
Table of Contents

Preface	xv
---------------	----

Chapter 1

Introduction

1.1 What Is Computational Statistics?	1
1.2 An Overview of the Book	2
Philosophy	2
What Is Covered	3
A Word About Notation	5
1.3 MATLAB Code	6
Computational Statistics Toolbox	7
Internet Resources	7
1.4 Further Reading	8

Chapter 2

Probability Concepts

2.1 Introduction	11
2.2 Probability	12
Background	12
Probability	14
Axioms of Probability	17
2.3 Conditional Probability and Independence	17
Conditional Probability	17
Independence	18
Bayes Theorem	19
2.4 Expectation	21
Mean and Variance	21
Skewness	23
Kurtosis	23
2.5 Common Distributions	24
Binomial	24
Poisson	26
Uniform	29
Normal	31
Exponential	33
Gamma	36
Chi-Square	37
Weibull	38
Beta	40

Multivariate Normal	41
2.6 MATLAB Code	45
2.7 Further Reading	46
Exercises	48

Chapter 3

Sampling Concepts

3.1 Introduction	51
3.2 Sampling Terminology and Concepts	51
Sample Mean and Sample Variance	53
Sample Moments	54
Covariance	56
3.3 Sampling Distributions	58
3.4 Parameter Estimation	60
Bias	61
Mean Squared Error	61
Relative Efficiency	62
Standard Error	62
Maximum Likelihood Estimation	63
Method of Moments	66
3.5 Empirical Distribution Function	68
Quantiles	69
3.6 MATLAB Code	73
3.7 Further Reading	74
Exercises	76

Chapter 4

Generating Random Variables

4.1 Introduction	79
4.2 General Techniques for Generating Random Variables	79
Uniform Random Numbers	79
Inverse Transform Method	82
Acceptance-Rejection Method	85
4.3 Generating Continuous Random Variables	89
Normal Distribution	89
Exponential Distribution	89
Gamma	91
Chi-Square	93
Beta	95
Multivariate Normal	96
Generating Variates on a Sphere	99
4.4 Generating Discrete Random Variables	100
Binomial	100
Poisson	102
Discrete Uniform	104

4.5 MATLAB Code	106
4.6 Further Reading	107
Exercises	109

Chapter 5

Exploratory Data Analysis

5.1 Introduction	111
5.2 Exploring Univariate Data	112
Histograms	113
Stem-and-Leaf	116
Quantile-Based Plots - Continuous Distributions	119
Q-Q Plot	119
Quantile Plots	123
Quantile Plots - Discrete Distributions	126
Poissonness Plot	126
Binomialness Plot	129
Box Plots	132
5.3 Exploring Bivariate and Trivariate Data	135
Scatterplots	135
Surface Plots	138
Contour Plots	138
Bivariate Histogram	141
3-D Scatterplot	145
5.4 Exploring Multi-Dimensional Data	147
Scatterplot Matrix	147
Slices and Isosurfaces	149
Star Plots	155
Andrews Curves	157
Parallel Coordinates	162
Projection Pursuit	168
Projection Pursuit Index	171
Finding the Structure	172
Structure Removal	174
Grand Tour	178
5.5 MATLAB Code	183
5.6 Further Reading	184
Exercises	187

Chapter 6

Monte Carlo Methods for Inferential Statistics

6.1 Introduction	191
6.2 Classical Inferential Statistics	192
Hypothesis Testing	192
Confidence Intervals	201
6.3 Monte Carlo Methods for Inferential Statistics	204

Basic Monte Carlo Procedure	204	
Monte Carlo Hypothesis Testing	205	
Monte Carlo Assessment of Hypothesis Testing	210	
6.4 Bootstrap Methods	214	
General Bootstrap Methodology	214	
Bootstrap Estimate of Standard Error	216	
Bootstrap Estimate of Bias	219	
Bootstrap Confidence Intervals	220	
Bootstrap Standard Confidence Interval	220	
Bootstrap- <i>t</i> Confidence Interval	221	
Bootstrap Percentile Interval	224	
6.5 MATLAB Code	226	
6.6 Further Reading	227	7
Exercises	228	8

Chapter 7

Data Partitioning

7.1 Introduction	231	1
7.2 Cross-Validation	232	
7.3 Jackknife	239	9
7.4 Better Bootstrap Confidence Intervals	247	
7.5 Jackknife-After-Bootstrap	251	
7.6 MATLAB Code	253	
7.7 Further Reading	254	1
Exercises	256	6

Chapter 8

Probability Density Estimation

8.1 Introduction	259	9
8.2 Histograms	261	
1-D Histograms	261	
Multivariate Histograms	267	
Frequency Polygons	269	
Averaged Shifted Histograms	274	
8.3 Kernel Density Estimation	280	
Univariate Kernel Estimators	280	
Multivariate Kernel Estimators	285	
8.4 Finite Mixtures	287	7
Univariate Finite Mixtures	289	
Visualizing Finite Mixtures	291	
Multivariate Finite Mixtures	293	
EM Algorithm for Estimating the Parameters	296	
Adaptive Mixtures	301	
8.5 Generating Random Variables	306	
8.6 MATLAB Code	311	

8.7 Further Reading	311
Exercises	314

Chapter 9

Statistical Pattern Recognition

9.1 Introduction	317
9.2 Bayes Decision Theory	319
Estimating Class-Conditional Probabilities: Parametric Method ...	321
Estimating Class-Conditional Probabilities: Nonparametric	322
Bayes Decision Rule	323
Likelihood Ratio Approach	329
9.3 Evaluating the Classifier	332
Independent Test Sample	333
Cross-Validation	335
Receiver Operating Characteristic (ROC) Curve	337
9.4 Classification Trees	342
Growing the Tree	347
Pruning the Tree	352
Choosing the Best Tree	356
Selecting the Best Tree Using an Independent Test Sample ...	357
Selecting the Best Tree Using Cross-Validation	361
9.5 Clustering	367
Measures of Distance	367
Hierarchical Clustering	369
K-Means Clustering	373
9.6 MATLAB Code	376
9.7 Further Reading	379
Exercises	381

Chapter 10

Nonparametric Regression

10.1 Introduction	385
10.2 Smoothing	390
Loess	391
Robust Loess Smoothing	396
Upper and Lower Smooths	400
10.3 Kernel Methods	401
Nadaraya-Watson Estimator	404
Local Linear Kernel Estimator	405
10.4 Regression Trees	407
Growing a Regression Tree	410
Pruning a Regression Tree	411
Selecting a Tree	412
10.5 MATLAB Code	419
10.6 Further Reading	420

Appendix E	
MATLAB Statistics Toolbox	547
Appendix F	
Computational Statistics Toolbox	557
Appendix G	
Data Sets	563
References	571
Index	585