

Contents

1	Overview of the Book	1
2	Introduction: Heterogeneity in Medicine	7
2.1	Example: Plasma Concentration of Beta-Carotene	11
2.1.1	Identification of a Latent Structure	11
2.1.2	Including Covariates	13
2.2	Computation	15
2.3	Example: Analysis of Heterogeneity in Drug Development	18
2.3.1	Basic Pharmacokinetic Concepts	18
2.3.2	Pharmacokinetic Parameters	19
2.3.3	First-Order Compartment Models	20
2.3.4	Population Pharmacokinetics	21
2.3.5	Theophylline Pharmacokinetics	22
2.4	A Note of Caution	28
3	Modeling Count Data	29
3.1	Example: Morbidity in Northeast Thailand	29
3.2	Parametric Mixture Models	30
3.3	Finite Mixture Models	33
3.3.1	Diagnostic Plots for Finite Mixture Models	34
3.3.2	A Finite Mixture Model for the Illness Spell Data	34
3.3.3	Estimating the Number of Components	36
3.4	Computation	39
3.4.1	Combination of VEM and EM Algorithms	39
3.4.2	Using the EM Algorithm	41
3.4.3	Estimating the Number of Components	42
3.5	Including Covariates	43
3.5.1	The Ames Test	43
3.5.2	Poisson and Negative Binomial Regression Models	47
3.5.3	Covariate-Adjusted Mixture Model for the Ames Test Data	49

3.6	Computation	51
3.6.1	Fitting Poisson and Negative Binomial Regression Models with SAS	51
3.6.2	Fitting Poisson and Negative Binomial Regression Models with R	52
3.6.3	Fitting Finite Mixture Models with the Package CAMAN	53
4	Theory and Algorithms	55
4.1	The Likelihood of Finite Mixture Models	55
4.2	Convex Geometry and Optimization	56
4.2.1	Derivatives and Directional Derivatives of Convex Functions	61
4.3	Application to the Flexible Support Size Case	64
4.3.1	Geometric Characterization	64
4.3.2	Algorithms for Flexible Support Size	69
4.3.3	VEM Algorithm: Computation	70
4.4	The Fixed Support Size Case	72
4.4.1	Fixed Support Size: The Newton–Raphson Algorithm	72
4.4.2	A General Description of the EM Algorithm	73
4.4.3	The EM Algorithm for Finite Mixture Models	74
4.4.4	EM Algorithm: Computation	77
4.4.5	A Hybrid Mixture Algorithm	79
4.4.6	The EM Algorithm with Gradient Update	80
4.5	Estimating the Number of Components	82
4.5.1	Graphical Techniques	82
4.5.2	Testing for the Number of Components	83
4.5.3	The Bootstrap Approach	84
4.6	Adjusting for Covariates	87
4.6.1	Generalized Linear Models	87
4.6.2	The EM Algorithm for Covariate-Adjusted Mixture Models	91
4.6.3	Computation: Vitamin A Supplementation Revisited	93
4.6.4	An Extension of the EM Algorithm with Gradient Update for Covariate-Adjusted Mixture Models	95
4.7	Case Study: EM Algorithm with Gradient Update for Nonlinear Finite Mixture Models	97
4.7.1	Introduction	97
4.7.2	Example: Dipyrene Pharmacokinetics	98
4.7.3	First-Order Compartment Models	98
4.7.4	Finite Mixture Model Analysis	102
5	Disease Mapping and Cluster Investigations	107
5.1	Introduction	107
5.2	Investigation of General Clustering	109
5.2.1	Traditional Approaches	110
5.2.2	The Empirical Bayes Approach	112

5.3	Computation	117
5.4	A Note on Autocorrelation Versus Heterogeneity	119
5.4.1	Heterogeneity	119
5.4.2	Autocorrelation	121
5.5	Focused Clustering	124
5.5.1	The Score Test for Focused Clustering	124
5.5.2	The Score Test Adjusted for Heterogeneity	128
5.5.3	The Score Test Based on the Negative Binomial Distribution	129
5.5.4	Estimation of α and ν	130
5.6	Case Study: Leukemia in Adults in the Vicinity of Krümmel	132
5.6.1	Background	133
5.6.2	The Retrospective Incidence Study Elbmarsch	134
5.6.3	Focused Analysis	135
5.6.4	Disease Mapping and Model-Based Methods	136
5.7	Mathematical Details of the Score Test	138
5.7.1	Expectation and Variance of the Score	138
5.7.2	The Score Test	139
6	Modeling Heterogeneity in Psychophysiology	143
6.1	The Electroencephalogram	143
6.1.1	Digitization	143
6.2	Modeling Spatial Heterogeneity Using Generalized Linear Mixed Models	144
6.2.1	The Periodogram and its Distributional Properties	144
6.3	Connection to Generalized Linear Models	148
6.3.1	Covariate-Adjusted Finite Mixture Models for the EEG Data	149
7	Investigating and Analyzing Heterogeneity in Meta-analysis	153
7.1	Introduction	153
7.1.1	Different Types of Overviews	154
7.2	Basic Statistical Analysis	155
7.2.1	Single Study Results	155
7.2.2	Publication Bias	157
7.2.3	Estimation of a Summary Effect	160
7.3	Analysis of Heterogeneity	162
7.3.1	The DerSimonian–Laird Approach	164
7.3.2	Maximum Likelihood Estimation of the Heterogeneity Variance τ^2	166
7.3.3	Another Estimator of τ^2 : The Simple Heterogeneity Variance Estimator	168
7.3.4	A Comment on Summary Estimates Under Heterogeneity	169
7.3.5	The Finite Mixture Model Approach	169
7.4	A Simulation Study Comparing Four Estimators of τ^2	171
7.4.1	Design of the Simulation Study	171

7.4.2	Simulation Results	173
7.4.3	Discussion	173
7.5	Metaregression	176
7.5.1	Finite Mixture Models Adjusted for Covariates	177
7.6	Interpretation of the Results of Meta-analysis of Observational Studies	180
7.6.1	Bias	181
7.6.2	Confounding	181
7.6.3	Heterogeneity	182
7.7	Case Study: Aspirin Use and Breast Cancer Risk – A Meta-analysis and Metaregression of Observational Studies from 2001 to 2007	183
7.7.1	Introduction	183
7.7.2	Literature Search and Data Extraction	184
7.7.3	Study Characteristics	184
7.7.4	Publication Bias	185
7.7.5	Results	186
7.7.6	Results of a Metaregression	186
7.7.7	Modeling Dose Response	187
7.7.8	A Metaregression Model for Dose-Response Analysis	190
7.7.9	Discussion	191
7.8	Computation	192
7.8.1	“Standard Meta-analysis”	192
7.8.2	Meta-analysis with SAS	194
7.8.3	Finite Mixture Models	196
7.8.4	Metaregression	197
8	Analysis of Gene Expression Data	201
8.1	DNA Microarrays	201
8.2	The Analysis of Differential Gene Expression	202
8.2.1	Analysis Based on Simultaneous Hypothesis Testing	202
8.2.2	A Mixture Model Approach	206
8.2.3	Computation	209
8.3	A Change of Perspective: Applying Methods from Meta-analysis	210
8.4	Case Study: Identification of a Gene Signature for Breast Cancer Prognosis	213
8.4.1	Introduction	213
8.4.2	Application of the Meta-analytic Mixture Model to the Breast Cancer Data	214
8.4.3	Validation of Results	215
	References	219
	Subject Index	237
	Author Index	243