.10 Exercises 523

References 528
Index 530