

Contents

Preface	ix
1 Probability Models	1
1.1 Probability: A Measure of Uncertainty	1
1.1.1 Why Do We Need Probability Theory?	2
1.2 Probability Models	4
1.2.1 Venn Diagrams and Subsets	7
1.3 Properties of Probability Models	10
1.4 Uniform Probability on Finite Spaces	14
1.4.1 Combinatorial Principles	15
1.5 Conditional Probability and Independence	20
1.5.1 Conditional Probability	20
1.5.2 Independence of Events	23
1.6 Continuity of P	28
1.7 Further Proofs (Advanced)	31
2 Random Variables and Distributions	33
2.1 Random Variables	34
2.2 Distributions of Random Variables	38
2.3 Discrete Distributions	41
2.3.1 Important Discrete Distributions	42
2.4 Continuous Distributions	51
2.4.1 Important Absolutely Continuous Distributions	53
2.5 Cumulative Distribution Functions	62
2.5.1 Properties of Distribution Functions	63
2.5.2 Cdfs of Discrete Distributions	64
2.5.3 Cdfs of Absolutely Continuous Distributions	65
2.5.4 Mixture Distributions	68
2.5.5 Distributions Neither Discrete Nor Continuous (Advanced)	70
2.6 One-Dimensional Change of Variable	74
2.6.1 The Discrete Case	75
2.6.2 The Continuous Case	75
2.7 Joint Distributions	79
2.7.1 Joint Cumulative Distribution Functions	80

2.7.2	Marginal Distributions	81
2.7.3	Joint Probability Functions	83
2.7.4	Joint Density Functions	85
2.8	Conditioning and Independence	93
2.8.1	Conditioning on Discrete Random Variables	94
2.8.2	Conditioning on Continuous Random Variables	95
2.8.3	Independence of Random Variables	97
2.8.4	Order Statistics	103
2.9	Multidimensional Change of Variable	109
2.9.1	The Discrete Case	109
2.9.2	The Continuous Case (Advanced)	110
2.9.3	Convolution	113
2.10	Simulating Probability Distributions	116
2.10.1	Simulating Discrete Distributions	117
2.10.2	Simulating Continuous Distributions	119
2.11	Further Proofs (Advanced)	125
3	Expectation	129
3.1	The Discrete Case	129
3.2	The Absolutely Continuous Case	141
3.3	Variance, Covariance, and Correlation	149
3.4	Generating Functions	162
3.4.1	Characteristic Functions (Advanced)	169
3.5	Conditional Expectation	173
3.5.1	Discrete Case	173
3.5.2	Absolutely Continuous Case	176
3.5.3	Double Expectations	177
3.5.4	Conditional Variance (Advanced)	179
3.6	Inequalities	184
3.6.1	Jensen's Inequality (Advanced)	187
3.7	General Expectations (Advanced)	191
3.8	Further Proofs (Advanced)	194
4	Sampling Distributions and Limits	199
4.1	Sampling Distributions	200
4.2	Convergence in Probability	203
4.2.1	The Weak Law of Large Numbers	205
4.3	Convergence with Probability 1	208
4.3.1	The Strong Law of Large Numbers	211
4.4	Convergence in Distribution	213
4.4.1	The Central Limit Theorem	215
4.4.2	The Central Limit Theorem and Assessing Error	220
4.5	Monte Carlo Approximations	224
4.6	Normal Distribution Theory	234
4.6.1	The Chi-Squared Distribution	236
4.6.2	The t Distribution	239

4.6.3	The F Distribution	240
4.7	Further Proofs (Advanced)	246
5	Statistical Inference	253
5.1	Why Do We Need Statistics?	254
5.2	Inference Using a Probability Model	258
5.3	Statistical Models	262
5.4	Data Collection	269
5.4.1	Finite Populations	270
5.4.2	Simple Random Sampling	271
5.4.3	Histograms	274
5.4.4	Survey Sampling	276
5.5	Some Basic Inferences	282
5.5.1	Descriptive Statistics	282
5.5.2	Plotting Data	287
5.5.3	Types of Inference	289
6	Likelihood Inference	297
6.1	The Likelihood Function	297
6.1.1	Sufficient Statistics	302
6.2	Maximum Likelihood Estimation	308
6.2.1	Computation of the MLE	310
6.2.2	The Multidimensional Case (Advanced)	316
6.3	Inferences Based on the MLE	320
6.3.1	Standard Errors, Bias, and Consistency	321
6.3.2	Confidence Intervals	326
6.3.3	Testing Hypotheses and P-Values	332
6.3.4	Inferences for the Variance	338
6.3.5	Sample-Size Calculations: Confidence Intervals	340
6.3.6	Sample-Size Calculations: Power	341
6.4	Distribution-Free Methods	349
6.4.1	Method of Moments	349
6.4.2	Bootstrapping	351
6.4.3	The Sign Statistic and Inferences about Quantiles	357
6.5	Large Sample Behavior of the MLE (Advanced)	364
7	Bayesian Inference	373
7.1	The Prior and Posterior Distributions	374
7.2	Inferences Based on the Posterior	384
7.2.1	Estimation	387
7.2.2	Credible Intervals	391
7.2.3	Hypothesis Testing and Bayes Factors	394
7.2.4	Prediction	400
7.3	Bayesian Computations	407
7.3.1	Asymptotic Normality of the Posterior	407
7.3.2	Sampling from the Posterior	407

7.3.3	Sampling from the Posterior Using Gibbs Sampling (Advanced)	413
7.4	Choosing Priors	421
7.4.1	Conjugate Priors	422
7.4.2	Elicitation	422
7.4.3	Empirical Bayes	423
7.4.4	Hierarchical Bayes	424
7.4.5	Improper Priors and Noninformativity	425
7.5	Further Proofs (Advanced)	430
7.5.1	Derivation of the Posterior Distribution for the Location-Scale Normal Model	430
7.5.2	Derivation of $J(\theta(\psi_0, \lambda))$ for the Location-Scale Normal	431
8	Optimal Inferences	433
8.1	Optimal Unbiased Estimation	434
8.1.1	The Rao–Blackwell Theorem and Rao–Blackwellization	435
8.1.2	Completeness and the Lehmann–Scheffé Theorem	438
8.1.3	The Cramer–Rao Inequality (Advanced)	440
8.2	Optimal Hypothesis Testing	446
8.2.1	The Power Function of a Test	446
8.2.2	Type I and Type II Errors	447
8.2.3	Rejection Regions and Test Functions	448
8.2.4	The Neyman–Pearson Theorem	449
8.2.5	Likelihood Ratio Tests (Advanced)	455
8.3	Optimal Bayesian Inferences	460
8.4	Decision Theory (Advanced)	464
8.5	Further Proofs (Advanced)	473
9	Model Checking	479
9.1	Checking the Sampling Model	479
9.1.1	Residual and Probability Plots	486
9.1.2	The Chi-Squared Goodness of Fit Test	491
9.1.3	Prediction and Cross-Validation	495
9.1.4	What Do We Do When a Model Fails?	496
9.2	Checking for Prior–Data Conflict	502
9.3	The Problem with Multiple Checks	509
10	Relationships Among Variables	511
10.1	Related Variables	512
10.1.1	The Definition of Relationship	512
10.1.2	Cause–Effect Relationships and Experiments	516
10.1.3	Design of Experiments	519
10.2	Categorical Response and Predictors	527
10.2.1	Random Predictor	527
10.2.2	Deterministic Predictor	530
10.2.3	Bayesian Formulation	533

10.3	Quantitative Response and Predictors	538
10.3.1	The Method of Least Squares	538
10.3.2	The Simple Linear Regression Model	540
10.3.3	Bayesian Simple Linear Model (Advanced)	554
10.3.4	The Multiple Linear Regression Model (Advanced)	558
10.4	Quantitative Response and Categorical Predictors	577
10.4.1	One Categorical Predictor (One-Way ANOVA)	577
10.4.2	Repeated Measures (Paired Comparisons)	584
10.4.3	Two Categorical Predictors (Two-Way ANOVA)	586
10.4.4	Randomized Blocks	594
10.4.5	One Categorical and One Quantitative Predictor	594
10.5	Categorical Response and Quantitative Predictors	602
10.6	Further Proofs (Advanced)	607
11	Advanced Topic — Stochastic Processes	615
11.1	Simple Random Walk	615
11.1.1	The Distribution of the Fortune	616
11.1.2	The Gambler's Ruin Problem	618
11.2	Markov Chains	623
11.2.1	Examples of Markov Chains	624
11.2.2	Computing with Markov Chains	626
11.2.3	Stationary Distributions	629
11.2.4	Markov Chain Limit Theorem	633
11.3	Markov Chain Monte Carlo	641
11.3.1	The Metropolis–Hastings Algorithm	644
11.3.2	The Gibbs Sampler	647
11.4	Martingales	650
11.4.1	Definition of a Martingale	650
11.4.2	Expected Values	651
11.4.3	Stopping Times	652
11.5	Brownian Motion	657
11.5.1	Faster and Faster Random Walks	657
11.5.2	Brownian Motion as a Limit	659
11.5.3	Diffusions and Stock Prices	661
11.6	Poisson Processes	665
11.7	Further Proofs	668
	Appendices	675
A	Mathematical Background	675
A.1	Derivatives	675
A.2	Integrals	676
A.3	Infinite Series	677
A.4	Matrix Multiplication	678
A.5	Partial Derivatives	678

A.6	Multivariable Integrals	679
B	Computations	683
B.1	Using R	683
B.2	Using Minitab	699
C	Common Distributions	705
C.1	Discrete Distributions	705
C.2	Absolutely Continuous Distributions	706
D	Tables	709
D.1	Random Numbers	710
D.2	Standard Normal Cdf	712
D.3	Chi-Squared Distribution Quantiles	713
D.4	t Distribution Quantiles	714
D.5	F Distribution Quantiles	715
D.6	Binomial Distribution Probabilities	724
E	Answers to Odd-Numbered Exercises	729
	Index	750