

# CONTENTS

<b>Preface</b>	<b>vii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Sets and Functions</b>	<b>3</b>
1.1 Sets, 3	
1.2 Ways of Specifying Sets, 4	
1.3 Universal Sets, 5	
1.4 Subsets, 6	
1.5 Finite and Infinite Sets, 7	
1.6 Venn Diagrams, 8	
1.7 The Empty Set, 9	
1.8 The Union of Sets, 9	
1.9 The Intersection of Sets, 10	
1.10 The Complement of a Set, 12	
1.11 The Difference between Two Sets, 14	
1.12 The Algebra of Sets, 15	
1.13 Postulates and Theorems about Sets, 15	
1.14 Set Products and Relations, 18	
1.15 Specifying Relations, 22	
1.16 The Domain and Range of a Relation, 24	
1.17 Functional Relations, 25	
1.18 Variables and Functional Notation, 27	
1.19 Continuous Variables and Functions, 30	
1.20 Functions with Several Variables, 31	
1.21 Functions and Precise Prediction, 32	
Exercises, 33	

<b>2</b>	<b>Elementary Probability Theory</b>	<b>38</b>
2.1	Simple Experiments, 39	
2.2	Events, 40	
2.3	Events as Sets, 41	
2.4	Families of Events, 42	
2.5	Probability Functions, 43	
2.6	A Special Case: Equally Probable Elementary Events, 48	
2.7	Computing Probabilities, 51	
2.8	Sequences of Events, 53	
2.9	Number of Possible Sequences for $N$ Trials: Counting Rule 1, 54	
2.10	Counting Rule 2, 54	
2.11	Counting Rule 3: Permutations, 55	
2.12	Counting Rule 4: Ordered Combinations, 56	
2.13	Counting Rules 5 and 6: Combinations, 57	
2.14	Some Examples: Poker Hands, 60	
2.15	The Frequency Interpretation of Probability, 62	
2.16	Bernoulli's Theorem, 63	
2.17	An Example of Simple Statistical Inference, 65	
2.18	The Subjective Interpretation of Probability, 68	
2.19	Probability and Betting Odds, 71	
2.20	A Simple Decision-Making Example, 74	
2.21	Joint Events, 77	
2.22	Combining Probabilities of Joint Events, 78	
2.23	Conditional Probability, 80	
2.24	Relations Among Conditional Probabilities, 83	
2.25	Bayes' Theorem, 84	
2.26	Independence, 87	
2.27	Representing Joint Events in Tables, 91	
	Exercises, 95	
 <b>3</b>	 <b>Probability Distributions</b>	 <b>103</b>
3.1	Random Variables, 103	
3.2	Probability Distributions, 106	
3.3	Discrete Random Variables, 108	
3.4	Graphs of Probability Distributions of Discrete Random Variables, 111	
3.5	Continuous Random Variables, 113	
3.6	Graphs of Probability Distributions of Continuous Random Variables, 116	
3.7	Continuous Random Variables as an Idealization, 120	
3.8	Cumulative Distribution Functions, 120	
3.9	Joint Probability Distributions, 125	
3.10	Independence of Random Variables, 132	
3.11	Independence of Functions of Random Variables, 134	
3.12	Summary Measures of Probability Distributions, 134	
3.13	The Expectation of a Random Variable, 136	

- 3.14 The Algebra of Expectations, 139
- 3.15 Measures of Central Tendency, 143
- 3.16 The Mean as the “Center of Gravity” of a Distribution, 144
- 3.17 The Median, 147
- 3.18 The Mode, 149
- 3.19 “Best Guess” Interpretations of the Mean, Mode, and Median, 150
- 3.20 Relations between Central Tendency Measures and the “Shapes” of Distributions, 152
- 3.21 Measures of Dispersion: The Variance, 155
- 3.22 The Standard Deviation, 158
- 3.23 The Mean as the “Origin” for the Variance, 159
- 3.24 Moments of a Distribution, 161
- 3.25 Moments of Joint Distributions, 163
- 3.26 The Relative Location of a Value in a Probability Distribution: Standardized Random Variables, 166
- 3.27 Tchebycheff’s Inequality, 167  
Exercises, 171

#### 4 Special Probability Distributions

178

- 4.1 Bernoulli Trials, 178
- 4.2 Sampling From a Bernoulli Process, 179
- 4.3 Number of Successes as a Random Variable: The Binomial Distribution, 182
- 4.4 The Binomial Distribution and the Binomial Expansion, 185
- 4.5 Probabilities of Intervals in the Binomial Distribution, 185
- 4.6 The Mean and Variance of the Binomial Distribution, 188
- 4.7 The Binomial Distribution of Proportions, 190
- 4.8 The Form of a Binomial Distribution, 191
- 4.9 The Binomial as a Sampling Distribution, 192
- 4.10 A Preview of a Use of the Binomial Distribution, 194
- 4.11 Number of Trials as a Random Variable: The Pascal and Geometric Distributions, 197
- 4.12 The Multinomial Distribution, 200
- 4.13 The Hypergeometric Distribution, 201
- 4.14 The Poisson Process and Distribution, 204
- 4.15 The Poisson Approximation to the Binomial, 208
- 4.16 Summary of Special Discrete Distributions, 209
- 4.17 Special Continuous Distributions, 212
- 4.18 The Normal Distribution, 212
- 4.19 Cumulative Probabilities and Areas for the Normal Distribution, 215
- 4.20 The Use of Tables of the Normal Distribution, 217
- 4.21 The Importance of the Normal Distribution, 220
- 4.22 The Normal Approximation to the Binomial, 222
- 4.23 The Theory of the Normal Distribution of Error, 227
- 4.24 The Exponential Distribution, 228

- 4.25 The Uniform Distribution, 230
- 4.26 The Beta Distribution, 233
- 4.27 Summary of Special Continuous Distributions, 235  
Exercises, 236

## 5 Frequency and Sampling Distributions 242

- 5.1 Random Sampling, 242
- 5.2 Random Number Tables, 243
- 5.3 Measurement Scales, 246
- 5.4 Frequency Distributions, 252
- 5.5 Frequency Distributions with a Small Number of Measurement  
Classes, 253
- 5.6 Grouped Distributions, 255
- 5.7 Class Interval Size and Class Limits, 257
- 5.8 Interval Size and the Number of Class Intervals, 259
- 5.9 Midpoints of Class Intervals, 260
- 5.10 Frequency Distributions with Open or Unequal Class Intervals, 261
- 5.11 Graphs of Distributions: Histograms, 261
- 5.12 Frequency Polygons, 264
- 5.13 Cumulative Frequency Distributions, 265
- 5.14 Joint Frequency Distributions and Inferences about Association, 266
- 5.15 Statistical Relations and Association, 268
- 5.16 Measures of Central Tendency for Frequency Distributions, 269
- 5.17 Measures of Dispersion for Frequency Distributions, 274
- 5.18 Populations, Parameters, and Statistics, 279
- 5.19 Sampling Distributions, 280
- 5.20 Characteristics of Single-Variate Sampling Distributions, 281
- 5.21 The Mean and Variance of a Sampling Distribution, 283
- 5.22 Statistical Properties of Normal Population Distributions—Inde-  
pendence of Sample Mean and Variance, 285
- 5.23 Distributions of Linear Combinations of Variables, 286
- 5.24 Means and Variances of Linear Combinations, 288
- 5.25 The Central Limit Theorem, 292
- 5.26 The Central Limit Theorem and Linear Combinations of Means, 296
- 5.27 Standardized Random Variables Corresponding To Sample Means,  
296
- 5.28 The Uses of Frequency and Sampling Distributions, 297
- 5.29 To What Populations Do Our Inferences Refer?, 298  
Exercises, 300

## 6 Estimation 307

- 6.1 Sample Statistics As Estimators, 307
- 6.2 Unbiased Estimators, 308
- 6.3 Consistency, 311
- 6.4 Relative Efficiency, 313
- 6.5 Sufficiency, 315

- 6.6 Methods For Determining Good Estimators, 316
- 6.7 The Principle of Maximum Likelihood, 318
- 6.8 The Method of Moments, 322
- 6.9 Parameter Estimates Based on Pooled Samples, 324
- 6.10 Sampling From Finite Populations, 326
- 6.11 Interval Estimation—Confidence Intervals, 327
- 6.12 Approximate Confidence Intervals For Proportions, 332
- 6.13 Sample Size and The Accuracy of Estimation of the Mean, 333
- 6.14 The Problem of Unknown  $\sigma^2$ : The  $t$  Distribution, 335
- 6.15 The  $t$  and the Standardized Normal Distribution, 338
- 6.16 Tables of the  $t$  Distribution, 340
- 6.17 The Concept of Degrees of Freedom, 342
- 6.18 Confidence Intervals for the Mean when  $\sigma^2$  Is Unknown, 343
- 6.19 The Sampling Distribution of Differences between Means, 344
- 6.20 Confidence Intervals For Differences in Means, 347
- 6.21 The Chi-Square Distribution, 348
- 6.22 Tables of the Chi-Square Distribution, 353
- 6.23 The Addition of Chi-Square Variables, 353
- 6.24 The Distribution of the Sample Variance from a Normal Population, 354
- 6.25 Confidence Intervals for the Variance and Standard Deviation, 357
- 6.26 The Importance of the Normality Assumption in Inferences About  $\sigma^2$ , 359
- 6.27 The Normal Approximation to the Chi-Square Distribution, 359
- 6.28 The  $F$  Distribution, 360
- 6.29 The Use of  $F$  Tables, 361
- 6.30 Confidence Intervals for Ratios of Variances, 363
- 6.31 Relationships among the Theoretical Sampling Distributions, 364  
Exercises, 367

## 7 Hypothesis Testing

375

- 7.1 Statistical Tests, 375
- 7.2 Statistical Hypotheses, 378
- 7.3 Assumptions in Hypothesis Testing, 379
- 7.4 Testing a Hypothesis in the Light of Sample Evidence, 379
- 7.5 Choosing a Way to Decide between Two Exact Hypotheses, 383
- 7.6 Errors and Losses, 387
- 7.7 Expected Loss as a Criterion for Choosing a Decision-Rule, 389
- 7.8 Subjective Expected Loss as a Criterion for Choosing among Decision-Rules, 392
- 7.9 Conventional Decision Rules: Type I and Type II Errors, 393
- 7.10 Deciding between Two Hypotheses about a Mean, 396
- 7.11 The Option of Suspending Judgment, 399
- 7.12 The Power of a Statistical Test, 401
- 7.13 Power and the Size of  $\alpha$ , 404
- 7.14 The Effect of Sample Size on Power, 405
- 7.15 Testing Inexact Hypotheses, 407

- 7.16 One-Tailed Rejection Regions, 411
- 7.17 Two-Tailed Tests of Hypotheses, 412
- 7.18 Relative Merits of One- and Two-Tailed Tests, 414
- 7.19 Some Remarks on the General Theory Underlying Tests of Hypotheses, 417
- 7.20 Reporting the Results of Tests, 419
- 7.21 Sample Size and  $p$ -Values, 422
- 7.22 Tests Concerning Means, 424
- 7.23 Tests Concerning Differences between Means, 425
- 7.24 Paired Observations, 427
- 7.25 The Power of  $t$  Tests, 429
- 7.26 Testing Hypotheses about a Single Variance, 430
- 7.27 Testing Hypotheses about Two Variances, 432  
Exercises, 433

## 8 Bayesian Inference

444

- 8.1 Introduction, 444
- 8.2 Bayes' Theorem for Discrete Random Variables, 446
- 8.3 Interpretation of the Prior and Posterior Distributions, 448
- 8.4 An Example of Bayesian Inference and Decision, 450
- 8.5 The Assessment of Prior Probabilities, 456
- 8.6 Bayes' Theorem for Continuous Random Variables, 457
- 8.7 Conjugate Prior Distributions, 459
- 8.8 Bayesian Inference for a Bernoulli Process, 460
- 8.9 Stopping Rules and the Likelihood Function, 463
- 8.10 The Use of Beta Prior Distributions: An Example, 466
- 8.11 Bayesian Inference for a Normal Process, 469
- 8.12 Relative Weights of Prior and Sample Information, 471
- 8.13 The Use of Normal Prior Distributions: An Example, 473
- 8.14 Bayesian Inference for Other Processes, 476
- 8.15 The Assessment of Prior Distributions, 476
- 8.16 Sensitivity Analysis, 481
- 8.17 Representing an "Informationless" Prior State, 482
- 8.18 Diffuse Beta and Normal Prior Distributions, 484
- 8.19 Diffuse Prior Distributions and Classical Statistics, 488
- 8.20 The Posterior Distribution and Estimation, 489
- 8.21 Prior and Posterior Odds Ratios, 492
- 8.22 The Importance of the Likelihood Ratio in Classical Tests, 494
- 8.23 The Posterior Distribution and Hypothesis Testing, 496
- 8.24 The Posterior Distribution and Two-Tailed Tests, 499
- 8.25 Bayesian Inference For Several Hypotheses, 502
- 8.26 The Posterior Distribution and Decision Theory, 503  
Exercises, 504

<b>9</b>	<b>Decision Theory</b>	<b>510</b>
9.1	Introduction, 510	
9.2	Certainty versus Uncertainty, 511	
9.3	Payoffs and Losses, 513	
9.4	Nonprobabilistic Criteria for Decision under Uncertainty, 516	
9.5	Probabilistic Criteria for Decision under Uncertainty, 521	
9.6	Utility, 524	
9.7	The Assessment of Utility Functions, 527	
9.8	Utility as a Function of Money, 529	
9.9	Utility and Decision-Making, 532	
9.10	A Formal Statement of the Decision Problem, 533	
9.11	Decision Theory and Point Estimation, 535	
9.12	Point Estimation: Linear Loss Functions, 538	
9.13	Point Estimation: Quadratic Loss Functions, 542	
9.14	Decision Theory and Hypothesis Testing, 545	
9.15	Hypothesis Testing: An Example, 548	
9.16	Inference and Decision, 549	
9.17	Terminal Decisions and Preposterior Decisions, 551	
9.18	The Value of Perfect Information, 551	
9.19	EVPI: An Example, 555	
9.20	The Value of Sample Information, 558	
9.21	Sample Size and ENGS, 562	
9.22	EVSI and ENGS: An Example, 564	
9.23	Linear Payoff Functions: The Two-Action Problem, 570	
9.24	Loss Functions and Linear Payoff Functions, 574	
9.25	Linear Loss Functions and the Normal Process, 577	
9.26	The General Finite-Action Problem with Linear Payoffs and Losses, 582	
9.27	Decision under Certainty, 583	
9.28	Game Theory, 584	
9.29	A Comment on the Different Approaches to Statistical Problems, 585 Exercises, 586	
<b>10</b>	<b>Regression and Correlation</b>	<b>597</b>
10.1	Review of Joint Distributions, 599	
10.2	Correlation, 600	
10.3	Correlation in Bivariate Normal Populations, 606	
10.4	Tests of Hypotheses in Correlation Problems, 609	
10.5	Confidence Intervals for $\rho_{XY}$ , 612	
10.6	An Example of a Correlation Problem, 613	
10.7	The Regression Curve, 616	
10.8	Linear Regression, 619	
10.9	Estimating the Regression Line, 622	

- 10.10 The Idea of Regression toward the Mean, 627
- 10.11 The Regression of  $X$  on  $Y$ , 629
- 10.12 Regression in Bivariate Normal Populations, 632
- 10.13 An Alternative Model for Simple Linear Regression, 634
- 10.14 Tests for Zero Linear Regression, 637
- 10.15 Interval Estimation in Regression Problems, 638
- 10.16 An Example of a Regression Problem, 640
- 10.17 Curvilinear Regression, 644
- 10.18 Applying Least Squares in Curvilinear Regression, 648
- 10.19 Transformation of Variables in Curvilinear Regression, 651
- 10.20 Multiple Regression, 655
- 10.21 The Coefficient of Multiple Correlation, 659
- 10.22 Partial Correlation, 661
- 10.23 Multiple Regression in Matrix Form: The General Linear Model, 663
- 10.24 Inferences about Multiple Regression and Correlation, 668
- 10.25 An Example of a Multiple Regression Problem, 669
- 10.26 Computers and Multiple Regression, 673
- 10.27 Bayesian Inference and Regression Analysis, 674
- 10.28 Time Series and Regression Analysis, 675
- 10.29 A Brief Look at Econometrics: Other Problems in Regression, 677  
Exercises, 679

**11 Sampling Theory, Experimental Design, and Analysis of Variance** **688**

- 11.1 Sampling from Finite Populations, 688
- 11.2 Simple Random Sampling, 691
- 11.3 Stratified Sampling, 693
- 11.4 Cluster Sampling, 698
- 11.5 Multistage Sampling, 700
- 11.6 Other Topics in Sampling Theory, 701
- 11.7 Sample Design and Experimental Design, 703
- 11.8 The Analysis of Variance, 706
- 11.9 The Linear Model for Simple One-Way Analysis of Variance, 707
- 11.10 The Partition of the Sum of Squares for Any Set of  $J$  Distinct Samples, 711
- 11.11 Assumptions Underlying Inferences about Treatment Effects, 713
- 11.12 The Mean Square between Groups, 713
- 11.13 The Mean Square within Groups, 716
- 11.14 The  $F$  Test in the Analysis of Variance, 718
- 11.15 Computational Forms for the Simple Analysis of Variance, 719
- 11.16 An Example of a Simple One-Way Analysis of Variance, 722
- 11.17 The Importance of the Assumptions, 724
- 11.18 Estimating the Strength of a Statistical Relation from the One-Way Analysis of Variance, 726



- 11.19 The Two-Way Analysis of Variance with Replication, 730
- 11.20 The Linear Model in the Two-Way Analysis, 732
- 11.21 The Importance of Interaction Effects, 735
- 11.22 The Concepts of Replication and Orthogonal Designs, 738
- 11.23 The Partition of the Sum of Squares for the Two-Way Analysis of Variance, 739
- 11.24 Assumptions in the Two-Way Fixed-Effects Model, 741
- 11.25 The Mean Squares and Their Expectations, 742
- 11.26 Computing Forms for the Two-Way Analysis with Replications, 745
- 11.27 An Example, 749
- 11.28 Estimating Strength of Association in Two-Way Analysis of Variance, 752
- 11.29 The Fixed-Effects Model and the Random-Effects Model, 754
- 11.30 Regression and Analysis of Variance: The General Linear Model, 757
- 11.31 Testing for Linear and Nonlinear Regression, 759
- 11.32 The Analysis of Variance as a Summarization of Data, 765
- 11.33 Analyzing Experiments with More Than Two Experimental Factors: The General Problem of Experimental Design, 767  
Exercises, 771

## **12     Nonparametric Methods**

**778**

- 12.1 Comparing Entire Distributions, 779
- 12.2 Comparing Sample and Population Distributions: Goodness of Fit, 781
- 12.3 The Relation to Large Sample Tests of a Single Proportion, 786
- 12.4 A Special Problem: A Goodness-of-Fit Test for a Normal Distribution, 788
- 12.5 Pearson  $\chi^2$  Tests of Association, 791
- 12.6 An Example of a Test for Independence in a Larger Table, 797
- 12.7 The Assumptions in  $\chi^2$  Tests for Association, 799
- 12.8 Measures of Association in Contingency Tables, 801
- 12.9 The Phi Coefficient and Indices of Contingency, 802
- 12.10 A Measure of Predictive Association for Categorical Data, 805
- 12.11 Information Theory and the Analysis of Contingency Tables, 810
- 12.12 Retrospect: The  $\chi^2$  Tests and Measures of Association, 812
- 12.13 Order Statistics, 813
- 12.14 Order Techniques as Substitutes for the Parametric Methods, 814
- 12.15 The Kolmogorov-Smirnov Tests for Goodness of Fit, 817
- 12.16 Comparing Two or More Independent Groups: The Median Test, 821
- 12.17 The Wald-Wolfowitz "Runs" Test for Two Samples, 824
- 12.18 The Mann-Whitney Test for Two Independent Samples, 827
- 12.19 The Sign Test for Matched Pairs, 830
- 12.20 The Wilcoxon Test for Two Matched Samples, 832
- 12.21 The Kruskal-Wallis "Analysis of Variance" by Ranks, 835

12.22	The Friedman Test for $J$ Matched Groups, 837
12.23	Rank-Order Correlation Methods, 839
12.24	The Spearman Rank Correlation Coefficient, 841
12.25	The Kendall Tau Coefficient, 845
12.26	Kendall's Tau versus the Spearman Rank Correlation, 848
12.27	Kendall's Coefficient of Concordance, 849
	Exercises, 852

**Appendix A: Some Common Differentiation and Integration Formulas 861**

**Appendix B: Matrix Algebra 864**

**Tables**

Table I.	Cumulative Normal Probabilities, 873
Table II.	Upper Percentage Points of the $t$ Distribution, 875
Table III.	Upper Percentage Points of the $\chi^2$ Distribution, 876
Table IV.	Percentage Points of the $F$ Distribution, 878
Table V.	Binomial Probabilities, 881
Table VI.	Poisson Probabilities, 886
Table VII.	Unit Normal Linear Loss Integral, 892
Table VIII.	Random Digits, 894
Table IX.	The Transformation of $r$ to $Z$ , 895
Table X.	Upper Percentage Points of $D$ in the Kolmogorov-Smirnov One-Sample Test, 897
Table XI.	Binomial Coefficients, 898
Table XII.	Factorials of Integers, 899
Table XIII.	Powers and Roots, 900

**References and Suggestions for Further Reading 911**

**Answers to Selected Exercises 917**

**Index 927**