

Contents

Preface	xi
1 Probability	1
1.1 The Laws of Probability	1
1.2 Probability Distributions	5
1.2.1 Discrete and Continuous Probability Distributions	5
1.2.2 Cumulative Probability Distribution Function	8
1.2.3 Change of Variables	8
1.3 Characterizations of Probability Distributions	9
1.3.1 Medians, Modes, and Full Width at Half Maximum	9
1.3.2 Moments, Means, and Variances	10
1.3.3 Moment Generating Function and the Characteristic Function	14
1.4 Multivariate Probability Distributions	16
1.4.1 Distributions with Two Independent Variables	16
1.4.2 Covariance	17
1.4.3 Distributions with Many Independent Variables	19
2 Some Useful Probability Distribution Functions	22
2.1 Combinations and Permutations	22
2.2 Binomial Distribution	24
2.3 Poisson Distribution	27
2.4 Gaussian or Normal Distribution	31
2.4.1 Derivation of the Gaussian Distribution—Central Limit Theorem	31
2.4.2 Summary and Comments on the Central Limit Theorem	34
2.4.3 Mean, Moments, and Variance of the Gaussian Distribution	36
2.5 Multivariate Gaussian Distribution	37
2.6 χ^2 Distribution	41
2.6.1 Derivation of the χ^2 Distribution	41
2.6.2 Mean, Mode, and Variance of the χ^2 Distribution	44
2.6.3 χ^2 Distribution in the Limit of Large n	45

2.6.4 Reduced χ^2	46
2.6.5 χ^2 for Correlated Variables	46
2.7 Beta Distribution	47
3 Random Numbers and Monte Carlo Methods	50
3.1 Introduction	50
3.2 Nonuniform Random Deviates	51
3.2.1 Inverse Cumulative Distribution Function Method	52
3.2.2 Multidimensional Deviates	53
3.2.3 Box-Müller Method for Generating Gaussian Deviates	53
3.2.4 Acceptance-Rejection Algorithm	54
3.2.5 Ratio of Uniforms Method	57
3.2.6 Generating Random Deviates from More Complicated Probability Distributions	59
3.3 Monte Carlo Integration	59
3.4 Markov Chains	63
3.4.1 Stationary, Finite Markov Chains	63
3.4.2 Invariant Probability Distributions	65
3.4.3 Continuous Parameter and Multiparameter Markov Chains	68
3.5 Markov Chain Monte Carlo Sampling	71
3.5.1 Examples of Markov Chain Monte Carlo Calculations	71
3.5.2 Metropolis-Hastings Algorithm	72
3.5.3 Gibbs Sampler	77
4 Elementary Frequentist Statistics	81
4.1 Introduction to Frequentist Statistics	81
4.2 Means and Variances for Unweighted Data	82
4.3 Data with Uncorrelated Measurement Errors	86
4.4 Data with Correlated Measurement Errors	91
4.5 Variance of the Variance and Student's t Distribution	95
4.5.1 Variance of the Variance	96
4.5.2 Student's t Distribution	98
4.5.3 Summary	100
4.6 Principal Component Analysis	101
4.6.1 Correlation Coefficient	101
4.6.2 Principal Component Analysis	102
4.7 Kolmogorov-Smirnov Test	107
4.7.1 One-Sample K-S Test	107
4.7.2 Two-Sample K-S Test	109
5 Linear Least Squares Estimation	111
5.1 Introduction	111
5.2 Likelihood Statistics	112
5.2.1 Likelihood Function	112
5.2.2 Maximum Likelihood Principle	115
5.2.3 Relation to Least Squares and χ^2 Minimization	119

5.3 Fits of Polynomials to Data	120
5.3.1 Straight Line Fits	120
5.3.2 Fits with Polynomials of Arbitrary Degree	126
5.3.3 Variances, Covariances, and Biases	128
5.3.4 Monte Carlo Error Analysis	136
5.4 Need for Covariances and Propagation of Errors	137
5.4.1 Need for Covariances	137
5.4.2 Propagation of Errors	139
5.4.3 Monte Carlo Error Propagation	142
5.5 General Linear Least Squares	144
5.5.1 Linear Least Squares with Nonpolynomial Functions	144
5.5.2 Fits with Correlations among the Measurement Errors	147
5.5.3 χ^2 Test for Goodness of Fit	149
5.6 Fits with More Than One Dependent Variable	152
6 Nonlinear Least Squares Estimation	155
6.1 Introduction	155
6.2 Linearization of Nonlinear Fits	157
6.2.1 Data with Uncorrelated Measurement Errors	158
6.2.2 Data with Correlated Measurement Errors	161
6.2.3 Practical Considerations	162
6.3 Other Methods for Minimizing S	163
6.3.1 Grid Mapping	163
6.3.2 Method of Steepest Descent, Newton's Method, and Marquardt's Method	164
6.3.3 Simplex Optimization	168
6.3.4 Simulated Annealing	168
6.4 Error Estimation	171
6.4.1 Inversion of the Hessian Matrix	171
6.4.2 Direct Calculation of the Covariance Matrix	173
6.4.3 Summary and the Estimated Covariance Matrix	176
6.5 Confidence Limits	176
6.6 Fits with Errors in Both the Dependent and Independent Variables	181
6.6.1 Data with Uncorrelated Errors	182
6.6.2 Data with Correlated Errors	184
7 Bayesian Statistics	187
7.1 Introduction to Bayesian Statistics	187
7.2 Single-Parameter Estimation: Means, Modes, and Variances	191
7.2.1 Introduction	191
7.2.2 Gaussian Priors and Likelihood Functions	192
7.2.3 Binomial and Beta Distributions	194
7.2.4 Poisson Distribution and Uniform Priors	195
7.2.5 <i>More about the Prior Probability Distribution</i>	198
7.3 Multiparameter Estimation	199
7.3.1 Formal Description of the Problem	199
7.3.2 Laplace Approximation	200

7.3.3 Gaussian Likelihoods and Priors: Connection to Least Squares	203
7.3.4 Difficult Posterior Distributions: Markov Chain Monte Carlo Sampling	211
7.3.5 Credible Intervals	212
7.4 Hypothesis Testing	214
7.5 Discussion	217
7.5.1 Prior Probability Distribution	217
7.5.2 Likelihood Function	218
7.5.3 Posterior Distribution Function	218
7.5.4 Meaning of Probability	219
7.5.5 Thoughts	219
8 Introduction to Fourier Analysis	221
8.1 Introduction	221
8.2 Complete Sets of Orthonormal Functions	221
8.3 Fourier Series	226
8.4 Fourier Transform	233
8.4.1 Fourier Transform Pairs	234
8.4.2 Summary of Useful Fourier Transform Pairs	241
8.5 Discrete Fourier Transform	242
8.5.1 Derivation from the Continuous Fourier Transform	243
8.5.2 Derivation from the Orthogonality Relations for Discretely Sampled Sine and Cosine Functions	245
8.5.3 Parseval's Theorem and the Power Spectrum	248
8.6 Convolution and the Convolution Theorem	249
8.6.1 Convolution	249
8.6.2 Convolution Theorem	254
9 Analysis of Sequences: Power Spectra and Periodograms	256
9.1 Introduction	256
9.2 Continuous Sequences: Data Windows, Spectral Windows, and Aliasing	256
9.2.1 Data Windows and Spectral Windows	257
9.2.2 Aliasing	263
9.2.3 Arbitrary Data Windows	265
9.3 Discrete Sequences	265
9.3.1 The Need to Oversample F_m	266
9.3.2 Nyquist Frequency	267
9.3.3 Integration Sampling	270
9.4 Effects of Noise	271
9.4.1 Deterministic and Stochastic Processes	271
9.4.2 Power Spectrum of White Noise	272
9.4.3 Deterministic Signals in the Presence of Noise	275
9.4.4 Nonwhite, Non-Gaussian Noise	277
9.5 Sequences with Uneven Spacing	278
9.5.1 Least Squares Periodogram	278
9.5.2 Lomb-Scargle Periodogram	280
9.5.3 Generalized Lomb-Scargle Periodogram	284
9.6 Signals with Variable Periods: The O–C Diagram	287

10 Analysis of Sequences: Convolution and Covariance	292
10.1 Convolution Revisited	292
10.1.1 Impulse Response Function	292
10.1.2 Frequency Response Function	297
10.2 Deconvolution and Data Reconstruction	301
10.2.1 Effect of Noise on Deconvolution	301
10.2.2 Wiener Deconvolution	305
10.2.3 Richardson-Lucy Algorithm	308
10.3 Autocovariance Functions	309
10.3.1 Basic Properties of Autocovariance Functions	309
10.3.2 Relation to the Power Spectrum	313
10.3.3 Application to Stochastic Processes	316
10.4 Cross-Covariance Functions	324
10.4.1 Basic Properties of Cross-Covariance Functions	324
10.4.2 Relation to χ^2 and to the Cross Spectrum	326
10.4.3 Detection of Pulsed Signals in Noise	329
Appendices	
A Some Useful Definite Integrals	333
B Method of Lagrange Multipliers	338
C Additional Properties of the Gaussian Probability Distribution	342
D The n -Dimensional Sphere	352
E Review of Linear Algebra and Matrices	354
F Limit of $[1 + f(x)/n]^n$ for Large n	374
G Green's Function Solutions for Impulse Response Functions	375
H Second-Order Autoregressive Process	379
Bibliography	383
Index	385