Contents

PART B

INTRODUCTION .................................................. xv
11. LINEAR MODELS I ............................................. 499
   11.1. Describing the model .................................. 499
      11.1.1. Elements of linear algebra ..................... 501
      11.1.2. A Binomial-Logistic model .................... 504
      11.1.3. The general linear model ...................... 506
      11.1.4. A Normal-Identity model ....................... 507
      11.1.5. The linear predictor and the link function ... 508
      11.1.6. The density function ........................... 509
      11.1.7. The likelihood function for the Exponential family 511
      11.1.8. The Binomial-Logistic example ................ 512
   11.2. Fitting a model to the data ......................... 513
      11.2.1. The data structure ............................. 513
      11.2.2. The likelihood function and its maximum ...... 514
      11.2.3. The deviance .................................... 517
      11.2.4. Regression through the origin ................ 518
      11.2.5. Partitioning the deviance ..................... 520
   11.3. Model specification and selection ................... 522
      11.3.1. Subspaces ....................................... 522
      11.3.2. Model formulae for the linear predictor ...... 524
      11.3.3. Models for qualitative data ................... 526
      11.3.4. The factors: main effects and interaction .... 528
      11.3.5. A two-factor model for a two-way analysis of variance model ................... 530
      11.3.6. Factorial models for cross-classified data ... 533
      11.3.7. Mixed models ................................... 536
      11.3.8. Factors with ordered levels ................... 538
   11.4. Contingency tables ..................................... 540
      11.4.1. Sampling models for contingency tables ....... 540
      11.4.2. Independence models for the linear predictor 542
      11.4.3. Three-way independence models ................ 545
      11.4.4. Response and treatment factors ................ 546
   11.5. Further reading and references ....................... 547
12. LINEAR MODELS II ........................................... 549
   12.1. Least squares ........................................ 549
       12.1.1. Introductory ideas on Least squares .......... 549
       12.1.2. Least squares in 2-space ..................... 551
       12.1.3. Maximum likelihood estimation for the Normal density ............ 553
       12.1.4. Least squares on two explanatory variables .......... 554
       12.1.5. The non-orthogonal case ....................... 557
       12.1.6. Extensions to further explanatory variables .......... 560
   12.2. Repeated sampling ................................... 561
       12.2.1. Normal distribution theory .................... 561
       12.2.2. First and second moments ..................... 562
       12.2.3. One explanatory variable ...................... 563
       12.2.4. Two orthogonal explanatory variables .......... 566
       12.2.5. Non-orthogonal explanatory variables .......... 568
   12.3. Least squares and the likelihood function .............. 571
       12.3.1. Weighted least squares ....................... 571
       12.3.2. Approximating the likelihood function .......... 572
       12.3.3. Example: fitting an independence model .......... 574
       12.3.4. An Exponential example ....................... 575
       12.3.5. An iterative procedure ....................... 576
       12.3.6. Asymptotic sampling theory ................... 576
   12.4. Further reading and references ...................... 577

13. SEQUENTIAL ANALYSIS ...................................... 579
   13.1. Sequential tests of hypotheses ...................... 580
       13.1.1. The Operating Characteristic (O.C.) function .......... 580
       13.1.2. Expected sample size ....................... 581
       13.1.3. Examples of sampling inspection schemes .......... 581
       13.2. The sequential probability ratio test (SPRT) ........ 583
       13.2.1. Approximate values for the stopping boundaries ..... 583
       13.2.2. Examples ................................... 585
       13.3. The O.C. function of the SPRT ................... 588
       13.4. Expected sample size for the SPRT ............... 590
       13.5. Examples ................................... 591
       13.6. The SPRT for composite hypotheses ............... 594
       13.7. Tests involving two Binomials ................... 596
       13.7.1. Wald's method ................................ 597
       13.7.2. Armitage's method ............................ 597
       13.7.3. Discussion ................................ 600
       13.8. Other sequential procedures ..................... 600
       13.8.1. Sequential designs for the 'Two-Binomials' problem .......... 600
       13.8.2. Bayesian methods ............................ 601
   13.9. Further reading and references ..................... 601
# Contents

## 14. DISTRIBUTION-FREE METHODS

14.1. Introduction ........................................ 603
14.2. Tests based on the empirical distribution function .... 604
   14.2.1. The Kolmogorov–Smirnov test: one sample .......... 606
   14.2.2. The Kolmogorov–Smirnov test: two samples ......... 608
14.3. Tests based on order statistics .......................... 610
14.4. One-sample tests ........................................ 611
   14.4.1. Sign test ......................................... 611
   14.4.2. The Wilcoxon signed ranks test ....................... 613
14.5. Matched pairs ........................................... 615
14.6. Two-sample tests ......................................... 615
   14.6.1. The two-sample median test .......................... 616
   14.6.2. The Wilcoxon–Mann–Whitney test ..................... 618
   14.6.3. Runs ................................................ 621
14.7. Several samples ......................................... 623
   14.7.1. Median test ......................................... 623
   14.7.2. The Kruskal–Wallis test ............................. 624
14.8. Randomization tests ..................................... 626
14.9. Rank correlation measures ................................ 628
   14.9.1. Spearman's rank correlation coefficient ............. 628
   14.9.2. Kendall's rank correlation coefficient ............... 629
14.10. Further reading and references ........................... 631

## 15. BAYESIAN STATISTICS

15.1. Introduction ........................................... 633
15.2. Bayes' theorem: the discrete case ....................... 635
   15.2.1. Learning from a single data set ...................... 635
   15.2.2. Learning from several data sets ...................... 639
   15.2.3. Bayes' theorem expressed in terms of odds .......... 643
   15.2.4. Extension to an infinite list of possible models .... 644
15.3. Bayes' theorem: the continuous case ..................... 644
   15.3.1. The continuous form of Bayes' theorem ............... 644
   15.3.2. The assessment of prior densities .................... 647
      (i) Smoothing of historical data; (ii) Judgemental curve
          fitting
   15.3.3. An illustration of Bayes' theorem for a single unknown
          parameter ............................................. 652
   15.3.4. An illustration of Bayes' theorem for two unknown
          parameters ............................................ 655
   15.3.5. Approximate analysis under great prior uncertainty .. 657
15.4. Bayesian approaches to typical statistical questions .... 665
   15.4.1. Point estimation .................................... 665
   15.4.2. Interval estimation .................................. 666
   15.4.3. Significance testing .................................. 668
   15.4.4. Prediction .......................................... 669
15.4.5. Summarizing data: sufficient statistics  669
15.4.6. The likelihood principle  670
15.5. Bayesian inference for some univariate probability models  672
  15.5.1. Inferences for Binomial and related distributions  672
  15.5.2. Inferences for the Poisson distribution  678
  15.5.3. Inferences for the Normal distribution  681
15.6. Bayes methods when models contain many parameters  685
  15.6.1. Inappropriateness of 'completely vague' prior specifications  685
  15.6.2. Simple examples  685
15.7. Further reading and references  687

16. MULTIVARIATE ANALYSIS: CLASSICAL METHODS  689
  16.1. Introduction  689
  16.2. Samples from multivariate Normal (MVN) distributions  695
    16.2.1. Maximum likelihood estimation  696
    16.2.2. Some sampling distributions  698
    16.2.3. Tests and confidence regions for the expectation vector  700
    16.2.4. Two-sample problems  701
    16.2.5. Inferences about the correlation coefficient  702
    16.2.6. Non-central distributions  704
  16.3. Principal components  704
    16.3.1. Introduction  704
    16.3.2. Population principal components  705
    16.3.3. Sample principal components  707
    16.3.4. Numerical example  708
    16.3.5. Some sampling distributions  710
  16.4. Factor Analysis  710
    16.4.1. Introduction  710
    16.4.2. Factor model  711
    16.4.3. Some properties  711
    16.4.4. Estimation  713
    16.4.5. Discussion  714
  16.5. Canonical correlation  715
    16.5.1. Introduction  715
    16.5.2. Population canonical correlations  715
    16.5.3. Sample canonical correlations  718
    16.5.4. Numerical example  719
  16.6. Discriminant analysis  720
    16.6.1. Introduction  720
    16.6.2. Discrimination in two known populations  720
    16.6.3. Discrimination in two multivariate populations  721
    16.6.4. Discrimination in several populations  723
  16.7. Further reading and references  724
### 17. MULTIVARIATE ANALYSIS: ORDINATION MULTIDIMENSIONAL SCALING AND ALLIED TOPICS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.1</td>
<td>Introduction</td>
<td>727</td>
</tr>
<tr>
<td>17.2</td>
<td>Principal components analysis</td>
<td>729</td>
</tr>
<tr>
<td>17.3</td>
<td>Multiplicative models and the Eckhart–Young theorem</td>
<td>733</td>
</tr>
<tr>
<td>17.4</td>
<td>Biplots</td>
<td>735</td>
</tr>
<tr>
<td>17.5</td>
<td>Correspondence analysis</td>
<td>739</td>
</tr>
<tr>
<td>17.6</td>
<td>Metric scaling: principal coordinates analysis and classical scaling</td>
<td>741</td>
</tr>
<tr>
<td>17.7</td>
<td>Metric scaling: other methods</td>
<td>746</td>
</tr>
<tr>
<td>17.8</td>
<td>Non-metric multidimensional scaling</td>
<td>751</td>
</tr>
<tr>
<td>17.9</td>
<td>Multidimensional unfolding</td>
<td>754</td>
</tr>
<tr>
<td>17.10</td>
<td>Orthogonal Procrustes analysis</td>
<td>761</td>
</tr>
<tr>
<td>17.11</td>
<td>The general comparison of scalings</td>
<td>764</td>
</tr>
<tr>
<td>17.12</td>
<td>Three-way scaling</td>
<td>771</td>
</tr>
<tr>
<td>17.13</td>
<td>The analysis of asymmetry</td>
<td>776</td>
</tr>
<tr>
<td>17.14</td>
<td>Further reading and references</td>
<td>779</td>
</tr>
</tbody>
</table>

### 18. TIME SERIES

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.1</td>
<td>Introduction</td>
<td>783</td>
</tr>
<tr>
<td>18.2</td>
<td>Classical regression models for time series</td>
<td>784</td>
</tr>
<tr>
<td>18.2.1</td>
<td>Seasonal effects model</td>
<td>784</td>
</tr>
<tr>
<td>18.2.2</td>
<td>Cyclical component model</td>
<td>786</td>
</tr>
<tr>
<td>18.3</td>
<td>The periodogram of a time series</td>
<td>790</td>
</tr>
<tr>
<td>18.3.1</td>
<td>Harmonic components of periodic time series</td>
<td>790</td>
</tr>
<tr>
<td>18.3.2</td>
<td>Harmonic regression analysis</td>
<td>792</td>
</tr>
<tr>
<td>18.3.3</td>
<td>The periodogram</td>
<td>793</td>
</tr>
<tr>
<td>18.3.4</td>
<td>Effect of mean correction on the periodogram</td>
<td>794</td>
</tr>
<tr>
<td>18.3.5</td>
<td>Periodograms for simple models</td>
<td>795</td>
</tr>
<tr>
<td>18.3.6</td>
<td>The interpretation of periodograms</td>
<td>797</td>
</tr>
<tr>
<td>18.4</td>
<td>Differencing operations</td>
<td>798</td>
</tr>
<tr>
<td>18.5</td>
<td>Stationary time series</td>
<td>800</td>
</tr>
<tr>
<td>18.5.1</td>
<td>The autocorrelation function (acf)</td>
<td>800</td>
</tr>
<tr>
<td>18.5.2</td>
<td>The relationship between the periodogram and the acf</td>
<td>803</td>
</tr>
<tr>
<td>18.5.3</td>
<td>The spectrum and sample spectrum</td>
<td>804</td>
</tr>
<tr>
<td>18.5.4</td>
<td>Properties of the spectrum and acf</td>
<td>805</td>
</tr>
<tr>
<td>18.5.5</td>
<td>The partial autocorrelation function (pacf)</td>
<td>806</td>
</tr>
<tr>
<td>18.5.6</td>
<td>The sample pacf</td>
<td>808</td>
</tr>
<tr>
<td>18.6</td>
<td>The general linear model (GLM)</td>
<td>811</td>
</tr>
<tr>
<td>18.6.1</td>
<td>Definition and properties</td>
<td>811</td>
</tr>
<tr>
<td>18.6.2</td>
<td>Linear operations on time series</td>
<td>811</td>
</tr>
<tr>
<td>18.6.3</td>
<td>Linear operations and the spectrum</td>
<td>813</td>
</tr>
<tr>
<td>18.6.4</td>
<td>Linear operations on a finite sample</td>
<td>815</td>
</tr>
<tr>
<td>18.6.5</td>
<td>Constraints on the GLM</td>
<td>816</td>
</tr>
<tr>
<td>18.6.6</td>
<td>Prediction using the GLM</td>
<td>818</td>
</tr>
</tbody>
</table>
### 18.7. The moving average (MA) model

18.7.1. Model definition
18.7.2. Characteristic property of the MA(q) model
18.7.3. Efficient estimation for MA(q) models
18.7.4. Forecasting using MA models
18.7.5. The EWMA predictor
18.7.6. Box–Jenkins seasonal model for the airline series

### 18.8. The autoregressive (AR) model

18.8.1. Model definition
18.8.2. First and second order examples
18.8.3. Characteristic properties of the AR(p) model
18.8.4. Efficient estimation for the AR(p) model
18.8.5. Forecasting using AR(p) models

### 18.9. Autoregressive moving average (ARMA) models

18.9.1. Model definition and properties
18.9.2. An example

### 18.10. Spectrum estimation

18.10.1. The difficulties
18.10.2. Direct methods of estimation
18.10.3. Indirect methods of estimation
18.10.4. The use of spectral analysis

### 18.11. Time series regression models

### 18.12. Further reading and references

### 19. DECISION THEORY

19.1. Basic ideas
19.1.1. Mathematical framework
19.1.2. Minimax and Bayes decision rules
19.1.3. Admissible decisions
19.1.4. Geometric interpretation
19.1.5. Some basic theorems

19.2. Statistics and decision theory
19.2.1. Estimation
19.2.2. Tests of simple hypotheses versus simple hypotheses
19.2.3. The Neyman–Pearson lemma

19.3. Risk attitudes and utility theory
19.3.1. Risk aversion
19.3.2. One-dimensional utility functions
19.3.3. Assessment of utility functions
19.3.4. Higher-dimensional utility functions

19.4. Sequential decisions
19.4.1. Basic ideas
19.4.2. Decision trees

19.5. Axiomatic approaches
19.5.1. Coherence axioms for decision-making . 891
19.5.2. Degrees of belief as probabilities . 892
19.6. Further reading and references . 896

20. KALMAN FILTERING . 897
20.1. Historical background . 897
20.2. State-space models . 899
20.2.1. System and measurement equations . 899
20.2.2. A simple example . 900
20.3. Derivation of the Kalman filter for a discrete linear dynamic system . 902
20.3.1. Assumptions . 902
20.3.2. The Kalman gain . 902
20.3.3. Optimum choice of the Kalman gain . 904
20.3.4. State prediction . 905
20.3.5. Interpretation of filter equations . 905
20.3.6. Summary . 906
20.4. Test of filter performance . 907
20.4.1. Innovation properties . 907
20.4.2. Filter divergence . 908
20.5. Kalman filtering for continuous time and non-linear systems . 909
20.5.1. Continuous time systems . 909
20.5.2. Non-linear systems: the extended Kalman filter . 910
20.6. Model identification and parameter estimation . 912
20.6.1. General . 912
20.6.2. Model identification . 915
20.6.3. Parameter estimation . 919
   (a) Linear estimation; (b) Non-linear estimation
20.7. Applications . 922
20.7.1. An application of the discrete Linear Kalman filter . 922
   (a) Introduction; (b) identification and parameter estimation; (c) Results
20.7.2. An application of the extended Kalman filter . 929
   (a) The hydrological model; (b) State-space formulation; (c) Parameter estimation; (d) Results
20.8. Concluding remarks . 936
20.9. Further reading and references . 936

Bibliography . 939
Appendix . A1
Index . xxii

Chapters 1–10 are contained in Part A.