Contents

INTRODUCTION xvii

1. INTRODUCTION: PROBABILITY AND STATISTICS . . 1

2. BASIC CONCEPTS OF PROBABILITY 3
2.1. Randomness 3
2.2. Frequency tables 5
2.3. Probability: intuitive ideas 6

3. THE CALCULUS OF PROBABILITY 9
3.1. Sample space. Events and sets 9
3.2. Axioms of probability for discrete sample spaces . . . 10
3.3. Randomization. Fundamental sampling theorem 11
3.4. Relationships between events 14
3.4.1. Negation 14
3.4.2. Conjunction 16
3.4.3. $P(AB)$ in terms of $P(A)$ and $P(B)$: Disjoint events . . 17
3.4.4. Disjunction 19
3.4.5. Note on notation 21
3.4.6. Conjunctions and disjunctions involving more than two events 22
3.4.7. Implication 24
3.5. Statistical independence 25
3.6. Some applications of the concept of independence . . . 28
3.6.1. Reduction in dimensionality of sample space . . . 28
3.6.2. Axiom on independence-conferring property of randomization 29
3.6.3. Sampling with replacement from a finite population . . 29
3.7. Bernoulli trials: sampling with replacement from a finite population containing only two kinds of articles 30
3.8. Bernoulli trials in other contexts including (as an approximation) sampling without replacement. The Bernoulli process . . . 32
3.9. Non-independent events and conditional probability . . . 33
3.9.1. Conditional probability 33

3.9.2. Conditional probability and a factorization rule for $P(A \text{ and } B)$ 36
3.9.3. Conditional probability and independence 37
3.9.4. Conditional probabilities are probabilities 38
3.9.5. Manipulative properties of conditional probabilities . . 38
3.9.6. Conditional probability statements involving more than two events 39

4. RANDOM VARIABLES AND THEIR REALIZATIONS; PROBABILITY DISTRIBUTIONS 41
4.1. Random variables 41
4.2. Random variables as functions 42
4.3. The probability distribution of integer-valued random variables . 43
4.3.1. Distribution specified by p.d.f. 43
4.3.2. Distribution specified by c.d.f. 45
4.4. Multivariate random variables: independent random variables . 47
4.5. The calculus of random variables: addition 49
4.5.1. Introduction 49
4.5.2. Addition of random variables 52
4.6. Constants regarded as random variables 55
4.7. Transformations of random variables 56

5. A FIRST CATALOGUE OF DISCRETE PROBABILITY DISTRIBUTIONS (INTEGER-VALUED VARIABLES) 59
5.1. The discrete uniform distribution 59
5.2. Distributions arising in connection with the Bernoulli sequence . 62
5.2.1. Bernoulli trial variables: trials with success-parameter p ('p-trials') 62
5.2.2. The binomial distribution with n trials and success parameter p (the binomial (n, p) distribution) 63
5.2.3. The geometric distribution 67
(i) first version 67
(ii) second version of the geometric distribution 69
5.2.4. The negative binomial distribution (the Pascal distribution) 70
5.3. The hypergeometric distribution 71
5.4. The Poisson distribution 74
5.5. The Poisson distribution as approximation to the binomial and to the hypergeometric distribution 76
5.6. Simulation: generating realizations of a discrete random variable 77

6. INDEPENDENT AND NON-INDEPENDENT RANDOM VARIABLES AND THEIR JOINT DISTRIBUTIONS, BIVARIATE AND MULTIVARIATE; MARGINAL AND CONDITIONAL DISTRIBUTIONS 83
6.1. Bivariate distributions 83

6.1.1. Bivariate random variables and their joint p.d.f. . . . 83
6.1.2. Bivariate distributions and their joint c.d.f. . . . 85
6.2. Bivariate frequency tables 87
6.3. Marginal distributions 88
6.3.1. Introduction 88
6.3.2. Marginal distributions in the case of independence . . 89
6.4. An important multivariate distribution: the multinomial and a special case: the trinomial 90
6.4.1. The trinomial distribution 90
6.4.2. Multinomial distributions and their marginals . . . 91
6.5. Conditional probabilities and conditional distributions . . 94
6.6. Bivariate and multivariate probabilities expressed in terms of conditional probabilities, and the special case of independence . 99
6.6.1. Introduction 99
6.6.2. The special case of independence. Distribution of a sample . 101
6.7. Truncated distributions 102
6.8. The 'residual waiting time' property of the geometric distribution . 105
6.9. The interval between successive failures in a Bernoulli process . 106

7. THE ADDITION OF RANDOM VARIABLES: CONVOLUTION 109
7.1. Introduction 109
7.2. Non-negative variables 111
7.3. Variables bounded above and below 112
7.4. Examples of the convolution operation 112
7.4.1. The sum of independent p-binomials is p binomial . . 112
7.4.2. The sum of geometrically distributed variables . . . 114

8. EXPECTATION (EXPECTED VALUE) 115
8.1. Motivation and definition: examples 115
8.2. Expectation of a sum 117
8.3. Expectation of a constant 119
8.4. Expectation of a linear function of random variables . . . 119
8.5. Expectation of a non-linear function of a random variable . . 120
8.6. Expectation of a product of independent variables . . . 123
8.7. Variables that do not have a finite expectation 124
8.8. A minimizing property of $E(X)$ 124
8.9. Conditional expectation 125

9. VARIANCE, COVARIANCE, AND HIGHER MOMENTS . . 129
9.1. Variance: motivation, and definition of sample variance . . 129
9.2. Standard deviation and variance of a random variable ('population values') 131
9.2.1. Definitions and examples 131

9.2.2. Effect of linear transformation 134
9.2.3. Variance of a sum of independent variables . . . 134
9.2.4. Variance of scaled sum of independent variables. . . 134
9.2.5. Variance of a linear function of independent variables. . 135
9.2.6. Coefficient of variation 135
9.3. Random variables that do not have a finite variance . . . 136
9.4. Scaling, standardization and skewness 136
9.5. Chebychev's inequality. Markov's inequality 137
9.6. Covariance. Variance of a linear function of random variables. Uncorrelated variables 139
9.6.1. Definition of covariance 139
9.6.2. Properties of the covariance operator 142
9.6.3. Uncorrelated variables 144
9.6.4. Covariance of one linear function with another: matrix form 146
9.6.5. Variance of a linear function 146
9.7. The variance–covariance inequality 148
9.8. The correlation coefficient 148
9.8.1. Definition and basic properties 148
9.8.2. The correlation coefficient and linear transformations . . 149
9.8.3. The correlation coefficient as a measure of statistical association 150
9.9. Approximations for the expectation, variance and covariance of nonlinear functions 152
9.9.1. Introduction 152
9.9.2. Approximations for $E\{h(X)\}$ 152
9.9.3. Approximations for var $\{h(X)\}$ 154
9.9.4. Approximations for cov $\{g(X_1), h(X_2)\}$ 154
9.10. Shape: skewness and kurtosis 155
9.10.1. Geometrical motivation 155
9.10.2. Coefficient of skewness of linear transformations and sums of random variables 156
9.10.3. Coefficient of skewness of a sum of independent and identically distributed variables 157
9.10.4. Kurtosis 157
9.11. Moments 157
9.12. Conditional variance and conditional moments 159

10. CONTINUOUS DISTRIBUTIONS 161
10.1. Introduction 161
10.1.1. Basic properties of the p.d.f. and the c.d.f. . . . 161
10.1.2. An intuitive interpretation of the graph of a p.d.f. . . 164
10.1.3. The mode of a continuous p.d.f. 165
10.2. Some examples of continuous random variables . . . 166
10.2.1. The continuous uniform distribution 166
10.2.2. *n* Uniformly distributed particles 168
10.2.3. The exponential distribution 170

10.3. The cumulative distribution function and the probability density function of a continuous random variable 171
10.3.1. The cumulative distribution function 171
10.3.2. The probability density function 173
10.3.3. Fractiles and percentiles 176
10.4. Measures of location and scale 177
10.4.1. Expected value and variance of a continuous random variable 177
10.4.2. Other measures of location and scale 180
10.5. Joint distributions: independence 181
10.6. Expected value of a sum 183
10.7. Transformations 185
10.8. Convolutions 188

11. A CATALOGUE OF CONTINUOUS PROBABILITY DISTRIBUTIONS 191
11.1. The continuous uniform distribution 191
11.1.1. Standard and other forms 191
11.1.2. Convolutions of uniformly distributed variables . . 193
11.2. The (negative) exponential distribution 196
11.2.1. Standard and other forms 196
11.2.2. Convolutions and special properties 198
11.3. The gamma family 199
11.3.1. Standard and scaled forms of the p.d.f. of the gamma distribution 199
11.3.2. Relation between gamma and exponential distributions; convolutions. Erlangian distributions 201
11.3.3. Occurrence and applications 202
11.4. The Normal (Gaussian) distribution 203
11.4.1. Standard form 203
11.4.2. Occurrence and applications: Introduction to the central limit theorem 204
11.4.3. The Normal family 206
11.4.4. The value of a certain definite integral 208
11.4.5. Convolutions of Normal distribution 208
11.4.6. Numerical values of Normal probabilities. Use of tables . 209
11.4.7. The Normal distribution as an approximation to the binomial 210
11.4.8. Normal probability graph paper ('arithmetic probability paper') 212
11.4.9. The Normal distribution and errors of measurement . . 214
11.4.10. The Normal distribution and errors of computation . . 214
11.4.11. The Chi-squared distribution 216
11.5. The log-Normal distribution 217
11.6. The Beta distribution 219
11.6.1. Standard form 219

11.6.2. Linearly transformed Beta distribution 220
11.6.3. Relation of Beta distribution with other distributions . 221
11.7. The Cauchy distribution 221
11.8. The Pareto distribution 222
11.9. The Weibull distribution 224
11.10. The Logistic distribution 225
11.11. The Powered Gamma distribution (Kritskii-Menkel) . . 226
11.12. The Pearson System 228
11.13. Simulation: generating realizations of continuous random variables 229

12. GENERATING FUNCTIONS 231
12.0. Introduction 231
12.1. The probability generating function (p.g.f.) 231
12.1.1. Definition of the p.g.f. 231
12.1.2. The p.g.f. as expectation 232
12.1.3. Convolutions 233
12.1.4. Tail probabilities 234
12.1.5. Expectation obtained from the p.g.f. 235
12.1.6. Variance obtained from the p.g.f. 235
12.1.7. Extended definition of the p.g.f.: variables taking positive and negative values. The p.g.f. of $R - T$ 236
12.1.8. Probability generating function of a discrete random variable whose values are not necessarily integers 238
12.1.9. The p.g.f. of a multivariate distribution 240
12.2. The moment generating function (m.g.f.) 242
12.2.1. Definition 242
12.2.2. Major properties 243
12.2.3. Bivariate and multivariate m.g.f.'s 244
12.3. The Laplace Transform (L.T.) 246
12.4. The characteristic function 246
12.4.1. Introduction and definition 246
12.4.2. Linear transformations 248
12.4.3. Principal properties 248
12.4.4. Bivariate and multivariate c.f.'s 249
12.5. The second characteristic function (cumulant generating function) 250
12.5.1. Definition 250
12.5.2. Another version of the c.g.f. 250
12.5.3. Relation between cumulants and moments . . . 251
12.5.4. Effect of shift of origin 252
12.5.5. Cumulants of a sum of random variables . . . 253
12.6. The factorial moment generating function (f.m.g.f.) . . . 253

13. MULTIVARIATE CONTINUOUS DISTRIBUTIONS . . 255
13.1. Joint, marginal and conditional distributions; independence. . . 255
13.1.1. Introduction: joint distributions 255

13.1.2. Marginal distributions 257
13.1.3. The moments of a multivariate distribution . . . 259
13.1.4. Conditional distributions 261
13.1.5. Regression 262
13.1.6. Independence: the distribution of a sample . . . 263
13.1.7. Multivariate moment generating functions . . . 264
13.2. Multivariate transformations 265
13.2.1. Introduction 265
13.2.2. Invertible transformations 265
13.3. Linear operations on jointly distributed random variables . . 267
13.3.1. Vector variables, expectation vector, dispersion matrix . 267
13.3.2. Linear transforms 269
13.3.3. More on variances and covariances 270
(i) The cross-covariance matrix 270
(ii) Dispersion matrix of a sum 270
13.3.4. The bivariate Normal distribution 270
13.4. The multivariate Normal (MVN) distribution 272
13.4.1. The MVN density function 272
13.4.2. The MVN distribution and independence . . . 274
13.4.3. Moment generating function of the MVN distribution . 275
13.4.4. The marginal distribution of the MVN 276
(i) Normal marginals do not imply a Multivariate Normal distribution 279
(ii) Marginal distribution of the MVN derived in terms of the moment generating function 279
13.4.5. The conditional distributions of the MVN . . . 280
13.4.6. Summary of properties of the bivariate and trivariate Normal distribution 281
(i) Bivariate Normal 281
(ii) Trivariate Normal 282
(iii) Partial Correlation coefficient 284
13.4.7. The joint distribution of a set of linear functions of MVN variables 284
Independence and Orthogonal transformations in the MVN 285
13.4.8. The distribution of a certain quadratic form . . . 286
13.5. An example of a non-Normal multivariate distribution. The Dirichlet distribution: a multivariate beta 286
13.6. Linear regression ('multiple regression') and the multiple correlation coefficient 287

14. MISCELLANEOUS TOPICS ON DISTRIBUTIONS . . . 289
14.1. Distributions that are part continuous, part discrete . . . 289
14.2. Mixtures of distributions 290
14.3. Compound distributions: the sum of a random number of random variables 292
14.3.1. Distribution of S_N 292

14.3.2. Mean and variance of S_N 293
14.3.3. Use of the Probability Generating Function for compound distributions 294
14.3.4. Moment generating function of S_N 295

15. ORDER STATISTICS 297
15.1. Introduction 297
15.2. The distribution of individual order statistics when the underlying distribution is continuous 299
(i) Largest value 299
(ii) Smallest value 300
(iii) A general value 301
(iv) Order statistics of the exponential and the logistic distributions 303
15.3. The Order statistics of discrete variables 303
15.4. The Joint distribution of several order statistics in samples from a continuous distribution 306
15.5. Median, range and extreme values for continuous distributions . 307
15.6. 'Extreme-value' distributions 310

16. MISCELLANEOUS TOPICS INVOLVING CONDITIONAL PROBABILITY 313
16.1. The conditioned variable $X \mid y$ 313
16.2. Theorem of Total Probability. 314
16.3. The conditional expectation theorem 316
16.4. Bayes' Theorem 319
16.4.1. Introduction: Bayes' Theorem with binomial sampling . 319
16.4.2. Bayes' Theorem with general sampling from a population selected from a finite or countable set 321
16.4.3. Bayes' Theorem for a continuous parameter 322

17. LAWS OF LARGE NUMBERS. STABLE DISTRIBUTIONS . 325
17.1. The (weak) law of large numbers 325
17.2. Strong law of large numbers 326
17.3. Central Limit Theorem 327
17.4. Stable distributions 328

18. STOCHASTIC PROCESSES: INTRODUCTION AND EXAMPLES 331
18.0. Introduction: The state space and parameter space 331
18.0.1. Univariate random variables 331
18.0.2. Multivariate random variables 331
18.0.3. Discrete-state or continuous-state stochastic processes with discrete parameter 332
18.0.4. Continuous-parameter stochastic processes 334
18.0.5. Other classifications of stochastic processes 334
18.0.6. Dimensionality 335

18.0.7. Sequence of sums of random variables 335
18.1. The branching process 336
18.1.1. General description 336
18.1.2. Mean size of the *n*th generation 338
18.1.3. Probabilities of ultimate extinction and of indefinite survival 338
18.1.4. Population size after a large number of generations . . 340
18.1.5. Effect of starting from more than a single initial particle . 340
18.1.6. Applications to survival of family names 340
18.2. The Gambler's Ruin 341
18.2.1. Introduction 341
18.2.2. The probability of ruin 342
18.2.3. Expected duration of play 343
18.2.4. Distribution of duration of play 344
18.3. The Random Walk 346
18.3.1. Introduction: the state space and the parameter space . 346
18.3.2. Simple random walk. First passages. Absorbing barriers . 346
(i) Unrestricted simple random walk 346
(ii) First passages 349
(iii) First passages by the method of unit descents . . . 350
(iv) An improper random variable 352
(v) First passage time by the method of images: (geometric version) 352
(vi) The method of images (analytic version) 354
(vii) First returns in an unrestricted random walk . . . 355
(viii) Time to absorption, by the method of images . . . 357
(ix) First passages when there are two absorbing barriers. . 358
18.4. More general one-dimensional random walks. Wald's identity . 362
18.5. Diffusion 370

19. PROCESSES IN DISCRETE TIME: MARKOV CHAINS . . 373
19.1. Generalities on stochastic processes 373
19.2. General dependence, independence and Markovian dependence . 374
19.3. Discrete-state Markov Chains 375
19.4. The s-step transition probability 382
19.5. The distribution of X_t 385
19.6. The equilibrium distribution, the stationary distribution and ergodic chains 387
19.7. Classification of states and chains 388
19.7.1. Communicating states 388
19.7.2. Irreducible chains 389
19.7.3. Periodic and aperiodic chains 389
19.7.4. Irreducible chains and equilibrium distributions . . 390
19.7.5. Further properties of Markov Chains. Closed sets and absorbing states: transient states 392
19.8. The correlation function of a Markov Chain 393
19.9. Markov Chains of the second (or higher) order 397

19.10. Continuous-state Markov Chains 400

20. CONTINUOUS-TIME PROCESSES 403
20.1. The Poisson process 403
20.1.1. General definition: the one-dimensional process . . 403
20.1.2. Waiting times from a fixed origin 406
20.1.3. Interval between consecutive events 408
20.1.4. Superposition of independent Poisson processes . . 408
20.1.5. Clustering 409
20.1.6. Non-homogeneous Poisson processes 409
20.1.7. The homogeneous Poisson process in two or more dimensions 410
20.1.8. Nearest neighbours 410
20.1.9. Spatio-temporal Poisson fields 411
20.1.10. Projections of multi-dimensional Poisson processes . . 411
20.2. Birth processes 411
20.3. The two-state process 413
20.4. The Birth-and-Death process 414
20.4.1. General description 414
20.4.2. Statistical equilibrium 415
20.4.3. Special case: 'Linear growth' 416
20.5. The simple queue 417
20.6. Markov Chains with continuous time parameters . . . 418

21. RENEWALS 421
21.1. Integer-valued random variables. Introduction, Definition, Examples 421
21.2. Fundamental Theorems on recurrence-time distributions . . 423
21.3. Number of occurrences 427
21.3. The Markov Chain as a renewal process 428
21.5. Modified renewal processes 431
21.6. Renewal Theory in continuous time. 431

22. COMPLEMENTS ON STOCHASTIC PROCESSES: SECOND ORDER PROCESSES 433
22.1. Definitions 433
22.2. The covariance function 434
22.2.1. Definitionss 434
22.2.2. Properties of the auto-covariance function . . . 435
22.3. Spectral density 436
22.4. Introduction to stochastic analysis 436
22.4.1. Convergence 436
22.4.2. Continuity 437
22.4.3. Differentiability 438

BIBLIOGRAPHY 439

INDEX 443