

Probability and Statistics: The Science of Uncertainty

Michael J. Evans and Jeffrey S. Rosenthal University of Toronto

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	204
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 Inferences	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	Asymptotics for 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 Via 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
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	490
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

CONTENTS

i

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	751

Preface

This book is an introductory text on probability and statistics, targeting students who have studied one year of calculus at the university level and are seeking an introduction to probability and statistics with mathematical content. Where possible, we provide mathematical details, and it is expected that students are seeking to gain some mastery over these, as well as to learn how to conduct data analyses. All the usual methodologies covered in a typical introductory course are introduced, as well as some of the theory that serves as their justification.

The text can be used with or without a statistical computer package. It is our opinion that students should see the importance of various computational techniques in applications, and the book attempts to do this. Accordingly, we feel that computational aspects of the subject, such as Monte Carlo, should be covered, even if a statistical package is not used. Almost any statistical package is suitable. A **Computations** appendix provides an introduction to the R language. This covers all aspects of the language needed to do the computations in the text. Furthermore, we have provided the R code for any of the more complicated computations. Students can use these examples as templates for problems that involve such computations, e.g., using Gibbs sampling. Also, we have provided, in a separate section of this appendix, Minitab code for those computations that are slightly involved, e.g., Gibbs sampling. No programming experience is required of students to do the problems.

We have organized the exercises in the book into groups, as an aid to users. **Exercises** are suitable for all students and offer practice in applying the concepts discussed in a particular section. **Problems** require greater understanding, and a student can expect to spend more thinking time on these. If a problem is marked (MV), then it will require some facility with multivariable calculus beyond the first calculus course, although these problems are not necessarily hard. **Challenges** are problems that most students will find difficult; these are only for students who have no trouble with the **Exercises** and the **Problems**. There are also **Computer Exercises** and **Computer Problems**, where it is expected that students will make use of a statistical package in deriving solutions.

We have included a number of **Discussion Topics** designed to promote critical thinking in students. Throughout the book, we try to point students beyond the mastery of technicalities to think of the subject in a larger frame of reference. It is important that students acquire a sound mathematical foundation in the basic techniques of probability and statistics, which we believe this book will help students accomplish. Ultimately, however, these subjects are applied in real-world contexts, so it is equally important that students understand how to go about their application and understand what issues arise. Often, there are no right answers to **Discussion Topics**; their purpose is to get a

student thinking about the subject matter. If these were to be used for evaluation, then they would be answered in essay format and graded on the maturity the student showed with respect to the issues involved. **Discussion Topics** are probably most suitable for smaller classes, but these will also benefit students who simply read them over and contemplate their relevance.

Some sections of the book are labelled **Advanced**. This material is aimed at students who are more mathematically mature (for example, they are taking, or have taken, a second course in calculus). All the **Advanced** material can be skipped, with no loss of continuity, by an instructor who wishes to do so. In particular, the final chapter of the text is labelled **Advanced** and would only be taught in a high-level introductory course aimed at specialists. Also, many proofs appear in the final section of many chapters, labelled **Further Proofs (Advanced)**. An instructor can choose which (if any) of these proofs they wish to present to their students.

As such, we feel that the material in the text is presented in a flexible way that allows the instructor to find an appropriate level for the students they are teaching. A **Mathematical Background** appendix reviews some mathematical concepts, from a first course in calculus, in case students could use a refresher, as well as brief introductions to partial derivatives, double integrals, etc.

Chapter 1 introduces the probability model and provides motivation for the study of probability. The basic properties of a probability measure are developed.

Chapter 2 deals with discrete, continuous, joint distributions, and the effects of a change of variable. It also introduces the topic of simulating from a probability distribution. The multivariate change of variable is developed in an Advanced section.

Chapter 3 introduces expectation. The probability-generating function is discussed, as are the moments and the moment-generating function of a random variable. This chapter develops some of the major inequalities used in probability. A section on characteristic functions is included as an Advanced topic.

Chapter 4 deals with sampling distributions and limits. Convergence in probability, convergence with probability 1, the weak and strong laws of large numbers, convergence in distribution, and the central limit theorem are all introduced, along with various applications such as Monte Carlo. The normal distribution theory, necessary for many statistical applications, is also dealt with here.

As mentioned, Chapters 1 through 4 include material on Monte Carlo techniques. Simulation is a key aspect of the application of probability theory, and it is our view that its teaching should be integrated with the theory right from the start. This reveals the power of probability to solve real-world problems and helps convince students that it is far more than just an interesting mathematical theory. No practitioner divorces himself from the theory when using the computer for computations or vice versa. We believe this is a more modern way of teaching the subject. This material can be skipped, however, if an instructor believes otherwise or feels there is not enough time to cover it effectively.

Chapter 5 is an introduction to statistical inference. For the most part, this is concerned with laying the groundwork for the development of more formal methodology in later chapters. So practical issues — such as proper data collection, presenting data via graphical techniques, and informal inference methods like descriptive statistics — are discussed here.

Chapter 6 deals with many of the standard methods of inference for one-sample problems. The theoretical justification for these methods is developed primarily through the likelihood function, but the treatment is still fairly informal. Basic methods of inference, such as the standard error of an estimate, confidence intervals, and P-values, are introduced. There is also a section devoted to distribution-free (nonparametric) methods like the bootstrap.

Chapter 7 involves many of the same problems discussed in Chapter 6, but now from a Bayesian perspective. The point of view adopted here is not that Bayesian methods are better or, for that matter, worse than those of Chapter 6. Rather, we take the view that Bayesian methods arise naturally when the statistician adds another ingredient — the prior — to the model. The appropriateness of this, or the sampling model for the data, is resolved through the model-checking methods of Chapter 9. It is not our intention to have students adopt a particular philosophy. Rather, the text introduces students to a broad spectrum of statistical thinking.

Subsequent chapters deal with both frequentist and Bayesian approaches to the various problems discussed. The Bayesian material is in clearly labelled sections and can be skipped with no loss of continuity, if so desired. It has become apparent in recent years, however, that Bayesian methodology is widely used in applications. As such, we feel that it is important for students to be exposed to this, as well as to the frequentist approaches, early in their statistical education.

Chapter 8 deals with the traditional optimality justifications offered for some statistical inferences. In particular, some aspects of optimal unbiased estimation and the Neyman–Pearson theorem are discussed. There is also a brief introduction to decision theory. This chapter is more formal and mathematical than Chapters 5, 6, and 7, and it can be skipped, with no loss of continuity, if an instructor wants to emphasize methods and applications.

Chapter 9 is on model checking. We placed model checking in a separate chapter to emphasize its importance in applications. In practice, model checking is the way statisticians justify the choices they make in selecting the ingredients of a statistical problem. While these choices are inherently subjective, the methods of this chapter provide checks to make sure that the choices made are sensible in light of the objective observed data.

Chapter 10 is concerned with the statistical analysis of relationships among variables. This includes material on simple linear and multiple regression, ANOVA, the design of experiments, and contingency tables. The emphasis in this chapter is on applications.

Chapter 11 is concerned with stochastic processes. In particular, Markov chains and Markov chain Monte Carlo are covered in this chapter, as are Brownian motion and its relevance to finance. Fairly sophisticated topics are introduced, but the treatment is entirely elementary. Chapter 11 depends only on the material in Chapters 1 through 4.

A one-semester course on probability would cover Chapters 1–4 and perhaps some of Chapter 11. A one-semester, follow-up course on statistics would cover Chapters 5–7 and 9–10. Chapter 8 is not necessary, but some parts, such as the theory of unbiased estimation and optimal testing, are suitable for a more theoretical course.

A basic two-semester course in probability and statistics would cover Chapters 1–6 and 9–10. Such a course covers all the traditional topics, including basic probability

theory, basic statistical inference concepts, and the usual introductory applied statistics topics. To cover the entire book would take three semesters, which could be organized in a variety of ways.

The Advanced sections can be skipped or included, depending on the level of the students, with no loss of continuity. A similar approach applies to Chapters 7, 8, and 11.

Students who have already taken an introductory, noncalculus-based, applied statistics course will also benefit from a course based on this text. While similar topics are covered, they are presented with more depth and rigor here. For example, *Introduction to the Practice of Statistics*, 6th ed., by D. Moore and G. McCabe (W. H. Freeman, 2009) is an excellent text, and we believe that our book would serve as a strong basis for a follow-up course.

There is an **Instructor's Solutions Manual** available from the publisher.

The second edition contains many more basic exercises than the first edition. Also, we have rewritten a number of sections, with the aim of making the material clearer to students. One goal in our rewriting was to subdivide the material into smaller, more digestible components so that key ideas stand out more boldly. There has been a complete typographical redesign that we feel aids in this as well. In the appendices, we have added material on the statistical package R as well as answers for the odd-numbered exercises that students can use to check their understanding.

Many thanks to the reviewers and users for their comments: Abbas Alhakim (Clarkson University), Michelle Baillargeon (McMaster University), Arne C. Bathke (University of Kentucky), Lisa A. Bloomer (Middle Tennessee State University), Christopher Brown (California Lutheran University), Jem N. Corcoran (University of Colorado), Guang Cheng (Purdue University), Yi Cheng (Indiana University South Bend), Eugene Demidenko (Dartmouth College), Robert P. Dobrow (Carleton College), John Ferdinands (Calvin College), Soledad A. Fernandez (The Ohio State University), Paramjit Gill (University of British Columbia Okanagan), Marvin Glover (Milligan College), Ellen Gundlach (Purdue University), Paul Gustafson (University of British Columbia), Jan Hannig (Colorado State University), Solomon W. Harrar (The University of Montana), Susan Herring (Sonoma State University), George F. Hilton (Pacific Union College), Chun Jin (Central Connecticut State University), Paul Joyce (University of Idaho), Hubert Lilliefors (George Washington University), Andy R. Magid (University of Oklahoma), Phil McDonough (University of Toronto), Julia Morton (Nipissing University), Jean D. Opsomer (Colorado State University), Randall H. Rieger (West Chester University), Robert L. Schaefer (Miami University), Osnat Stramer (University of Iowa), Tim B. Swartz (Simon Fraser University), Glen Takahara (Queen's University), Robert D. Thompson (Hunter College), David C. Vaughan (Wilfrid Laurier University), Joseph J. Walker (Georgia State University), Chad Westerland (University of Arizona), Dongfeng Wu (Mississippi State University), Yuehua Wu (York University), Nicholas Zaino (University of Rochester). In particular, Professor Chris Andrews (State University of New York) provided many corrections to the first edition.

The authors would also like to thank many who have assisted in the development of this project. In particular, our colleagues and students at the University of Toronto have been very supportive. Ping Gao, Aysha Hashim, Gun Ho Jang, Hadas Moshonov, and Mahinda Samarakoon helped in many ways. A number of the data sets in Chapter

10 have been used in courses at the University of Toronto for many years and were, we believe, compiled through the work of the late Professor Daniel B. DeLury. Professor David Moore of Purdue University was of assistance in providing several of the tables at the back of the text. Patrick Farace, Anne Scanlan-Rohrer, Chris Spavins, Danielle Swearingin, Brian Tedesco, Vivien Weiss, and Katrina Wilhelm of W. H. Freeman provided much support and encouragement. Our families helped us with their patience and care while we worked at what seemed at times an unending task; many thanks to Rosemary and Heather Evans and Margaret Fulford.

Michael Evans and Jeffrey Rosenthal
Toronto, 2009

We have made the book freely available on the web for several years now. This version corrects an number of issues pointed out to us by users. We are grateful to all those who have brought these to our attention (see the older webpage for individual acknowledgements) and we will continue to make corrections as we are so informed.

Michael Evans and Jeffrey Rosenthal
Toronto, 2023

