

CONTENTS

PREFACE TO THE THIRD EDITION	vi
ACKNOWLEDGMENTS	ix
CHAPTER 1. THE SCIENTIFIC METHOD	1
1.1 The Logic of Scientific Reasoning	1
1.2 Variability of Phenomena Requires Statistical Analysis	6
1.3 Inductive Inference: Statistics as the Technology of the Scientific Method	7
1.4 Design of Studies	8
1.5 How to Quantify Variables	10
1.6 The Null Hypothesis	11
1.7 Why Do We Test the Null Hypothesis?	12
1.8 Types of Errors	14
1.9 Significance Level and Types of Error	15
1.10 Consequences of Type I and Type II Errors	16
CHAPTER 2. A LITTLE BIT OF PROBABILITY	19
2.1 What Is Probability?	19
2.2 Combining Probabilities	20
2.3 Conditional Probability	23
2.4 Bayesian Probability	24
2.5 Odds and Probability	25
2.6 Likelihood Ratio	26
2.7 Summary of Probability	27
CHAPTER 3. MOSTLY ABOUT STATISTICS	29
3.1 Chi-Square for 2×2 Tables	29
3.2 McNemar Test	34
3.3 Kappa	35

3.4 Description of a Population: Use of the Standard Deviation	36
3.5 Meaning of the Standard Deviation: The Normal Distribution	40
3.6 The Difference Between Standard Deviation and Standard Error	42
3.7 Standard Error of the Difference Between Two Means	46
3.8 Z Scores and the Standardized Normal Distribution	47
3.9 The t Statistic	51
3.10 Sample Values and Population Values Revisited	52
3.11 A Question of Confidence	53
3.12 Confidence Limits and Confidence Intervals	55
3.13 Degrees of Freedom	56
3.14 Confidence Intervals for Proportions	57
3.15 Confidence Intervals Around the Difference Between Two Means	58
3.16 Comparisons Between Two Groups	60
3.17 Z-Test for Comparing Two Proportions	60
3.18 t-Test for the Difference Between Means of Two Independent Groups: Principles	62
3.19 How to Do a t-Test: An Example	64
3.20 Matched Pair t-Test	66
3.21 When Not to Do a Lot of t-Tests: The Problem of Multiple Tests of Significance	67
3.22 Analysis of Variance: Comparison Among Several Groups	69
3.23 Principles	69
3.24 Bonferroni Procedure: An Approach to Making Multiple Comparisons	72
3.25 Analysis of Variance When There Are Two Independent Variables: The Two-Factor ANOVA	74
3.26 Interaction Between Two Independent Variables	75
3.27 Example of a Two-Way ANOVA	76
3.28 Kruskal-Wallis Test to Compare Several Groups	77
3.29 Association and Causation: The Correlation Coefficient	78
3.30 How High Is High?	79

Contents	xiii
3.31 Causal Pathways	79
3.32 Regression	82
3.33 The Connection Between Linear Regression and the Correlation Coefficient	84
3.34 Multiple Linear Regression	84
3.35 Summary So Far	86
CHAPTER 4. MOSTLY ABOUT EPIDEMIOLOGY	87
4.1 The Uses of Epidemiology	87
4.2 Some Epidemiologic Concepts: Mortality Rates	88
4.3 Age-Adjusted Rates	90
4.4 Incidence and Prevalence Rates	92
4.5 Standardized Mortality Ratio	94
4.6 Person-Years of Observation	94
4.7 Dependent and Independent Variables	96
4.8 Types of Studies	96
4.9 Cross-Sectional Versus Longitudinal Looks at Data	97
4.10 Measures of Relative Risk: Inferences From Prospective Studies: the Framingham Study	101
4.11 Calculation of Relative Risk from Prospective Studies	103
4.12 Odds Ratio: Estimate of Relative Risk from Case-Control Studies	104
4.13 Attributable Risk	107
4.14 Response Bias	109
4.15 Confounding Variables	111
4.16 Matching	112
4.17 Multiple Logistic Regression	113
4.18 Confounding By Indication	116
4.19 Survival Analysis: Life Table Methods	117
4.20 Cox Proportional Hazards Model	120
4.21 Selecting Variables For Multivariate Models	122
4.22 Interactions: Additive and Multiplicative Models	124
Summary:	127

CHAPTER 5. MOSTLY ABOUT SCREENING	129
5.1 Sensitivity, Specificity, and Related Concepts	129
5.2 Cutoff Point and Its Effects on Sensitivity and Specificity	136
CHAPTER 6. MOSTLY ABOUT CLINICAL TRIALS	141
6.1 Features of Randomized Clinical Trials	141
6.2 Purposes of Randomization	143
6.3 How to Perform Randomized Assignment	144
6.4 Two-Tailed Tests Versus One-Tailed Test	145
6.5 Clinical Trial as “Gold Standard”	146
6.6 Regression Toward the Mean	147
6.7 Intention-to-Treat Analysis	150
6.8 How Large Should the Clinical Trial Be?	151
6.9 What Is Involved in Sample Size Calculation?	153
6.10 How to Calculate Sample Size for the Difference Between Two Proportions	157
6.11 How to Calculate Sample Size for Testing the Difference Between Two Means	158
CHAPTER 7. MOSTLY ABOUT QUALITY OF LIFE	161
7.1 Scale Construction	162
7.2 Reliability	162
7.3 Validity	164
7.4 Responsiveness	165
7.5 Some Potential Pitfalls	167
CHAPTER 8. MOSTLY ABOUT GENETIC EPIDEMIOLOGY	171
8.1 A New Scientific Era	171
8.2 Overview of Genetic Epidemiology	172
8.3 Twin Studies	173
8.4 Linkage and Association Studies	175
8.5 LOD Score: Linkage Statistic	178
8.6 Association Studies	179
8.7 Transmission Disequilibrium Tests (TDT)	181
8.8 Some Additional Concepts and Complexities of Genetic Studies	185

Contents	xv
CHAPTER 9. RESEARCH ETHICS AND STATISTICS	189
9.1 What does statistics have to do with it?	189
9.2 Protection of Human Research Subjects	190
9.3 Informed Consent	192
9.4 Equipoise	194
9.5 Research Integrity	194
9.6 Authorship policies	195
9.7 Data and Safety Monitoring Boards	196
9.8 Summary	196
Postscript A FEW PARTING COMMENTS ON THE IMPACT OF EPIDEMIOLOGY ON HUMAN LIVES	197
Appendix A. CRITICAL VALUES OF CHI-SQUARE, Z, AND t	199
Appendix B. FISHER’S EXACT TEST	201
Appendix C. KRUSKAL–WALLIS NONPARAMETRIC TEST TO COMPARE SEVERAL GROUPS	203
Appendix D. HOW TO CALCULATE A CORRELATION COEFFICIENT	205
Appendix E. AGE-ADJUSTMENT	207
Appendix F. CONFIDENCE LIMITS ON ODDS RATIOS	211
Appendix G. “J” OR “U” SHAPED RELATIONSHIP BETWEEN TWO VARIABLES	213
Appendix H. DETERMINING APPROPRIATENESS OF CHANGE SCORES	217
Appendix I. GENETIC PRINCIPLES	221

REFERENCES	227
SUGGESTED READINGS	233
INDEX	237