

Bayes and Empirical Bayes Methods for Data Analysis

BRADLEY P. CARLIN

*Associate Professor,
Division of Biostatistics,
School of Public Health,
University of Minnesota,
Minneapolis, USA*

and

THOMAS A. LOUIS

*Professor and Head of Department,
Division of Biostatistics,
School of Public Health,
University of Minnesota,
Minneapolis, USA*



CHAPMAN & HALL

London · Weinheim · New York · Tokyo · Melbourne · Madras

Contents

Preface	xiii
1 Procedures and their properties	1
1.1 Introduction and motivation	1
1.2 Structures for inference	6
1.3 The general decision problem	7
1.3.1 Risk and admissibility	9
1.3.2 Unbiased rules	10
1.3.3 Bayes rules	12
1.3.4 Minimax rules	14
1.4 Procedure evaluation and other unifying concepts	14
1.4.1 Mean squared error	15
1.4.2 The variance-bias tradeoff	15
1.5 Exercises	17
2 The Bayes approach	21
2.1 Introduction	21
2.2 Prior distributions	26
2.2.1 Elicited priors	27
2.2.2 Conjugate priors	30
2.2.3 Noninformative priors	33
2.2.4 Other prior construction methods	37
2.3 Bayesian inference	38
2.3.1 Point estimation	39
2.3.2 Interval estimation	42
2.3.3 Hypothesis testing and Bayes factors	45
2.3.4 Example: Consumer preference data	50
2.4 Model assessment	54
2.4.1 Diagnostic measures	54
2.4.2 Model averaging	57
2.5 Nonparametric methods	59

2.6	Exercises	62
3	The empirical Bayes approach	67
3.1	Introduction	67
3.2	Nonparametric EB point estimation	68
3.2.1	Compound sampling models	68
3.2.2	Simple NPEB (Robbins' method)	69
3.2.3	Estimation of the prior (NPML method)	71
3.2.4	Example: Accident data	72
3.3	Parametric EB point estimation	74
3.3.1	Gaussian/Gaussian models	75
3.3.2	Beta/binomial model	80
3.3.3	EB performance of the PEB	82
3.3.4	Stein estimation	84
3.4	Computation via the EM algorithm	89
3.5	Interval estimation	93
3.5.1	Morris' approach	95
3.5.2	Marginal posterior approach	95
3.5.3	Bias correction approach	98
3.6	Generalization to regression structures	100
3.7	Exercises	102
4	Performance of Bayes procedures	107
4.1	Bayesian processing	108
4.1.1	Univariate stretching with a two-point prior	108
4.1.2	Multivariate Gaussian model	109
4.2	Frequentist performance: Point estimates	110
4.2.1	Gaussian/Gaussian model	111
4.2.2	Beta/binomial model	113
4.2.3	Generalization	116
4.3	Frequentist performance: Confidence intervals	117
4.3.1	Beta/binomial model	117
4.3.2	Fieller-Creasy problem	124
4.4	Empirical Bayes performance	128
4.4.1	Point estimation	129
4.4.2	Interval estimation	132
4.5	Design of experiments	135
4.5.1	Bayesian design for Bayesian analysis	135
4.5.2	Bayesian design for frequentist analysis	137
4.6	Exercises	139

5 Bayesian computation	141
5.1 Introduction	141
5.2 Asymptotic methods	142
5.2.1 Normal approximation	142
5.2.2 Laplace's method	145
5.3 Noniterative Monte Carlo methods	150
5.3.1 Direct sampling	150
5.3.2 Indirect methods	152
5.4 Markov chain Monte Carlo methods	159
5.4.1 Substitution sampling and data augmentation	159
5.4.2 Gibbs sampling	163
5.4.3 Metropolis-Hastings algorithm	173
5.4.4 Hybrid forms and other algorithms	180
5.4.5 Convergence monitoring and acceleration	185
5.5 Exercises	197
6 Model criticism and selection	205
6.1 Bayesian robustness	206
6.1.1 Sensitivity analysis	206
6.1.2 Prior partitioning	212
6.2 Model assessment	218
6.3 Bayes factors	220
6.3.1 Direct methods	221
6.3.2 Marginal densities from the Gibbs sampler	223
6.3.3 Sampling over model and parameter space	225
6.3.4 Sampling over the model space alone	229
6.4 Predictive model selection	230
6.5 Exercises	233
7 Special methods and models	237
7.1 Ensemble estimates	237
7.2 Order restricted inference	246
7.3 Nonlinear models	247
7.4 Longitudinal data models	251
7.5 Continuous and categorical time series	253
7.6 Survival analysis and frailty models	255
7.6.1 Statistical models	256
7.6.2 Treatment effect prior determination	257
7.6.3 Computation and advanced models	258
7.7 Spatial and spatio-temporal models	260

7.8	Exercises	265
8	Case studies	273
8.1	Analysis of longitudinal AIDS data	274
8.1.1	Introduction and background	274
8.1.2	Modeling of longitudinal CD4 counts	275
8.1.3	CD4 response to treatment at two months	285
8.1.4	Survival analysis	287
8.1.5	Discussion	289
8.2	Robust analysis of clinical trials	289
8.2.1	Clinical background	289
8.2.2	Interim monitoring	290
8.2.3	Prior robustness and prior scoping	296
8.2.4	Discussion	302
8.3	Spatio-temporal mapping of lung cancer rates	303
8.3.1	Introduction	303
8.3.2	Data and model description	305
8.3.3	Computational considerations	307
8.3.4	Model fitting, validation and comparison	308
8.3.5	Discussion	315
	Appendices	317
A	Distributional catalog	319
A.1	Discrete	319
A.1.1	Univariate	319
A.1.2	Multivariate	321
A.2	Continuous	321
A.2.1	Univariate	321
A.2.2	Multivariate	325
B	Software guide	327
B.1	Prior elicitation	328
B.2	Random effects models/Empirical Bayes analysis	330
B.3	Bayesian analysis	334
B.3.1	Early developments	335
B.3.2	Special purpose programs	336
B.3.3	Teaching programs	341
B.3.4	Markov chain Monte Carlo programs	342
C	Answers to selected exercises	349

CONTENTS	xi
References	363
Author index	387
Subject index	393