
Contents

1	Probability Theory	1
1.1	Set Theory	1
1.2	Basics of Probability Theory	5
1.2.1	Axiomatic Foundations	5
1.2.2	The Calculus of Probabilities	9
1.2.3	Counting	13
1.2.4	Enumerating Outcomes	16
1.3	Conditional Probability and Independence	20
1.4	Random Variables	27
1.5	Distribution Functions	29
1.6	Density and Mass Functions	34
1.7	Exercises	37
1.8	Miscellanea	44
2	Transformations and Expectations	47
2.1	Distributions of Functions of a Random Variable	47
2.2	Expected Values	55
2.3	Moments and Moment Generating Functions	59
2.4	Differentiating Under an Integral Sign	68
2.5	Exercises	76
2.6	Miscellanea	82
3	Common Families of Distributions	85
3.1	Introduction	85
3.2	Discrete Distributions	85
3.3	Continuous Distributions	98
3.4	Exponential Families	111
3.5	Location and Scale Families	116

3.6	Inequalities and Identities	121
3.6.1	Probability Inequalities	122
3.6.2	Identities	123
3.7	Exercises	127
3.8	Miscellanea	135
4	Multiple Random Variables	139
4.1	Joint and Marginal Distributions	139
4.2	Conditional Distributions and Independence	147
4.3	Bivariate Transformations	156
4.4	Hierarchical Models and Mixture Distributions	162
4.5	Covariance and Correlation	169
4.6	Multivariate Distributions	177
4.7	Inequalities	186
4.7.1	Numerical Inequalities	186
4.7.2	Functional Inequalities	189
4.8	Exercises	192
4.9	Miscellanea	203
5	Properties of a Random Sample	207
5.1	Basic Concepts of Random Samples	207
5.2	Sums of Random Variables from a Random Sample	211
5.3	Sampling from the Normal Distribution	218
5.3.1	Properties of the Sample Mean and Variance	218
5.3.2	The Derived Distributions: Student's t and Snedecor's F	222
5.4	Order Statistics	226
5.5	Convergence Concepts	232
5.5.1	Convergence in Probability	232
5.5.2	Almost Sure Convergence	234
5.5.3	Convergence in Distribution	235
5.5.4	The Delta Method	240
5.6	Generating a Random Sample	245
5.6.1	Direct Methods	247
5.6.2	Indirect Methods	251
5.6.3	The Accept/Reject Algorithm	253
5.7	Exercises	255
5.8	Miscellanea	267
6	Principles of Data Reduction	271
6.1	Introduction	271
6.2	The Sufficiency Principle	272
6.2.1	Sufficient Statistics	272
6.2.2	Minimal Sufficient Statistics	279
6.2.3	Ancillary Statistics	282
6.2.4	Sufficient, Ancillary, and Complete Statistics	284

6.3	The Likelihood Principle	290
6.3.1	The Likelihood Function	290
6.3.2	The Formal Likelihood Principle	292
6.4	The Equivariance Principle	296
6.5	Exercises	300
6.6	Miscellanea	307
7	Point Estimation	311
7.1	Introduction	311
7.2	Methods of Finding Estimators	312
7.2.1	Method of Moments	312
7.2.2	Maximum Likelihood Estimators	315
7.2.3	Bayes Estimators	324
7.2.4	The EM Algorithm	326
7.3	Methods of Evaluating Estimators	330
7.3.1	Mean Squared Error	330
7.3.2	Best Unbiased Estimators	334
7.3.3	Sufficiency and Unbiasedness	342
7.3.4	Loss Function Optimality	348
7.4	Exercises	355
7.5	Miscellanea	367
8	Hypothesis Testing	373
8.1	Introduction	373
8.2	Methods of Finding Tests	374
8.2.1	Likelihood Ratio Tests	374
8.2.2	Bayesian Tests	379
8.2.3	Union-Intersection and Intersection-Union Tests	380
8.3	Methods of Evaluating Tests	382
8.3.1	Error Probabilities and the Power Function	382
8.3.2	Most Powerful Tests	387
8.3.3	Sizes of Union-Intersection and Intersection-Union Tests	394
8.3.4	p-Values	397
8.3.5	Loss Function Optimality	400
8.4	Exercises	402
8.5	Miscellanea	413
9	Interval Estimation	417
9.1	Introduction	417
9.2	Methods of Finding Interval Estimators	420
9.2.1	Inverting a Test Statistic	420
9.2.2	Pivotal Quantities	427
9.2.3	Pivoting the CDF	430
9.2.4	Bayesian Intervals	435

9.3	Methods of Evaluating Interval Estimators	440
9.3.1	Size and Coverage Probability	440
9.3.2	Test-Related Optimality	444
9.3.3	Bayesian Optimality	447
9.3.4	Loss Function Optimality	449
9.4	Exercises	451
9.5	Miscellanea	463
10	Asymptotic Evaluations	467
10.1	Point Estimation	467
10.1.1	Consistency	467
10.1.2	Efficiency	470
10.1.3	Calculations and Comparisons	473
10.1.4	Bootstrap Standard Errors	478
10.2	Robustness	481
10.2.1	The Mean and the Median	482
10.2.2	M-Estimators	484
10.3	Hypothesis Testing	488
10.3.1	Asymptotic Distribution of LRTs	488
10.3.2	Other Large-Sample Tests	492
10.4	Interval Estimation	496
10.4.1	Approximate Maximum Likelihood Intervals	496
10.4.2	Other Large-Sample Intervals	499
10.5	Exercises	504
10.6	Miscellanea	515
11	Analysis of Variance and Regression	521
11.1	Introduction	521
11.2	Oneway Analysis of Variance	522
11.2.1	Model and Distribution Assumptions	524
11.2.2	The Classic ANOVA Hypothesis	525
11.2.3	Inferences Regarding Linear Combinations of Means	527
11.2.4	The ANOVA F Test	530
11.2.5	Simultaneous Estimation of Contrasts	534
11.2.6	Partitioning Sums of Squares	536
11.3	Simple Linear Regression	539
11.3.1	Least Squares: A Mathematical Solution	542
11.3.2	Best Linear Unbiased Estimators: A Statistical Solution	544
11.3.3	Models and Distribution Assumptions	548
11.3.4	Estimation and Testing with Normal Errors	550
11.3.5	Estimation and Prediction at a Specified $x = x_0$	557
11.3.6	Simultaneous Estimation and Confidence Bands	559
11.4	Exercises	563
11.5	Miscellanea	572

12 Regression Models	577
12.1 Introduction	577
12.2 Regression with Errors in Variables	577
12.2.1 Functional and Structural Relationships	579
12.2.2 A Least Squares Solution	581
12.2.3 Maximum Likelihood Estimation	583
12.2.4 Confidence Sets	588
12.3 Logistic Regression	591
12.3.1 The Model	591
12.3.2 Estimation	593
12.4 Robust Regression	597
12.5 Exercises	602
12.6 Miscellanea	608
 Appendix: Computer Algebra	 613
 Table of Common Distributions	 621
 References	 629
 Author Index	 645
 Subject Index	 649

List of Tables

1.2.1	Number of arrangements	16
4.1.1	Values of the joint pmf $f(x, y)$	141
7.3.1	Three estimators for a binomial p	354
	Counts of leukemia cases	360
8.3.1	Two types of errors in hypothesis testing	383
9.2.1	Location–scale pivots	427
9.2.2	Sterne’s acceptance region and confidence set	431
	Three 90% normal confidence intervals	441
10.1.1	Bootstrap and Delta Method variances	480
	Median/mean asymptotic relative efficiencies	484
10.2.1	Huber estimators	485
	Huber estimator asymptotic relative efficiencies, $k = 1.5$	487
	Poisson LRT statistic	490
10.3.1	Power of robust tests	497
10.4.1	Confidence coefficient for a pivotal interval	500
10.4.2	Confidence coefficients for intervals based on Huber’s M-estimator	504
11.2.1	ANOVA table for oneway classification	538
11.3.1	Data pictured in Figure 11.3.1	542
11.3.2	ANOVA table for simple linear regression	556
12.3.1	Challenger data	594
12.4.1	Potoroo data	598
	Regression M-estimator asymptotic relative efficiencies	601

List of Figures

1.2.1	Dart board for Example 1.2.7	9
1.2.2	Histogram of averages	19
1.5.1	Cdf of Example 1.5.2	30
1.5.2	Geometric cdf, $p = .3$	32
1.6.1	Area under logistic curve	36
2.1.1	Transformation of Example 2.1.2	49
2.1.2	Increasing and nondecreasing cdfs	54
2.3.1	Exponential densities	60
2.3.2	Two pdfs with the same moments	65
2.3.3	Poisson approximation to the binomial	68
3.3.1	Standard normal density	105
3.3.2	Normal approximation to the binomial	106
3.3.3	Beta densities	107
3.3.4	Symmetric beta densities	108
3.3.5	Standard normal density and Cauchy density	109
3.3.6	Lognormal and gamma pdfs	110
3.5.1	Location densities	117
3.5.2	Exponential location densities	118
3.5.3	Members of the same scale family	119
3.5.4	Location-scale families	120
4.1.1	Regions for Example 4.1.12	147
4.5.1	Regions for Examples 4.5.4 and 4.5.8	170
4.5.2	Regions for Example 4.5.9	175
4.7.1	Convex function	189
4.7.2	Jensen's Inequality	190
5.4.1	Region on which $f_{R,V}(r, v) > 0$ for Example 5.4.7	232
5.6.1	Histogram of exponential pdf	248
5.6.2	Histogram of Poisson sample variances	251
5.6.3	Beta distribution in accept/reject sampling	252

7.3.1	Binomial MSE comparison	333
7.3.2	Risk functions for variance estimators	351
8.2.1	LRT statistic	377
8.3.1	Power functions for Example 8.3.2	384
8.3.2	Power functions for Example 8.3.3	384
8.3.3	Power functions for three tests in Example 8.3.19	394
8.3.4	Risk function for test in Example 8.3.31	401
9.2.1	Confidence interval–acceptance region relationship	421
9.2.2	Acceptance region and confidence interval for Example 9.2.3	423
9.2.3	Credible and confidence intervals from Example 9.2.16	437
9.2.4	Credible probabilities of the intervals from Example 9.2.16	438
9.2.5	Coverage probabilities of the intervals from Example 9.2.16	439
9.3.1	Three interval estimators from Example 9.2.16	449
10.1.1	Asymptotic relative efficiency for gamma mean estimators	478
10.3.1	Poisson LRT histogram	490
10.4.1	LRT intervals for a binomial proportion	502
10.4.2	Coverage probabilities for nominal .9 binomial confidence procedures	503
11.3.1	Vertical distances that are measured by RSS	542
11.3.2	Geometric description of the BLUE	547
11.3.3	Scheffé bands, t interval, and Bonferroni intervals	562
12.2.1	Distance minimized by orthogonal least squares	581
12.2.2	Three regression lines	583
12.2.3	Creasy–Williams F statistic	590
12.3.1	Challenger data logistic curve	595
12.4.1	Least squares, LAD, and M-estimate fits	599

List of Examples

1.1.3	Event operations	3
1.2.2	Sigma algebra-I	6
1.2.3	Sigma algebra-II	6
1.2.5	Defining probabilities-I	7
1.2.7	Defining probabilities-II	8
1.2.10	Bonferroni's Inequality	11
1.2.12	Lottery-I	13
1.2.13	Tournament	13
1.2.15	Lottery-II	14
1.2.18	Poker	16
1.2.19	Sampling with replacement	17
1.2.20	Calculating an average	18
1.3.1	Four aces	20
1.3.3	Continuation of Example 1.3.1	20
1.3.4	Three prisoners	21
1.3.6	Coding	23
1.3.8	Chevalier de Meré	24
1.3.10	Tossing two dice	25
1.3.11	Letters	26
1.3.13	Three coin tosses-I	27
1.4.2	Random variables	28
1.4.3	Three coin tosses-II	28
1.4.4	Distribution of a random variable	29
1.5.2	Tossing three coins	30
1.5.4	Tossing for a head	31
1.5.5	Continuous cdf	32
1.5.6	Cdf with jumps	33
1.5.9	Identically distributed random variables	33
1.6.2	Geometric probabilities	34
1.6.4	Logistic probabilities	36
2.1.1	Binomial transformation	48
2.1.2	Uniform transformation	49
2.1.4	Uniform-exponential relationship-I	51
2.1.6	Inverted gamma pdf	51

2.1.7	Square transformation	52
2.1.9	Normal-chi squared relationship	53
2.2.2	Exponential mean	55
2.2.3	Binomial mean	56
2.2.4	Cauchy mean	56
2.2.6	Minimizing distance	58
2.2.7	Uniform-exponential relationship-II	58
2.3.3	Exponential variance	59
2.3.5	Binomial variance	61
2.3.8	Gamma mgf	63
2.3.9	Binomial mgf	64
2.3.10	Nonunique moments	64
2.3.13	Poisson approximation	66
2.4.5	Interchanging integration and differentiation-I	71
2.4.6	Interchanging integration and differentiation-II	72
2.4.7	Interchanging summation and differentiation	73
2.4.9	Continuation of Example 2.4.7	74
3.2.1	Acceptance sampling	88
3.2.3	Dice probabilities	91
3.2.4	Waiting time	93
3.2.5	Poisson approximation	94
3.2.6	Inverse binomial sampling	96
3.2.7	Failure times	98
3.3.1	Gamma-Poisson relationship	100
3.3.2	Normal approximation	105
3.4.1	Binomial exponential family	111
3.4.3	Binomial mean and variance	112
3.4.4	Normal exponential family	113
3.4.6	Continuation of Example 3.4.4	114
3.4.8	A curved exponential family	115
3.4.9	Normal approximations	115
3.5.3	Exponential location family	118
3.6.2	Illustrating Chebychev	122
3.6.3	A normal probability inequality	123
3.6.6	Higher-order normal moments	125
3.6.9	Higher-order Poisson moments	127
4.1.2	Sample space for dice	140
4.1.4	Continuation of Example 4.1.2	141
4.1.5	Joint pmf for dice	142
4.1.7	Marginal pmf for dice	143
4.1.8	Dice probabilities	144
4.1.9	Same marginals, different joint pmf	144
4.1.11	Calculating joint probabilities-I	145

4.1.12	Calculating joint probabilities-II	146
4.2.2	Calculating conditional probabilities	148
4.2.4	Calculating conditional pdfs	150
4.2.6	Checking independence-I	152
4.2.8	Checking independence-II	153
4.2.9	Joint probability model	154
4.2.11	Expectations of independent variables	155
4.2.13	Mgf of a sum of normal variables	156
4.3.1	Distribution of the sum of Poisson variables	157
4.3.3	Distribution of the product of beta variables	158
4.3.4	Sum and difference of normal variables	159
4.3.6	Distribution of the ratio of normal variables	162
4.4.1	Binomial-Poisson hierarchy	163
4.4.2	Continuation of Example 4.4.1	163
4.4.5	Generalization of Example 4.4.1	165
4.4.6	Beta-binomial hierarchy	167
4.4.8	Continuation of Example 4.4.6	168
4.5.4	Correlation-I	170
4.5.8	Correlation-II	173
4.5.9	Correlation-III	174
4.6.1	Multivariate pdfs	178
4.6.3	Multivariate pmf	181
4.6.8	Mgf of a sum of gamma variables	183
4.6.13	Multivariate change of variables	185
4.7.4	Covariance inequality	188
4.7.8	An inequality for means	191
5.1.2	Sample pdf-exponential	208
5.1.3	Finite population model	210
5.2.8	Distribution of the mean	215
5.2.10	Sum of Cauchy random variables	216
5.2.12	Sum of Bernoulli random variables	217
5.3.5	Variance ratio distribution	224
5.3.7	Continuation of Example 5.3.5	225
5.4.5	Uniform order statistic pdf	230
5.4.7	Distribution of the midrange and range	231
5.5.3	Consistency of S^2	233
5.5.5	Consistency of S	233
5.5.7	Almost sure convergence	234
5.5.8	Convergence in probability, not almost surely	234
5.5.11	Maximum of uniforms	235
5.5.16	Normal approximation to the negative binomial	239
5.5.18	Normal approximation with estimated variance	240
5.5.19	Estimating the odds	240
5.5.22	Continuation of Example 5.5.19	242

5.5.23	Approximate mean and variance	242
5.5.25	Continuation of Example 5.5.23	243
5.5.27	Moments of a ratio estimator	244
5.6.1	Exponential lifetime	246
5.6.2	Continuation of Example 5.6.1	246
5.6.3	Probability Integral Transform	247
5.6.4	Box-Muller algorithm	249
5.6.5	Binomial random variable generation	249
5.6.6	Distribution of the Poisson variance	250
5.6.7	Beta random variable generation—I	251
5.6.9	Beta random variable generation—II	254
6.2.3	Binomial sufficient statistic	274
6.2.4	Normal sufficient statistic	274
6.2.5	Sufficient order statistics	275
6.2.7	Continuation of Example 6.2.4	277
6.2.8	Uniform sufficient statistic	277
6.2.9	Normal sufficient statistic, both parameters unknown	279
6.2.12	Two normal sufficient statistics	280
6.2.14	Normal minimal sufficient statistic	281
6.2.15	Uniform minimal sufficient statistic	282
6.2.17	Uniform ancillary statistic	282
6.2.18	Location family ancillary statistic	283
6.2.19	Scale family ancillary statistic	284
6.2.20	Ancillary precision	285
6.2.22	Binomial complete sufficient statistic	285
6.2.23	Uniform complete sufficient statistic	286
6.2.26	Using Basu's Theorem—I	288
6.2.27	Using Basu's Theorem—II	289
6.3.2	Negative binomial likelihood	290
6.3.3	Normal fiducial distribution	291
6.3.4	Evidence function	292
6.3.5	Binomial/negative binomial experiment	293
6.3.7	Continuation of Example 6.3.5	295
6.4.1	Binomial equivariance	297
6.4.3	Continuation of Example 6.4.1	298
6.4.5	Conclusion of Example 6.4.1	299
6.4.6	Normal location invariance	299
7.2.1	Normal method of moments	313
7.2.2	Binomial method of moments	313
7.2.3	Satterthwaite approximation	314
7.2.5	Normal likelihood	316
7.2.6	Continuation of Example 7.2.5	317
7.2.7	Bernoulli MLE	317

7.2.8	Restricted range MLE	318
7.2.9	Binomial MLE, unknown number of trials	318
7.2.11	Normal MLEs, μ and σ unknown	321
7.2.12	Continuation of Example 7.2.11	322
7.2.13	Continuation of Example 7.2.2	323
7.2.14	Binomial Bayes estimation	324
7.2.16	Normal Bayes estimators	326
7.2.17	Multiple Poisson rates	326
7.2.18	Continuation of Example 7.2.17	327
7.2.19	Conclusion of Example 7.2.17	328
7.3.3	Normal MSE	331
7.3.4	Continuation of Example 7.3.3	331
7.3.5	MSE of binomial Bayes estimator	332
7.3.6	MSE of equivariant estimators	333
7.3.8	Poisson unbiased estimation	335
7.3.12	Conclusion of Example 7.3.8	338
7.3.13	Unbiased estimator for the scale uniform	339
7.3.14	Normal variance bound	340
7.3.16	Continuation of Example 7.3.14	341
7.3.18	Conditioning on an insufficient statistic	343
7.3.21	Unbiased estimators of zero	345
7.3.22	Continuation of Example 7.3.13	346
7.3.24	Binomial best unbiased estimation	347
7.3.25	Binomial risk functions	350
7.3.26	Risk of normal variance	350
7.3.27	Variance estimation using Stein's loss	351
7.3.28	Two Bayes rules	353
7.3.29	Normal Bayes estimates	353
7.3.30	Binomial Bayes estimates	354
8.2.2	Normal LRT	375
8.2.3	Exponential LRT	376
8.2.5	LRT and sufficiency	378
8.2.6	Normal LRT with unknown variance	378
8.2.7	Normal Bayesian test	379
8.2.8	Normal union-intersection test	380
8.2.9	Acceptance sampling	382
8.3.2	Binomial power function	383
8.3.3	Normal power function	384
8.3.4	Continuation of Example 8.3.3	385
8.3.7	Size of LRT	386
8.3.8	Size of union-intersection test	387
8.3.10	Conclusion of Example 8.3.3	387
8.3.14	UMP binomial test	390
8.3.15	UMP normal test	390

8.3.18	Continuation of Example 8.3.15	392
8.3.19	Nonexistence of UMP test	392
8.3.20	Unbiased test	393
8.3.22	An equivalence	395
8.3.25	Intersection-union test	396
8.3.28	Two-sided normal p-value	398
8.3.29	One-sided normal p-value	398
8.3.30	Fisher's Exact Test	399
8.3.31	Risk of UMP test	401
9.1.2	Interval estimator	418
9.1.3	Continuation of Example 9.1.2	418
9.1.6	Scale uniform interval estimator	419
9.2.1	Inverting a normal test	420
9.2.3	Inverting an LRT	423
9.2.4	Normal one-sided confidence bound	425
9.2.5	Binomial one-sided confidence bound	425
9.2.7	Location-scale pivots	427
9.2.8	Gamma pivot	428
9.2.9	Continuation of Example 9.2.8	429
9.2.10	Normal pivotal interval	429
9.2.11	Shortest length binomial set	431
9.2.13	Location exponential interval	433
9.2.15	Poisson interval estimator	434
9.2.16	Poisson credible set	436
9.2.17	Poisson credible and coverage probabilities	437
9.2.18	Coverage of a normal credible set	438
9.3.1	Optimizing length	441
9.3.3	Optimizing expected length	443
9.3.4	Shortest pivotal interval	443
9.3.6	UMA confidence bound	445
9.3.8	Continuation of Example 9.3.6	446
9.3.11	Poisson HPD region	448
9.3.12	Normal HPD region	449
9.3.13	Normal interval estimator	450
10.1.2	Consistency of \bar{X}	468
10.1.4	Continuation of Example 10.1.2	469
10.1.8	Limiting variances	470
10.1.10	Large-sample mixture variances	471
10.1.13	Asymptotic normality and consistency	472
10.1.14	Approximate binomial variance	474
10.1.15	Continuation of Example 10.1.14	475
10.1.17	AREs of Poisson estimators	476
10.1.18	Estimating a gamma mean	477

10.1.19	Bootstrapping a variance	478
10.1.20	Bootstrapping a binomial variance	479
10.1.21	Conclusion of Example 10.1.20	479
10.1.22	Parametric bootstrap	480
10.2.1	Robustness of the sample mean	482
10.2.3	Asymptotic normality of the median	483
10.2.4	AREs of the median to the mean	484
10.2.5	Huber estimator	485
10.2.6	Limit distribution of the Huber estimator	486
10.2.7	ARE of the Huber estimator	487
10.3.2	Poisson LRT	489
10.3.4	Multinomial LRT	491
10.3.5	Large-sample binomial tests	493
10.3.6	Binomial score test	495
10.3.7	Tests based on the Huber estimator	496
10.4.1	Continuation of Example 10.1.14	497
10.4.2	Binomial score interval	498
10.4.3	Binomial LRT interval	499
10.4.4	Approximate interval	499
10.4.5	Approximate Poisson interval	500
10.4.6	More on the binomial score interval	501
10.4.7	Comparison of binomial intervals	502
10.4.8	Intervals based on the Huber estimator	503
10.4.9	Negative binomial interval	504
10.6.2	Influence functions of the mean and median	518
11.2.1	Oneway ANOVA	522
11.2.3	The ANOVA hypothesis	525
11.2.6	ANOVA contrasts	529
11.2.9	Pairwise differences	534
11.2.12	Continuation of Example 11.2.1	538
11.3.1	Predicting grape crops	540
11.3.4	Continuation of Example 11.3.1	555
12.2.1	Estimating atmospheric pressure	579
12.3.1	Challenger data	594
12.3.2	Challenger data continued	596
12.4.1	Robustness of least squares estimates	597
12.4.2	Catastrophic observations	598
12.4.3	Asymptotic normality of the LAD estimator	599
12.4.4	Regression M-estimator	601
12.4.5	Simulation of regression AREs	601
A.0.1	Unordered sampling	613
A.0.2	Univariate transformation	614
A.0.3	Bivariate transformations	614

A.0.4	Normal probability	616
A.0.5	Density of a sum	616
A.0.6	Fourth moment of sum of uniforms	617
A.0.7	ARE for a gamma mean	618
A.0.8	Limit of chi squared mgfs	619