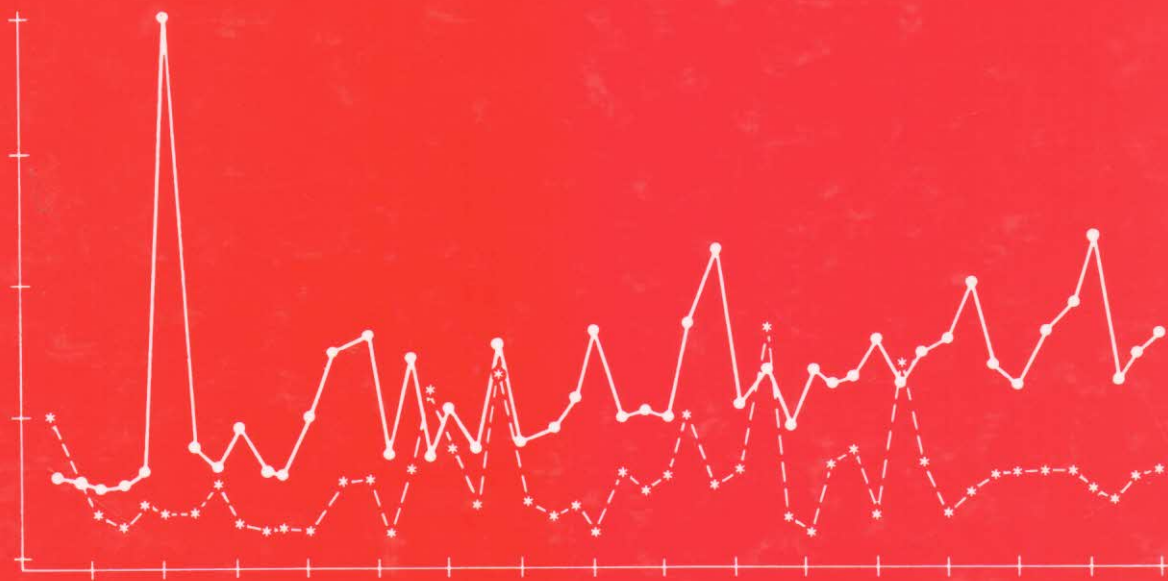


# STATISTICAL METHODS FOR ENVIRONMENTAL POLLUTION MONITORING



RICHARD O. GILBERT

# Contents

## Preface / ix

## 1 Introduction / 1

- 1.1 Types and Objectives of Environmental Pollution Studies / 1
- 1.2 Statistical Design and Analysis Problems / 2
- 1.3 Overview of the Design and Analysis Process / 3
- 1.4 Summary / 4

## 2 Sampling Environmental Populations / 5

- 2.1 Sampling in Space and Time / 5
- 2.2 Target and Sampled Populations / 7
- 2.3 Representative Units / 9
- 2.4 Choosing a Sampling Plan / 10
- 2.5 Variability and Error in Environmental Studies / 10
- 2.6 Case Study / 13
- 2.7 Summary / 15



## 3 Environmental Sampling Design / 17

- 3.1 Introduction / 17
- 3.2 Criteria for Choosing a Sampling Plan / 17
- 3.3 Methods for Selecting Sampling Locations and Times / 19
- 3.4 Summary / 24

## 4 Simple Random Sampling / 26

- 4.1 Basic Concepts / 26
- 4.2 Estimating the Mean and Total Amount / 27
- 4.3 Effect of Measurement Errors / 30
- 4.4 Number of Measurements: Independent Data / 30
- 4.5 Number of Measurements: Correlated Data / 35
- 4.6 Estimating  $\text{Var}(\bar{x})$  / 42
- 4.7 Summary / 43

## 5 Stratified Random Sampling / 45

- 5.1 Basic Concepts / 45
- 5.2 Estimating the Mean / 46
- 5.3 Estimating the Total Amount / 47
- 5.4 Arbitrary Selection of Strata / 48
- 5.5 Allocation of Samples to Strata / 50

5.6 Number of Samples / 51

5.7 Case Study / 53

5.8 Summary / 54

## **6 Two-Stage Sampling / 58**

6.1 Basic Concepts / 58

6.2 Primary Units of Equal Size / 59

6.3 Primary Units of Unequal Size / 64

6.4 Summary / 69

## **7 Compositing and Three-Stage Sampling / 71**

7.1 Basic Concepts / 72

7.2 Equal-Sized Units / 72

7.3 Unequal-Sized Units / 79

7.4 Summary / 85

## **8 Systematic Sampling / 89**

8.1 Sampling along a Line / 90

8.2 Sampling over Space / 93

8.3 Comparing Systematic with Random Sampling / 94

8.4 Estimating the Mean and Variance / 96

8.5 Populations with Trends / 99

8.6 Estimating  $\text{Var}(\bar{x})$  from a Single Systematic Sample / 100

8.7 Estimating Spatial Distributions / 102

8.8 Summary / 103

## **9 Double Sampling / 106**

9.1 Linear Regression Double Sampling / 106

9.2 Ratio Double Sampling / 111

9.3 Case Study / 112

9.4 Summary / 117

## **10 Locating Hot Spots / 119**

10.1 Determining Grid Spacing / 121

10.2 Size of Hot Spot Likely to be Hit / 125

10.3 Probability of Not Hitting a Hot Spot / 125

10.4 Taking Prior Information into Account / 127

10.5 Probability that a Hot Spot Exists When None Has Been Found / 128

10.6 Choosing the Consumer's Risk / 129

10.7 Summary / 131

## **11 Quantiles, Proportions, and Means / 132**

11.1 Basic Concepts / 132

11.2 Estimating Quantiles (Percentiles) / 134

11.3 Confidence Limits for Quantiles / 136

11.4 Estimating Proportions / 136

- 11.5 Two-Sided Confidence Limits for the Mean / 137
- 11.6 One-Sided Confidence Limits for the Mean / 138
- 11.7 Approximate Confidence Limits for the Mean / 139
- 11.8 Alternative Estimators for the Mean and Standard Deviation / 140
- 11.9 Nonparametric Estimators of Quantiles / 141
- 11.10 Nonparametric Confidence Limits for Quantiles / 141
- 11.11 Nonparametric Confidence Limits for Proportions / 142
- 11.12 Confidence Limits When Data Are Correlated / 144
- 11.13 Rank Von Neumann Test for Serial Correlation / 146
- 11.14 Data Transformations / 148
- 11.15 Summary / 149

## **12 Skewed Distributions and Goodness-of-Fit Tests / 152**

- 12.1 Lognormal Distribution / 152
- 12.2 Weibull, Gamma, and Beta Distributions / 155
- 12.3 Goodness-of-Fit Tests / 157
- 12.4 Summary / 162

## **13 Characterizing Lognormal Populations / 164**

- 13.1 Estimating the Mean and Variance / 164
- 13.2 Confidence Limits for the Mean / 169
- 13.3 Estimating the Median / 171
- 13.4 Confidence Limits for the Median / 173
- 13.5 Choosing  $n$  for Estimating the Median / 174
- 13.6 Estimating Quantiles / 174
- 13.7 Summary / 175

## **14 Estimating the Mean and Variance from Censored Data Sets / 177**

- 14.1 Data Near Detection Limits / 177
- 14.2 Estimators of the Mean and Variance / 178
- 14.3 Two-Parameter Lognormal Distribution / 181
- 14.4 Three-Parameter Lognormal Distribution / 184
- 14.5 Summary / 184

## **15 Outlier Detection and Control Charts / 186**

- 15.1 Data Screening and Validation / 186
- 15.2 Treatment of Outliers / 187
- 15.3 Rosner's Test for Detecting up to  $k$  Outliers / 188
- 15.4 Detecting Outliers in Correlated Variables / 191
- 15.5 Other Outlier Tests / 192
- 15.6 Control Charts / 193
- 15.7 Summary / 202

## **16 Detecting and Estimating Trends / 204**

- 16.1 Types of Trends / 204
- 16.2 Statistical Complexities / 205
- 16.3 Methods / 207
- 16.4 Mann-Kendall Test / 208

## *viii* Contents

16.5 Sen's Nonparametric Estimator of Slope / 217

16.6 Case Study / 219

16.7 Summary / 223

## **17 Trends and Seasonality / 225**

17.1 Seasonal Kendall Test / 225

17.2 Seasonal Kendall Slope Estimator / 227

17.3 Homogeneity of Trends in Different Seasons / 228

17.4 Sen's Test for Trend / 230

17.5 Testing for Global Trends / 231

17.6 Summary / 236

## **18 Comparing Populations / 241**

18.1 Tests Using Paired Data / 241

18.2 Independent Data Sets / 247

18.3 Summary / 252

## **Appendix A: Statistical Tables / 254**

## **Appendix B: TREND / 274**

**Symbols / 296**

**Glossary / 299**

**Bibliography / 301**

**Index / 315**