

# Contents

## PART B

<i>INTRODUCTION</i>	xv
<i>11. LINEAR MODELS I</i>	499
<i>11.1. Describing the model</i>	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
<i>11.2. Fitting a model to the data</i>	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
<i>11.3. Model specification and selection</i>	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
<i>11.4. Contingency tables</i>	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
<i>11.5. Further reading and references</i>	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

<b>14. DISTRIBUTION-FREE METHODS</b>	603
14.1. <i>Introduction</i>	603
14.2. <i>Tests based on the empirical distribution function</i>	604
14.2.1. The Kolmogorov–Smirnov test: one sample	606
14.2.2. The Kolmogorov–Smirnov test: two samples	608
14.3. <i>Tests based on order statistics</i>	610
14.4. <i>One-sample tests</i>	611
14.4.1. Sign test	611
14.4.2. The Wilcoxon signed ranks test	613
14.5. <i>Matched pairs</i>	615
14.6. <i>Two-sample tests</i>	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. <i>Several samples</i>	623
14.7.1. Median test	623
14.7.2. The Kruskal–Wallis test	624
14.8. <i>Randomization tests</i>	626
14.9. <i>Rank correlation measures</i>	628
14.9.1. Spearman’s rank correlation coefficient	628
14.9.2. Kendall’s rank correlation coefficient	629
14.10. <i>Further reading and references</i>	631
<b>15. BAYESIAN STATISTICS</b>	633
15.1. <i>Introduction</i>	633
15.2. <i>Bayes’ theorem: the discrete case</i>	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. <i>Bayes’ theorem: the continuous case</i>	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. <i>Bayesian approaches to typical statistical questions</i>	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. <i>Bayesian inference for some univariate probability models</i>	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. <i>Bayes methods when models contain many parameters</i>	685
15.6.1. Inappropriateness of 'completely vague' prior specifications . . . . .	685
15.6.2. Simple examples . . . . .	685
15.7. <i>Further reading and references</i> . . . . .	687
16. <i>MULTIVARIATE ANALYSIS: CLASSICAL METHODS</i>	689
16.1. <i>Introduction</i> . . . . .	689
16.2. <i>Samples from multivariate Normal (MVN) distributions</i>	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. <i>Principal components</i>	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. <i>Factor Analysis</i>	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. <i>Canonical correlation</i>	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. <i>Discriminant analysis</i>	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. <i>Further reading and references</i> . . . . .	724

<b>17. MULTIVARIATE ANALYSIS: ORDINATION MULTI-DIMENSIONAL SCALING AND ALLIED TOPICS</b>	<b>727</b>
17.1 Introduction	727
17.2 Principal components analysis	729
17.3 Multiplicative models and the Eckhart–Young theorem	733
17.4 Biplots	735
17.5 Correspondence analysis	739
17.6 Metric scaling: principal coordinates analysis and classical scaling	741
17.7 Metric scaling: other methods	746
17.8 Non-metric multidimensional scaling	751
17.9 Multidimensional unfolding	754
17.10 Orthogonal Procrustes analysis	761
17.11 The general comparison of scalings	764
17.12 Three-way scaling	771
17.13 The analysis of asymmetry	776
17.14 Further reading and references	779
<b>18. TIME SERIES</b>	<b>783</b>
18.1 Introduction	783
18.2 Classical regression models for time series	784
18.2.1. Seasonal effects model	784
18.2.2. Cyclical component model	786
18.3 The periodogram of a time series	790
18.3.1. Harmonic components of periodic time series	790
18.3.2. Harmonic regression analysis	792
18.3.3. The periodogram	793
18.3.4. Effect of mean correction on the periodogram	794
18.3.5. Periodograms for simple models	795
18.3.6. The interpretation of periodograms	797
18.4 Differencing operations	798
18.5 Stationary time series	800
18.5.1. The autocorrelation function (acf)	800
18.5.2. The relationship between the periodogram and the acf	803
18.5.3. The spectrum and sample spectrum	804
18.5.4. Properties of the spectrum and acf	805
18.5.5. The partial autocorrelation function (pacf)	806
18.5.6. The sample pacf	808
18.6 The general linear model (GLM)	811
18.6.1. Definition and properties	811
18.6.2. Linear operations on time series	811
18.6.3. Linear operations and the spectrum	813
18.6.4. Linear operations on a finite sample	815
18.6.5. Constraints on the GLM	816
18.6.6. Prediction using the GLM	818

18.7. <i>The moving average (MA) model</i>	818
18.7.1. Model definition	818
18.7.2. Characteristic property of the MA( $q$ ) model	819
18.7.3. Efficient estimation for MA( $q$ ) models	820
18.7.4. Forecasting using MA models	824
18.7.5. The EWMA predictor	826
18.7.6. Box–Jenkins seasonal model for the airline series	827
18.8. <i>The autoregressive (AR) model</i>	829
18.8.1. Model definition	829
18.8.2. First and second order examples	830
18.8.3. Characteristic properties of the AR( $p$ ) model	831
18.8.4. Efficient estimation for the AR( $p$ ) model	832
18.8.5. Forecasting using AR( $p$ ) models	834
18.9. <i>Autoregressive moving average (ARMA) models</i>	835
18.9.1. Model definition and properties	835
18.9.2. An example	836
18.10. <i>Spectrum estimation</i>	837
18.10.1. The difficulties	837
18.10.2. Direct methods of estimation	837
18.10.3. Indirect methods of estimation	840
18.10.4. The use of spectral analysis	841
18.11. <i>Time series regression models</i>	842
18.12. <i>Further reading and references</i>	845

19. <i>DECISION THEORY</i>	847
19.1. <i>Basic ideas</i>	847
19.1.1. Mathematical framework	847
19.1.2. Minimax and Bayes decision rules	849
19.1.3. Admissible decisions	856
19.1.4. Geometric interpretation	857
19.1.5. Some basic theorems	863
19.2. <i>Statistics and decision theory</i>	868
19.2.1. Estimation	868
19.2.2. Tests of simple hypotheses versus simple hypotheses	870
19.2.3. The Neyman–Pearson lemma	871
19.3. <i>Risk attitudes and utility theory</i>	873
19.3.1. Risk aversion	873
19.3.2. One-dimensional utility functions	879
19.3.3. Assessment of utility functions	882
19.3.4. Higher-dimensional utility functions	883
19.4. <i>Sequential decisions</i>	884
19.4.1. Basic ideas	884
19.4.2. Decision trees	886
19.5. <i>Axiomatic approaches</i>	891

19.5.1. Coherence axioms for decision-making . . . . .	891
19.5.2. Degrees of belief as probabilities . . . . .	892
19.6. <i>Further reading and references</i> . . . . .	896
<b>20. KALMAN FILTERING</b> . . . . .	897
20.1. <i>Historical background</i> . . . . .	897
20.2. <i>State-space models</i> . . . . .	899
20.2.1. System and measurement equations . . . . .	899
20.2.2. A simple example . . . . .	900
20.3. <i>Derivation of the Kalman filter for a discrete linear dynamic system</i> . . . . .	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. <i>Test of filter performance</i> . . . . .	907
20.4.1. Innovation properties . . . . .	907
20.4.2. Filter divergence . . . . .	908
20.5. <i>Kalman filtering for continuous time and non-linear systems</i> . . . . .	909
20.5.1. Continuous time systems . . . . .	909
20.5.2. Non-linear systems: the extended Kalman filter . . . . .	910
20.6. <i>Model identification and parameter estimation</i> . . . . .	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. <i>Applications</i> . . . . .	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 formula- tion; (c) Parameter estimation; (d) Results	
20.8. <i>Concluding remarks</i> . . . . .	936
20.9. <i>Further reading and references</i> . . . . .	936
<b>Bibliography</b> . . . . .	939
<b>Appendix</b> . . . . .	A1
<b>Index</b> . . . . .	xxi

Chapters 1–10 are contained in Part A.