Contents

Preface to Second Edition ..... **xi**

1 The Scope of Statistics **1**
1.1 General 1
1.2 Diagrams 5
1.3 Tabulation and data processing 8
1.4 Summarizing numerical data 15
1.5 Means and other measures of location 26
1.6 Measures of variation 32

2 Probability **42**
2.1 The meaning of probability 42
2.2 Probability calculations 45
2.3 Probability distributions 49
2.4 Expectation 53
2.5 The binomial distribution 56
2.6 The Poisson distribution 60
2.7 The normal (or Gaussian) distribution 67
2.8 Bayes' theorem 73
2.9 Subjective probability 76

3 Sampling **78**
3.1 Population and sample 78
3.2 The sampling error of a mean 80
3.3 The sampling error of a proportion 84
3.4 The sampling error of a variance 85
3.5 The sampling error of a difference 88
3.6 Some other variance formulae 90

4 Statistical Inference **93**
4.1 General 93
4.2 Significance tests on a sample mean 94
4.3 Interval estimation of a mean 100
4.4 Comparison of two means 104
4.5 Inferences from variances 112
4.6 Comparison of two variances 112
4.7 Inferences from proportions 115
4.8 Comparison of two proportions 120
4.9 2×2 tables and χ^2 tests 125
4.10 Inferences from counts 132
4.11 Comparison of two counts 134
4.12 Likelihood and Bayesian methods 137

5 Regression and Correlation 141
5.1 Association 141
5.2 Linear regression 143
5.3 Correlation 150
5.4 Sampling errors in regression and correlation 152

6 The Planning of Statistical Investigations 160
6.1 General 160
6.2 The planning of surveys: estimation of population parameters 161
6.3 Surveys to investigate associations 168
6.4 The design of experiments 172
6.5 Clinical trials 175
6.6 The size of a statistical investigation 179

7 Comparison of Several Groups 186
7.1 One-way analysis of variance 186
7.2 The method of weighting 194
7.3 Components of variance 196
7.4 Multiple comparisons 200
7.5 Comparison of several proportions: the $2 \times k$ contingency table 205
7.6 General contingency tables 208
7.7 Comparison of several variances 209
7.8 Comparison of several counts: the Poisson heterogeneity test 211

8 Further Analysis of Variance 214
8.1 Two-way analysis of variance: randomized blocks 214
8.2 The simple crossover design 222
8.3 Factorial designs 227
8.4 Latin squares 239
8.5 Other incomplete designs 245
8.6 Split-unit designs 251
8.7 Nonorthogonal two-way tables: some simple cases 257

9 Further Analysis of Straight-Line Data 264
9.1 Analysis of variance applied to regression 264
9.2 Errors in both variables 270
9.3 Straight lines through the origin 272
9.4 Regression in groups 273
9.5 The analysis of covariance 282

10 Multiple Measurements 296
10.1 Multiple regression 296
10.2 Multiple regression in groups 312
10.3 Polynomial and other curvilinear regressions 318
10.4 Multiple regression in the analysis of nonorthogonal data 324
10.5 Multivariate methods 326
10.6 Time series 347
10.7 Repeated measurements and growth curves 353

11 Data Editing 358
11.1 Preliminary remarks 358
11.2 Transformations in general 358
11.3 Logarithmic and power transformations 360
11.4 Transformations for proportions 364
11.5 Normal plots 368
11.6 Outlying observations 368

12 Further Analysis of Qualitative Data 371
12.1 Introduction 371
12.2 Components of χ^2 371
12.3 Combination of 2×2 tables 380
12.4 Combination of $r \times c$ tables 385
12.5 Generalized linear models 386
12.6 Standardization 399
12.7 Goodness of fit of frequency distributions 405

13 Distribution-Free Methods 408
13.1 Introduction 408
13.2 One-sample tests for location 409
13.3 Two-sample tests for location 411
13.4 Rank correlation 417
13.5 Estimation and general comments 419

14 Survival Analysis 421
14.1 Introduction 421
14.2 Life tables 422
14.3 Follow-up studies 424
14.4 Sampling errors in the life table 427
14.5 The product-limit estimate of survival 428
14.6 The logrank test 429
14.7 Parametric models 433
14.8 Regression and proportional-hazards models 435
14.9 The subject-years method 438

15 Sequential Methods **440**
15.1 General 440
15.2 Sequential tests for binary data 442
15.3 Normal approximations 449
15.4 Group sequential plans 451
15.5 Concluding remarks 454

16 Statistical Methods in Epidemiology **455**
16.1 Introduction 455
16.2 Relative risk 456
16.3 Attributable risk 469
16.4 Diagnostic tests and screening procedures 472
16.5 Disease clustering 478

17 Biological Assay **484**
17.1 Introduction 484
17.2 Parallel-line assays 486
17.3 Slope-ratio assays 492
17.4 Quantal-response assays 495

18 Statistical Computation **499**

Appendix Tables **506**

References **527**

Author Index **539**

Subject Index **543**