

Contents

1	Introduction	1
1.1	Examples	2
1.1.1	Inference for a Proportion	2
1.1.2	Comparison of Proportions	2
1.1.3	The Capture–Recapture Method	4
1.1.4	Hardy–Weinberg Equilibrium	4
1.1.5	Estimation of Diagnostic Tests Characteristics	5
1.1.6	Quantifying Disease Risk from Cancer Registry Data	6
1.1.7	Predicting Blood Alcohol Concentration	8
1.1.8	Analysis of Survival Times	8
1.2	Statistical Models	9
1.3	Contents and Notation of the Book	11
1.4	References	11
2	Likelihood	13
2.1	Likelihood and Log-Likelihood Function	13
2.1.1	Maximum Likelihood Estimate	14
2.1.2	Relative Likelihood	22
2.1.3	Invariance of the Likelihood	23
2.1.4	Generalised Likelihood	26
2.2	Score Function and Fisher Information	27
2.3	Numerical Computation of the Maximum Likelihood Estimate	31
2.3.1	Numerical Optimisation	31
2.3.2	The EM Algorithm	34
2.4	Quadratic Approximation of the Log-Likelihood Function	37
2.5	Sufficiency	40
2.5.1	Minimal Sufficiency	45
2.5.2	The Likelihood Principle	47
2.6	Exercises	48
2.7	Bibliographic Notes	50
3	Elements of Frequentist Inference	51
3.1	Unbiasedness and Consistency	51
3.2	Standard Error and Confidence Interval	55
3.2.1	Standard Error	56

3.2.2	Confidence Interval	56
3.2.3	Pivots	59
3.2.4	The Delta Method	63
3.2.5	The Bootstrap	65
3.3	Significance Tests and P -Values	70
3.4	Exercises	75
3.5	References	78
4	Frequentist Properties of the Likelihood	79
4.1	The Expected Fisher Information and the Score Statistic	80
4.1.1	The Expected Fisher Information	81
4.1.2	Properties of the Expected Fisher Information	84
4.1.3	The Score Statistic	87
4.1.4	The Score Test	89
4.1.5	Score Confidence Intervals	91
4.2	The Distribution of the ML Estimator and the Wald Statistic	94
4.2.1	Cramér–Rao Lower Bound	95
4.2.2	Consistency of the ML Estimator	96
4.2.3	The Distribution of the ML Estimator	97
4.2.4	The Wald Statistic	99
4.3	Variance Stabilising Transformations	101
4.4	The Likelihood Ratio Statistic	105
4.4.1	The Likelihood Ratio Test	106
4.4.2	Likelihood Ratio Confidence Intervals	106
4.5	The p^* Formula	112
4.6	A Comparison of Likelihood-Based Confidence Intervals	113
4.7	Exercises	119
4.8	References	122
5	Likelihood Inference in Multiparameter Models	123
5.1	Score Vector and Fisher Information Matrix	124
5.2	Standard Error and Wald Confidence Interval	128
5.3	Profile Likelihood	130
5.4	Frequentist Properties of the Multiparameter Likelihood	143
5.4.1	The Score Statistic	144
5.4.2	The Wald Statistic	145
5.4.3	The Multivariate Delta Method	146
5.4.4	The Likelihood Ratio Statistic	146
5.5	The Generalised Likelihood Ratio Statistic	148
5.6	Conditional Likelihood	153
5.7	Exercises	155
5.8	References	165
6	Bayesian Inference	167
6.1	Bayes' Theorem	168
6.2	Posterior Distribution	170
6.3	Choice of the Prior Distribution	179

6.3.1	Conjugate Prior Distributions	179
6.3.2	Improper Prior Distributions	183
6.3.3	Jeffreys' Prior Distributions	185
6.4	Properties of Bayesian Point and Interval Estimates	192
6.4.1	Loss Function and Bayes Estimates	192
6.4.2	Compatible and Invariant Bayes Estimates	195
6.5	Bayesian Inference in Multiparameter Models	196
6.5.1	Conjugate Prior Distributions	196
6.5.2	Jeffreys' and Reference Prior Distributions	198
6.5.3	Elimination of Nuisance Parameters	200
6.5.4	Compatibility of Uni- and Multivariate Point Estimates	204
6.6	Some Results from Bayesian Asymptotics	204
6.6.1	Discrete Asymptotics	205
6.6.2	Continuous Asymptotics	206
6.7	Empirical Bayes Methods	209
6.8	Exercises	214
6.9	References	219
7	Model Selection	221
7.1	Likelihood-Based Model Selection	224
7.1.1	Akaike's Information Criterion	224
7.1.2	Cross Validation and AIC	227
7.1.3	Bayesian Information Criterion	230
7.2	Bayesian Model Selection	231
7.2.1	Marginal Likelihood and Bayes Factor	232
7.2.2	Marginal Likelihood and BIC	236
7.2.3	Deviance Information Criterion	239
7.2.4	Model Averaging	240
7.3	Exercises	243
7.4	References	245
8	Numerical Methods for Bayesian Inference	247
8.1	Standard Numerical Techniques	248
8.2	Laplace Approximation	253
8.3	Monte Carlo Methods	258
8.3.1	Monte Carlo Integration	258
8.3.2	Importance Sampling	265
8.3.3	Rejection Sampling	267
8.4	Markov Chain Monte Carlo	269
8.5	Numerical Calculation of the Marginal Likelihood	280
8.5.1	Calculation Through Numerical Integration	280
8.5.2	Monte Carlo Estimation of the Marginal Likelihood	282
8.6	Exercises	286
8.7	References	289

9	Prediction	291
9.1	Plug-in Prediction	292
9.2	Likelihood Prediction	292
9.2.1	Predictive Likelihood	293
9.2.2	Bootstrap Prediction	295
9.3	Bayesian Prediction	299
9.3.1	Posterior Predictive Distribution	299
9.3.2	Computation of the Posterior Predictive Distribution	303
9.3.3	Model Averaging	305
9.4	Assessment of Predictions	306
9.4.1	Discrimination and Calibration	307
9.4.2	Scoring Rules	311
9.5	Exercises	315
9.6	References	316
Appendix A	Probabilities, Random Variables and Distributions	317
A.1	Events and Probabilities	318
A.1.1	Conditional Probabilities and Independence	318
A.1.2	Bayes' Theorem	319
A.2	Random Variables	319
A.2.1	Discrete Random Variables	319
A.2.2	Continuous Random Variables	320
A.2.3	The Change-of-Variables Formula	321
A.2.4	Multivariate Normal Distributions	323
A.3	Expectation, Variance and Covariance	324
A.3.1	Expectation	324
A.3.2	Variance	325
A.3.3	Moments	325
A.3.4	Conditional Expectation and Variance	325
A.3.5	Covariance	326
A.3.6	Correlation	327
A.3.7	Jensen's Inequality	328
A.3.8	Kullback–Leibler Discrepancy and Information Inequality	329
A.4	Convergence of Random Variables	329
A.4.1	Modes of Convergence	329
A.4.2	Continuous Mapping and Slutsky's Theorem	330
A.4.3	Law of Large Numbers	330
A.4.4	Central Limit Theorem	331
A.4.5	Delta Method	331
A.5	Probability Distributions	332
A.5.1	Univariate Discrete Distributions	333
A.5.2	Univariate Continuous Distributions	335
A.5.3	Multivariate Distributions	339

Appendix B	Some Results from Matrix Algebra and Calculus	341
B.1	Some Matrix Algebra	341
B.1.1	Trace, Determinant and Inverse	341
B.1.2	Cholesky Decomposition	343
B.1.3	Inversion of Block Matrices	344
B.1.4	Sherman–Morrison Formula	345
B.1.5	Combining Quadratic Forms	345
B.2	Some Results from Mathematical Calculus	345
B.2.1	The Gamma and Beta Functions	345
B.2.2	Multivariate Derivatives	346
B.2.3	Taylor Approximation	347
B.2.4	Leibniz Integral Rule	348
B.2.5	Lagrange Multipliers	348
B.2.6	Landau Notation	349
Appendix C	Some Numerical Techniques	351
C.1	Optimisation and Root Finding Algorithms	351
C.1.1	Motivation	351
C.1.2	Bisection Method	352
C.1.3	Newton–Raphson Method	354
C.1.4	Secant Method	357
C.2	Integration	358
C.2.1	Newton–Cotes Formulas	358
C.2.2	Laplace Approximation	361
Notation		363
References		367
Index		371