Chapman & Hall/CRC
Applied Environmental Statistics

Statistics for Environmental Science and Management

Second Edition

BRYAN F. J. MANLY
Western EcoSystem Technology, Inc.
Cheyenne, Wyoming, U.S.A.

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business

A CHAPMAN & HALL BOOK

Contents

Preface to the Second Edition xi
Preface to the First Edition xiii

1. The Role of Statistics in Environmental Science 1
1.1 Introduction 1
1.2 Some Examples 1
1.3 The Importance of Statistics in the Examples 19
1.4 Chapter Summary 19
Exercises 20

2. Environmental Sampling 23
2.1 Introduction 23
2.2 Simple Random Sampling 24
2.3 Estimation of Population Means 24
2.4 Estimation of Population Totals 29
2.5 Estimation of Proportions 30
2.6 Sampling and Nonsampling Errors 32
2.7 Stratified Random Sampling 33
2.8 Post-Stratification 38
2.9 Systematic Sampling 39
2.10 Other Design Strategies 44
2.11 Ratio Estimation 46
2.12 Double Sampling 50
2.13 Choosing Sample Sizes 51
2.14 Unequal-Probability Sampling 53
2.15 The Data Quality Objectives Process 55
2.16 Chapter Summary 56
Exercises 58

3. Models for Data 61
3.1 Statistical Models 61
3.2 Discrete Statistical Distributions 61
3.2.1 The Hypergeometric Distribution 62
3.2.2 The Binomial Distribution 63
3.2.3 The Poisson Distribution 64
3.3 Continuous Statistical Distributions 65
3.3.1 The Exponential Distribution 66
3.3.2 The Normal or Gaussian Distribution 67
3.3.3 The Lognormal Distribution 67
3.4 The Linear Regression Model 68
3.5 Factorial Analysis of Variance 74

3.5.1 One-Factor Analysis of Variance 76
3.5.2 Two-Factor Analysis of Variance 76
3.5.3 Three-Factor Analysis of Variance 78
3.5.4 Repeated-Measures Designs 82
3.5.5 Multiple Comparisons and Contrasts 83
3.6 Generalized Linear Models 84
3.7 Chapter Summary 90
Exercises 91

4. Drawing Conclusions from Data 97
4.1 Introduction 97
4.2 Observational and Experimental Studies 97
4.3 True Experiments and Quasi-Experiments 99
4.4 Design-Based and Model-Based Inference 101
4.5 Tests of Significance and Confidence Intervals 103
4.6 Randomization Tests 105
4.7 Bootstrapping 108
4.8 Pseudoreplication 110
4.9 Multiple Testing 112
4.10 Meta-Analysis 114
4.11 Bayesian Inference 119
4.12 Chapter Summary 120
Exercises 122

5. Environmental Monitoring 125
5.1 Introduction 125
5.2 Purposely Chosen Monitoring Sites 126
5.3 Two Special Monitoring Designs 126
5.4 Designs Based on Optimization 129
5.5 Monitoring Designs Typically Used 129
5.6 Detection of Changes by Analysis of Variance 131
5.7 Detection of Changes Using Control Charts 133
5.8 Detection of Changes Using CUSUM Charts 140
5.9 Chi-Squared Tests for a Change in a Distribution 145
5.10 Chapter Summary 149
Exercises 150

6. Impact Assessment 153
6.1 Introduction 153
6.2 The Simple Difference Analysis with BACI Designs 155
6.3 Matched Pairs with a BACI Design 158
6.4 Impact-Control Designs 161
6.5 Before–After Designs 162
6.6 Impact-Gradient Designs 163
6.7 Inferences from Impact Assessment Studies 163

6.8 Chapter Summary 164
Exercises 165

7. Assessing Site Reclamation 167
7.1 Introduction 167
7.2 Problems with Tests of Significance 167
7.3 The Concept of Bioequivalence 168
7.4 Two-Sided Tests of Bioequivalence 171
7.5 Chapter Summary 176
Exercises 177

8. Time Series Analysis 179
8.1 Introduction 179
8.2 Components of Time Series 180
8.3 Serial Correlation 182
8.4 Tests for Randomness 186
8.5 Detection of Change Points and Trends 190
8.6 More-Complicated Time Series Models 194
8.7 Frequency Domain Analysis 201
8.8 Forecasting 202
8.9 Chapter Summary 203
Exercises 204

9. Spatial-Data Analysis 207
9.1 Introduction 207
9.2 Types of Spatial Data 207
9.3 Spatial Patterns in Quadrat Counts 211
9.4 Correlation between Quadrat Counts 217
9.5 Randomness of Point Patterns 219
9.6 Correlation between Point Patterns 221
9.7 Mantel Tests for Autocorrelation 222
9.8 The Variogram 224
9.9 Kriging 228
9.10 Correlation between Variables in Space 230
9.11 Chapter Summary 231
Exercises 233

10. Censored Data 237
10.1 Introduction 237
10.2 Single Sample Estimation 237
10.3 Estimation of Quantiles 244
10.4 Comparing the Means of Two or More Samples 244
10.5 Regression with Censored Data 247
10.6 Chapter Summary 247
Exercises 248

11. Monte Carlo Risk Assessment 249
11.1 Introduction 249
11.2 Principles for Monte Carlo Risk Assessment 250
11.3 Risk Analysis Using a Spreadsheet 251
11.4 Chapter Summary 253
Exercises 253

12. Final Remarks 255

Appendices 257

References 279

Index 291

Preface to the Second Edition

The main changes for the second edition of the book have been the correction of a few errors that have either been pointed out by readers of the first edition or noticed by me in the updating of references and the text, particularly in terms of the software needed for calculations and changes to the web sites, and the addition of some exercises at the end of chapters. I would particularly like to thank students attending my workshops and courses at statistics.com for helping me to clarify parts of the text that were not altogether clear in the first edition.

The aims of the book are still the same as for the first edition; namely, to introduce environmental scientists and managers to the statistical methods that will be useful for them in their work, and also as a text suitable for a course in statistics for graduate students in the environmental science area.

Bryan Manly
March 2008

Preface to the First Edition

This book is intended to introduce environmental scientists and managers to the statistical methods that will be useful for them in their work. A secondary aim was to produce a text suitable for a course in statistics for graduate students in the environmental science area. I wrote the book because it seemed to me that these groups should really learn about statistical methods in a special way. It is true that their needs are similar in many respects to those working in other areas. However, there are some special topics that are relevant to environmental science to the extent that they should be covered in an introductory text, although they would probably not be mentioned at all in such a text for a more general audience. I refer to environmental monitoring, impact assessment, which all have their own chapters here.

The book is not intended to be a complete introduction to statistics. Rather, it is assumed that readers have already taken a course or read a book on basic methods, covering the ideas of random variation, statistical distributions, tests of significance, and confidence intervals. For those who have done this some time ago, Appendix A is meant to provide a quick refresher course.

A number of people have contributed directly or indirectly to this book. I must first mention Lyman McDonald of West, Inc., Cheyenne, WY, who first stimulated my interest in environmental statistics, as distinct from ecological statistics. Much of the contents of the book are influenced by the discussions that we have had on matters statistical. Jennifer Brown from the University of Canterbury in New Zealand has influenced the contents because we have shared the teaching of several short courses on statistics for environmental scientists and managers. Likewise, sharing a course on statistics for MSc students of environmental science with Caryn Thompson and David Fletcher has also had an effect on the book. Other people are too numerous to name, so I would just like to thank generally those who have contributed data sets, helped me check references and equations, etc.

Most of this book was written in the Department of Mathematics and Statistics at the University of Otago. As usual, the university was generous with the resources that are needed for the major effort of writing a book, including periods of sabbatical leave that enabled me to write large parts of the text without interruptions, and an excellent library.

However, the manuscript would definitely have taken longer to finish if I had not been invited to spend part of the year 2000 as a visiting researcher at the Max Planck Institute for Limnology at Plön in Germany. This enabled me to write the final chapters and put the whole book together. I am very grateful to Winfried Lampert, the Director of the Institute, for his kind invitation to come to Plön, and for allowing me to use the excellent facilities at the Institute while I was there.

The Saul Bellow quotation above may need some explanation. It results from attending meetings where an environmental matter is argued at length, with everyone being ignorant of the true facts of the case. Furthermore, one suspects that some people there would prefer not to know the true facts because this would be likely to end the arguments.

Bryan F.J. Manly
May 2000