

STATISTICS FOR EARTH AND ENVIRONMENTAL SCIENTISTS

John H. Schuenemeyer

*Southwest Statistical Consulting, LLC
Cortez, Colorado*

Lawrence J. Drew

*United States Geological Survey
Reston, Virginia*



WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

CONTENTS

Preface	ix
1 Role of Statistics and Data Analysis	1
1.1 Introduction	1
1.2 Case Studies	1
1.3 Data	2
1.4 Samples Versus the Population: Some Notation	3
1.5 Vector and Matrix Notation	4
1.6 Frequency Distributions and Histograms	5
1.7 Distribution as a Model	6
1.8 Sample Moments	9
1.9 Normal (Gaussian) Distribution	12
1.10 Exploratory Data Analysis	13
1.11 Estimation	17
1.12 Bias	18
1.13 Causes of Variance	21
1.14 About Data	21
1.15 Reasons to Conduct Statistically Based Studies	24
1.16 Data Mining	25
1.17 Modeling	25
1.18 Transformations	27
1.19 Statistical Concepts	28
1.20 Statistics Paradigms	30
1.21 Summary	33
Exercises	34
2 Modeling Concepts	37
2.1 Introduction	37
2.2 Why Construct a Model?	37
2.3 What Does a Statistical Model Do?	38
2.4 Steps in Modeling	39
2.5 Is a Model a Unique Solution to a Problem?	44
2.6 Model Assumptions	45

2.7	Designed Experiments	47
2.8	Replication	49
2.9	Summary	49
	Exercises	49
3	Estimation and Hypothesis Testing on Means and Other Statistics	51
3.1	Introduction	51
3.2	Independence of Observations	51
3.3	Central Limit Theorem	52
3.4	Sampling Distributions	53
3.5	Confidence Interval Estimate on a Mean	59
3.6	Confidence Interval on the Difference Between Means	64
3.7	Hypothesis Testing on Means	70
3.8	Bayesian Hypothesis Testing	79
3.9	Nonparametric Hypothesis Testing	82
3.10	Bootstrap Hypothesis Testing on Means	83
3.11	Testing Multiple Means via Analysis of Variance	85
3.12	Multiple Comparisons of Means	87
3.13	Nonparametric ANOVA	90
3.14	Paired Data	91
3.15	Kolmogorov–Smirnov Goodness-of-Fit Test	92
3.16	Comments on Hypothesis Testing	94
3.17	Summary	95
	Exercises	97
4	Regression	99
4.1	Introduction	99
4.2	Pittsburgh Coal Quality Case Study	99
4.3	Correlation and Covariance	100
4.4	Simple Linear Regression	105
4.5	Multiple Regression	125
4.6	Other Regression Procedures	139
4.7	Nonlinear Models	143
4.8	Summary	146
	Exercises	147
5	Time Series	151
5.1	Introduction	151
5.2	Time Domain	152
5.3	Frequency Domain	181
5.4	Wavelets	189

5.5	Summary	189
	Exercises	190
6	Spatial Statistics	193
6.1	Introduction	193
6.2	Data	193
6.3	Three-Dimensional Data Visualization	196
6.4	Spatial Association	199
6.5	Effect of Trend	208
6.6	Semivariogram Models	210
6.7	Kriging	218
6.8	Space-Time Models	237
6.9	Summary	239
	Exercises	240
7	Multivariate Analysis	243
7.1	Introduction	243
7.2	Multivariate Graphics	244
7.3	Principal Components Analysis	246
7.4	Factor Analysis	257
7.5	Cluster Analysis	263
7.6	Multidimensional Scaling	276
7.7	Discriminant Analysis	276
7.8	Tree-Based Modeling	286
7.9	Summary	289
	Exercises	290
8	Discrete Data Analysis and Point Processes	293
8.1	Introduction	293
8.2	Discrete Process and Distributions	293
8.3	Point Processes	301
8.4	Lattice Data and Models	308
8.5	Proportions	309
8.6	Contingency Tables	312
8.7	Generalized Linear Models	318
8.8	Summary	329
	Exercises	330
9	Design of Experiments	335
9.1	Introduction	335
9.2	Sampling Designs	335
9.3	Design of Experiments	347
9.4	Comments on Field Studies and Design	364

viii CONTENTS

9.5	Missing Data	366
9.6	Summary	367
	Exercises	368
10	Directional Data	371
10.1	Introduction	371
10.2	Circular Data	371
10.3	Spherical Data	379
10.4	Summary	386
	Exercises	387
	References	389
	Index	399