CONTENTS

1. Introduction 1

1.1 Factors, levels, cells and effects 2
1.2 Balanced and unbalanced data 4
a. Balanced data, 4
b. Special cases of unbalanced data, 4
- i. Planned unbalancedness, 4
-ii. Estimating missing observations, 6
c. Unbalanced data, 6
1.3 Fixed effects and random effects 7
a. Fixed effects models, 7
Example 1 (Tomato varieties), 7
Example 2 (Medications), 8
Example 3 (Soils and fertilizers), 9
b. Random effects models, 9
Example 4 (Clinics), 9
Example 5 (Dairy bulls), 12
Example 6 (Ball bearings and calipers), 13
c. Mixed models, 14
Example 7 (Medications and clinics), 14
Example 8 (Varieties and gardens), 14
1.4 Fixed or random? 15
Example 9 (Mice and technicians), 15
1.5 Finite populations 16
1.6 Summary 17
a. Characteristics of the fixed effects model and the random effects model for the 1-way classification, 17
b. Examples, 17
c. Fixed or random, 18

2. History and Comment 19

2.1 Analysis of variance 19

2.2 Early years: 1861–1949 22
a. Sources, 22
b. Pre-1900, 23
c. 1900–1939, 25
-i. R. A. Fisher, 25
-ii. L. C. Tippett, 27
-iii. The late 1930s, 29
-iv. Unbalanced data, 31
d. The 1940s, 32
2.3 Great strides: 1950–1969 33
a. The Henderson methods, 34
b. ANOVA estimation, in general, 35
-i. Negative estimates, 35
-ii. Unbiasedness, 36
-iii. Best unbiasedness, 37
-iv. Minimal sufficient statistics, 38
-v. Lack of uniqueness, 38
2.4 Into the 1970s and beyond 40
a. Maximum likelihood (ML), 40
b. Restricted maximum likelihood (REML), 41
c. Minimum norm estimation, 41
d. The dispersion-mean model, 42
e. Bayes estimation, 42
f. The recent decade, 43

3. The 1-Way Classification 44

3.1 The model 44
a. The model equation, 44
b. First moments, 45
c. Second moments, 45
3.2 Matrix formulation of the model 47
a. Example 1, 47
b. The general case, 48
c. Dispersion matrices, 49
-i. The traditional random model, 49
-ii. Other alternatives, 50
d. Unbalanced data, 51
-i. Example 2, 51
-ii. The general case, 52
-iii. Dispersion matrix, 52
3.3 Estimating the mean 52
3.4 Predicting random effects 54
3.5 ANOVA estimation—balanced data 57
a. Expected sums of squares, 57
-i. A direct derivation, 58
-ii. Using the matrix formulation, 58
b. ANOVA estimators, 59

c. Negative estimates, 60
d. Normality assumptions, 62
-i. χ^2-distributions of sums of squares, 62
-ii. Independence of sums of squares, 63
-iii. Sampling variances of estimators, 63
-iv. An F*-statistic to test* H: $\sigma_\alpha^2 = 0$, 64
-v. Confidence intervals, 65
-vi. Probability of a negative estimate, 66
-vii. Distribution of estimators, 69
3.6 ANOVA estimation—unbalanced data 69
a. Expected sums of squares, 69
-i. A direct derivation, 69
-ii. Using the matrix formulation, 70
b. ANOVA estimators, 71
c. Negative estimates, 72
d. Normality assumptions, 73
-i. χ^2-distributions of sums of squares, 73
-ii. Independence of sums of squares, 73
-iii. Sampling variances of estimators, 74
-iv. The effect of unbalancedness on sampling variances, 75
-v. F*-statistics,* 76
-vi. Confidence intervals, 76
3.7 Maximum likelihood estimation (MLE) 78
a. Balanced data, 79
-i. Likelihood, 79
-ii. ML equations and their solutions, 80
-iii. ML estimators, 81
-iv. Expected values and bias, 84
-v. Sampling variances, 85
b. Unbalanced data, 86
-i. Likelihood, 86
-ii. ML equations and their solutions, 87
-iii. ML estimators, 88
-iv. Bias, 88
-v. Sampling variances, 88
3.8 Restricted maximum likelihood estimation (REML) 90
a. Balanced data, 91
-i. Likelihood, 91
-ii. REML equations and their solutions, 91
-iii. REML estimators, 92
-iv. Comparison with ANOVA and ML, 92
-v. Bias, 93
-vi. Sampling variances, 93
b. Unbalanced data, 93
3.9 Bayes estimation 94
a. A simple sample, 94
b. The 1-way classification, random model, 97
c. Balanced data, 99

3.10 A summary 103
a. Balanced data, 103
b. Unbalanced data, 106
3.11 Exercises 108

4. Balanced Data 112

4.1 Establishing analysis of variance tables 113
a. Factors and levels, 113
b. Lines in the analysis of variance tables, 113
c. Interactions, 113
d. Degrees of freedom, 114
e. Sums of squares, 114
f. Calculating sums of squares, 115
4.2 Expected mean squares, E(MS) 116
4.3 The 2-way crossed classification 118
a. Introduction, 118
b. Analysis of variance table, 118
c. Expected mean squares, 119
-i. The fixed effects model, 121
-ii. The random effects model, 122
-iii. The mixed model, 122
-iv. A mixed model with Σ-restrictions, 123
d. ANOVA estimators of variance components, 127
4.4 ANOVA estimation 128
4.5 Normality assumptions 131
a. Distribution of mean squares, 131
b. Distribution of estimators, 132
c. Tests of hypotheses, 133
d. Confidence intervals, 135
e. Probability of a negative estimate, 137
f. Sampling variances of estimators, 137
4.6 A matrix formulation of mixed models 138
a. The general mixed model, 138
b. The 2-way crossed classification, 140
-i. Model equation, 140
-ii. Random or mixed?, 142
-iii. Dispersion matrix, 142
c. The 2-way nested classification, 142
d. Interaction or nested factor?, 143
e. The general case, 143
-i. Model equation, 144
-ii. Dispersion matrix, 144
4.7 Maximum likelihood estimation (ML) 146
a. Estimating the mean in random models, 146
b. Four models with closed form estimators, 147
-i. The 1-way random model, 147
-ii. The 2-way nested random model, 148
-iii. The 2-way crossed, with interaction, mixed model, 149

-iv. *The 2-way crossed, no interaction, mixed model*, 150
c. *Unbiasedness*, 151
d. *The 2-way crossed classification, random model*, 151
-i. *With interaction*, 152
-ii. *No interaction*, 153
e. *Existence of explicit solutions*, 153
f. *Asymptotic sampling variances for the 2-way crossed classification*, 154
-i. *The 2-way crossed classification, with interaction, random model*, 155
-ii. *The 2-way crossed classification, no interaction, random model*, 155
-iii. *The 2-way crossed classification, with interaction, mixed model*, 156
-iv. *The 2-way crossed classification, no interaction, mixed model*, 156
g. *Asymptotic sampling variances for two other models*, 157
-i. *The 2-way nested classification, random model*, 157
-ii. *The 1-way classification, random model*, 158
h. *Locating results*, 158
4.8 Restricted maximum likelihood 158
4.9 Estimating fixed effects in mixed models 159
4.10 Summary 161
4.11 Exercises 163

5. Analysis of Variance Estimation for Unbalanced Data 168

5.1 Model formulation 169
a. *Data*, 169
b. *A general model*, 170
-i. *Example 1—the 2-way crossed classification random model*, 170
-ii. *Dispersion matrix*, 171
-iii. *Example 1 (continued)*, 172
5.2 ANOVA estimation 172
a. *Example 2—the 1-way random model, balanced data*, 172
b. *Estimation*, 172
-i. *The general case*, 172
-ii. *Example 2 (continued)*, 173
-iii. *Example 1 (continued)*, 173
c. *Unbiasedness*, 175
d. *Sampling variances*, 176
-i. *A general result*, 176
-ii. *Example 2 (continued)*, 177
-iii. *A direct approach*, 178
e. *Unbiased estimation of sampling variances*, 179
5.3 Henderson's Method I 181
a. *The quadratic forms*, 181
b. *Estimation*, 183

c. *Negative estimates*, 184
d. *Sampling variances*, 184
e. *A general coefficient for Method I*, 186
f. *Synthesis*, 187
g. *Mixed models*, 188
h. *Merits and demerits*, 189
i. *A numerical example*, 189
5.4 Henderson's Method II 190
a. *Estimating the fixed effects*, 191
b. *Calculation*, 192
c. *Verification*, 194
-i. *The matrix* $\mathbf{X}'\mathbf{L}\mathbf{Z}$ *is null*, 194
-ii. *Row sums of* $\mathbf{X}'\mathbf{L}$ *are all the same*, 195
-iii. *All rows of* $\mathbf{X}-\mathbf{X}'\mathbf{L}\mathbf{X}$ *are the same*, 195
d. *Invariance*, 196
-i. *Rank properties*, 196
-ii. *An estimable function*, 197
-iii. *Two solutions*, 198
-iv. *The quadratic forms*, 198
e. *Coefficients of* σ_e^2, 198
f. *No fixed-by-random interactions*, 199
g. *Merits and demerits*, 201
5.5 Henderson's Method III 202
a. *Borrowing from fixed effects models*, 202
-i. *Reductions in sum of squares*, 202
-ii. *Expected sums of squares*, 203
b. *Mixed models*, 204
c. *A general result*, 206
d. *Sampling variances*, 207
e. *Merits and demerits*, 208
5.6 Method III applied to the 2-way crossed classification 208
a. *No interaction, random model*, 209
-i. *One set of sums of squares*, 209
-ii. *Three sets of sums of squares*, 210
-iii. *Calculation*, 213
-iv. *Sampling variances*, 213
b. *No interaction, mixed model*, 213
c. *With interaction, random model*, 213
-i. *One set of sums of squares*, 214
-ii. *Three sets of sums of squares*, 215
-iii. *Calculation*, 217
-iv. *Sampling variances*, 217
d. *With interaction, mixed models*, 218
5.7 Nested models 218
5.8 Other forms of ANOVA estimation 219
a. *Unweighted means method: all cells filled*, 219
b. *Weighted squares of means: all cells filled*, 220
5.9 Comparing different forms of ANOVA estimation 221
5.10 Estimating fixed effects in mixed models 225

5.11 Summary 226
5.12 Exercises 227

6. Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML) 232

6.1 The model and likelihood function 233
6.2 The ML estimation equations 234
a. A direct derivation, 234
b. An alternative form, 236
c. The Hartley–Rao form, 237
6.3 Asymptotic dispersion matrices for ML estimators 238
a. For variance components, 238
b. For ratios of components, 240
c. Maximum?, 241
6.4 Some remarks on computing 242
6.5 ML results for 2-way crossed classification, balanced data 243
a. 2-way crossed, random model, with interaction, 243
-i. Notation, 243
-ii. Inverse of V, 244
-iii. The estimation equations, 245
-iv. Information matrix, 247
b. 2-way crossed, random model, no interaction, 249
6.6 Restricted maximum likelihood (REML) 249
a. Linear combinations of observations, 250
b. The REML equations, 251
c. An alternative form, 251
d. Invariance to choice of error contrasts, 252
e. The information matrix, 252
f. Balanced data, 253
g. Using cell means models for the fixed effects, 253
6.7 Estimating fixed effects in mixed models 254
a. ML, 254
b. REML, 254
6.8 ML or REML? 254
6.9 Summary 255
6.10 Exercises 256

7. Prediction of Random Variables 258

7.1 Introduction 258
7.2 Best prediction (BP) 261
a. The best predictor, 261
b. Mean and variance properties, 262
c. Two properties of the best predictor of a scalar, 263
-i. Maximizing a correlation, 264
-ii. Maximizing the mean of a selected proportion, 264
d. Normality, 265

7.3 Best linear prediction (BLP) 265
a. BLP(**u**), 265
b. Example, 266
c. Derivation, 267
d. Ranking, 268
7.4 Mixed model prediction (BLUP) 269
a. Combining fixed and random effects, 269
b. Example, 270
c. Derivation of BLUP, 271
d. Variances and covariances, 272
e. Normality, 273
7.5 Other derivations of BLUP 273
a. A two-stage derivation, 273
b. A direct derivation assuming linearity, 274
c. Partitioning **y** *into two parts*, 274
d. A Bayes estimator, 275
7.6 Henderson's mixed model equations (MME) 275
a. Derivation, 275
b. Solutions, 276
c. Calculations for ML estimation, 278
-i. The estimation equations, 278
-ii. The information matrix, 280
d. Calculations for REML estimation, 282
-i. The estimation equations, 282
-ii. The information matrix, 284
e. Iterative procedures summarized, 284
-i. Adapting the MME, 284
-ii. Using the MME, 284
-iii. Iterating for ML, 285
-iv. Iterating for REML, 285
f. A summary, 286
7.7 Summary 286
7.8 Exercises 287

8. Computing ML and REML estimates 290

8.1 Introduction 290
8.2 Iterative methods based on derivatives 292
a. The basis of the methods, 292
b. The Newton–Raphson and Marquardt methods, 293
c. Method of scoring, 295
d. Quasi-Newton methods, 295
e. Obtaining starting values, 295
f. Termination rules, 296
g. Incorporation of non-negativity constraints, 296
h. Easing the computational burden, 296
8.3 The EM algorithm 297
a. A general formulation, 297
b. Distributional derivations needed for the EM algorithm, 298

c. EM algorithm for ML estimation (Version 1), 300
d. EM algorithm for ML estimation (Version 2), 301
e. Equivalence of the EM algorithm to the ML equations, 302
f. EM algorithm for REML estimation, 302
g. A Bayesian justification for REML, 303
h. Non-zero correlations among the $\mathbf{u}_i$*s*, 304
8.4 General methods that converge rapidly for balanced data 304
8.5 Pooling estimators from subsets of a large data set 305
8.6 Example: the 1-way random model 307
a. The EM algorithm (Version 1), 308
b. The method of scoring algorithm, 310
8.7 Discussion 311
a. Computing packages, 311
b. Evaluation of algorithms, 312
8.8 Summary 313
8.9 Exercises 314

9. Hierarchical Models and Bayesian Estimation 315

9.1 Basic principles 315
a. Introduction, 315
b. Simple examples, 316
c. The mixed model hierarchy, 317
d. The normal hierarchy, 318
e. Point estimator of variance or variance of point estimator, 319
9.2 Variance estimation in the normal hierarchy 321
a. Formal hierarchical estimation, 321
b. Likelihood methods, 321
c. Empirical Bayes estimation, 325
-i. General strategies, 325
-ii. Estimation, 325
-iii. Connections with likelihood, 326
9.3 Estimation of effects 327
a. Hierarchical estimation, 327
-i. Estimation of $\boldsymbol{\beta}$, 328
-ii. Estimation of $\mathbf{u}$, 330
b. An alternative derivation, 331
-i. Exploiting the multivariate normal structure, 331
-ii. Relationship to BLUP, 333
c. The 1-way classification, random model, 333
-i. Estimation of μ, 334
-ii. Estimation of $\boldsymbol{\alpha}$, 335
d. Empirical Bayes estimation, 337
-i. The 1-way classification, 337
-ii. Cautions, 338
-iii. Variance approximations, 341
9.4 Other types of hierarchies 343
a. A beta–binomial model, 344
b. A generalized linear model, 347

9.5 Practical considerations in hierarchical modeling 349
a. Computational problems, 349
b. Hierarchical EM, 350
9.6 Philosophical considerations in hierarchical modeling 352
9.7 Summary 355
9.8 Exercises 359

10. Binary and Discrete Data 367

10.1 Introduction 367
10.2 ANOVA methods 369
10.3 Beta–binomial models 369
a. Introduction, 369
b. Model specification, 370
c. Likelihood, 371
d. Discussion, 371
10.4 Logit–normal models 372
10.5 Probit–normal models 373
a. Introduction, 373
b. An example, 373
10.6 Discussion 374
10.7 Summary 375
10.8 Exercises 376

11. Other Procedures 378

11.1 Estimating components of covariance 378
a. Easy ANOVA estimation for certain models, 379
b. Examples of covariance components models, 380
-i. Covariances between effects of the same random factor, 381
-ii. Covariances between effects of different random factors, 381
-iii. Covariances between error terms, 382
c. Combining variables into a single vector, 382
d. Genetic covariances, 383
e. Maximum likelihood (ML) estimation, 383
-i. Estimation equations, 383
-ii. Large-sample dispersion matrix, 385
f. Restricted maximum likelihood (REML) estimation, 386
-i. Estimation equations, 386
-ii. Large-sample dispersion matrix, 387
11.2 Modeling variance components as covariances 387
a. All-cells-filled data, 388
b. Balanced data, 389
c. Diagnostic opportunities, 389
d. Some-cells-empty data, 390
11.3 Criteria-based procedures 391
a. Three criteria, 392
-i. Unbiasedness, 392
-ii. Translation invariance, 392

-iii. *Minimum variance*, 393
b. *LaMotte's minimum mean square procedures*, 393
-i. *Class C_0: unrestricted*, 393
-ii. *Class C_1: expectation of* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *containing no* $\boldsymbol{\beta}$, 394
-iii. *Class C_2: translation-invariant*, 394
-iv. *Class C_3: unbiased*, 394
-v. *Class C_4: translation-invariant and unbiased*, 394
c. *Minimum variance estimation (MINVAR)*, 394
d. *Minimum norm estimation (MINQUE)*, 397
e. *REML, MINQUE and I-MINQUE*, 398
f. *REML for balanced data*, 399
g. *MINQUEO*, 399
h. *MINQUE for the 1-way classification*, 399
11.4 Summary 400
11.5 Exercises 402

12. The Dispersion-Mean Model 405

12.1 The model 405
12.2 Ordinary least squares (OLS) yields MINQUEO 406
12.3 Fourth moments in the mixed model 407
a. *Dispersion matrix of* $\mathbf{u} \otimes \mathbf{u}$, 407
-i. *A normalizing transformation*, 407
-ii. *Example*, 408
-iii. *The general form of* $\mathrm{E}(\mathbf{w}\mathbf{w}' \otimes \mathbf{w}\mathbf{w}')$, 410
b. *Fourth central moments of* $\mathbf{y}$, 411
-i. *General case*, 411
-ii. *Under normality*, 411
c. *Dispersion matrix of* $\mathscr{Y}$, 411
-i. *General case*, 411
-ii. *Under normality*, 412
d. *Variance of a translation-invariant quadratic form*, 412
12.4 Generalized least squares (GLS) 413
a. *GLS yields REML equations under normality*, 413
b. *Excursus on estimating fixed effects*, 415
c. *REML is BLUE*, 415
12.5 Modified GLS yields ML 416
12.6 Balanced data 417
a. *Estimation under zero kurtosis*, 417
-i. *History*, 417
-ii. *The model*, 419
-iii. *Conclusion*, 421
b. *Estimation under non-zero kurtosis*, 421
-i. *The model*, 421
-ii. *ANOVA estimation*, 422
-iii. *Conclusion*, 424
12.7 Non-negative estimation 424
12.8 Summary 425
12.9 Exercises 426

Appendix F. Estimation Formulae for Unbalanced Data 427

PART I. THREE NESTED MODELS 427

F.1 The 1-way classification 427
a. Model, 427
b. Analysis of variance estimators, 428
c. Variances of analysis of variance estimators (under normality), 428
d. Maximum likelihood estimation (under normality), 428
e. Large-sample variances of maximum likelihood estimators (under normality), 428

F.2 The 2-way nested classification 429
a. Model, 429
b. Analysis of variance estimators, 429
c. Variances of analysis of variance estimators (under normality), 429
d. Large-sample variances of maximum likelihood estimators (under normality), 430

F.3 The 3-way nested classification 431
a. Model, 431
b. Analysis of variance estimators, 431
c. Variances of analysis of variance estimators (under normality), 432

PART II. THE 2-WAY CROSSED CLASSIFICATION 434

F.4 With interaction, random model 434
a. Model, 434
b. Henderson Method I estimators, 434
c. Variances of Henderson Method I estimators (under normality), 435
d. Henderson Method III estimators, 437

F.5 With interaction, mixed model 438
a. Model, 438
b. Henderson Method III, 438

F.6 No interaction, random model 439
a. Model, 439
b. Henderson Method I, 439
c. Variances of Henderson Method I estimators (under normality), 439
d. Henderson Method III, 440
e. Variances of Henderson Method III estimators (under normality), 440

F.7 No interaction, mixed model 441
a. Model, 441
b. Henderson Method III, 441

Appendix M. Some Results in Matrix Algebra 442

M.1 Summing vectors, and J-matrices 442
M.2 Direct sums and products 443
M.3 A matrix notation in terms of elements 445
M.4 Generalized inverses 447
a. Definitions, 447

b. Generalized inverses of $\mathbf{X}'\mathbf{X}$, 448
c. Partitioning $\mathbf{X}'\mathbf{X}$, 450
d. Rank results, 451
e. Vectors orthogonal to columns of $\mathbf{X}$, 451
f. A theorem involving $\mathbf{K}'$ *of maximum row rank for* $\mathbf{K}'\mathbf{X}$ *being null*, 451
M.5 The Schur complement 453
M.6 The trace of a matrix 454
M.7 Differentiation of matrix expressions 454
a. Scalars, 454
b. Vectors, 455
c. Inner products, 455
d. Quadratic forms, 455
e. Inverses, 456
f. Determinants, 456
g. Traces, 457
M.8 The operators vec and vech 458
M.9 Vec permutation matrices 459
M.10 The equality $\mathbf{V}\mathbf{V}^{-}\mathbf{X} = \mathbf{X}$ 460

Appendix S. Some Results in Statistics **461**
S.1 Conditional first and second moments 461
S.2 Least squares estimation 462
S.3 Normal and χ^2-distributions 463
a. Central χ^2, 464
b. Mean squares, 464
c. Non-central χ^2, 465
S.4 F-distributions 465
S.5 Quadratic forms 466
S.6 Bayes estimation 467
a. Density functions, 467
b. Bayes Theorem, 468
c. Bayes estimation, 469
d. Example, 469
e. Empirical Bayes estimation, 471
S.7 Maximum likelihood 472
a. The likelihood function, 472
b. Maximum likelihood estimation, 472
c. Asymptotic dispersion matrix, 473
d. Transforming parameters, 474

References **475**

List of Tables and Figures **490**

Author Index **493**

Subject Index **497**