



Robert C. Elston | William D. Johnson

Basic Biostatistics

for Geneticists and Epidemiologists

A Practical Approach

 WILEY

CONTENTS

PREFACE

ix

1 INTRODUCTION: THE ROLE AND RELEVANCE OF STATISTICS, GENETICS AND EPIDEMIOLOGY IN MEDICINE

3

Why Biostatistics?

3

What Exactly Is (Are) Statistics?

5

Reasons for Understanding Statistics

6

What Exactly is Genetics?

8

What Exactly is Epidemiology?

10

How Can a Statistician Help Geneticists and Epidemiologists?

11

Disease Prevention versus Disease Therapy

12

A Few Examples: Genetics, Epidemiology and Statistical Inference

12

Summary

14

References

15

2 POPULATIONS, SAMPLES, AND STUDY DESIGN

19

The Study of Cause and Effect

19

Populations, Target Populations and Study Units

21

Probability Samples and Randomization

23

Observational Studies

25

Family Studies

27

Experimental Studies

28

Quasi-Experimental Studies

36

Summary

37

Further Reading

38

Problems

38

3 DESCRIPTIVE STATISTICS

45

Why Do We Need Descriptive Statistics?

45

Scales of Measurement

46

Tables

47

Graphs

49

Proportions and Rates

55

Relative Measures of Disease Frequency

58

Sensitivity, Specificity and Predictive Values

61

Measures of Central Tendency	62
Measures of Spread or Variability	64
Measures of Shape	67
Summary	68
Further Reading	70
Problems	70
4 THE LAWS OF PROBABILITY	79
Definition of Probability	79
The Probability of Either of Two Events: A or B	82
The Joint Probability of Two Events: A and B	83
Examples of Independence, Nonindependence and Genetic Counseling	86
Bayes' Theorem	89
Likelihood Ratio	97
Summary	98
Further Reading	99
Problems	99
5 RANDOM VARIABLES AND DISTRIBUTIONS	107
Variability and Random Variables	107
Binomial Distribution	109
A Note about Symbols	112
Poisson Distribution	113
Uniform Distribution	114
Normal Distribution	116
Cumulative Distribution Functions	119
The Standard Normal (Gaussian) Distribution	120
Summary	122
Further Reading	123
Problems	123
6 ESTIMATES AND CONFIDENCE LIMITS	131
Estimates and Estimators	131
Notation for Population Parameters, Sample Estimates, and Sample Estimators	133
Properties of Estimators	134
Maximum Likelihood	135
Estimating Intervals	137
Distribution of the Sample Mean	138
Confidence Limits	140
Summary	146
Problems	148
7 SIGNIFICANCE TESTS AND TESTS OF HYPOTHESES	155
Principle of Significance Testing	155
Principle of Hypothesis Testing	156
Testing a Population Mean	157

One-Sided versus Two-Sided Tests	160
Testing a Proportion	161
Testing the Equality of Two Variances	165
Testing the Equality of Two Means	167
Testing the Equality of Two Medians	169
Validity and Power	172
Summary	176
Further Reading	178
Problems	178
8 LIKELIHOOD RATIOS, BAYESIAN METHODS AND MULTIPLE HYPOTHESES	187
Likelihood Ratios	187
Bayesian Methods	190
Bayes' Factors	192
Bayesian Estimates and Credible Intervals	194
The Multiple Testing Problem	195
Summary	198
Problems	199
9 THE MANY USES OF CHI-SQUARE	203
The Chi-Square Distribution	203
Goodness-of-Fit Tests	206
Contingency Tables	209
Inference About the Variance	219
Combining p -Values	220
Likelihood Ratio Tests	221
Summary	223
Further Reading	225
Problems	225
10 CORRELATION AND REGRESSION	233
Simple Linear Regression	233
The Straight-Line Relationship When There is Inherent Variability	240
Correlation	242
Spearman's Rank Correlation	246
Multiple Regression	246
Multiple Correlation and Partial Correlation	250
Regression toward the Mean	251
Summary	253
Further Reading	254
Problems	255
11 ANALYSIS OF VARIANCE AND LINEAR MODELS	265
Multiple Treatment Groups	265
Completely Randomized Design with a Single Classification of Treatment Groups	267

Data with Multiple Classifications	269
Analysis of Covariance	281
Assumptions Associated with the Analysis of Variance	282
Summary	283
Further Reading	284
Problems	285
12 SOME SPECIALIZED TECHNIQUES	293
Multivariate Analysis	293
Discriminant Analysis	295
Logistic Regression	296
Analysis of Survival Times	299
Estimating Survival Curves	301
Permutation Tests	304
Resampling Methods	309
Summary	312
Further Reading	313
Problems	313
13 GUIDES TO A CRITICAL EVALUATION OF PUBLISHED REPORTS	321
The Research Hypothesis	321
Variables Studied	321
The Study Design	322
Sample Size	322
Completeness of the Data	323
Appropriate Descriptive Statistics	323
Appropriate Statistical Methods for Inferences	323
Logic of the Conclusions	324
Meta-analysis	324
Summary	326
Further Reading	327
Problems	328
EPILOGUE	329
REVIEW PROBLEMS	331
ANSWERS TO ODD-NUMBERED PROBLEMS	345
APPENDIX	353
INDEX	365