Experimental Design and Data Analysis for Biologists

Gerry P. Quinn
Michael J. Keough

CAMBRIDGE

Contents

Preface *page* xv

1 Introduction 1

1.1 Scientific method 1
1.1.1 Pattern description 2
1.1.2 Models 2
1.1.3 Hypotheses and tests 3
1.1.4 Alternatives to falsification 4
1.1.5 Role of statistical analysis 5
1.2 Experiments and other tests 5
1.3 Data, observations and variables 7
1.4 Probability 7
1.5 Probability distributions 9
1.5.1 Distributions for variables 10
1.5.2 Distributions for statistics 12

2 Estimation 14

2.1 Samples and populations 14
2.2 Common parameters and statistics 15
2.2.1 Center (location) of distribution 15
2.2.2 Spread or variability 16
2.3 Standard errors and confidence intervals for the mean 17
2.3.1 Normal distributions and the Central Limit Theorem 17
2.3.2 Standard error of the sample mean 18
2.3.3 Confidence intervals for population mean 19
2.3.4 Interpretation of confidence intervals for population mean 20
2.3.5 Standard errors for other statistics 20
2.4 Methods for estimating parameters 23
2.4.1 Maximum likelihood (ML) 23
2.4.2 Ordinary least squares (OLS) 24
2.4.3 ML vs OLS estimation 25
2.5 Resampling methods for estimation 25
2.5.1 Bootstrap 25
2.5.2 Jackknife 26
2.6 Bayesian inference – estimation 27
2.6.1 Bayesian estimation 27
2.6.2 Prior knowledge and probability 28
2.6.3 Likelihood function 28
2.6.4 Posterior probability 28
2.6.5 Examples 29
2.6.6 Other comments 29

3 Hypothesis testing 32

3.1 Statistical hypothesis testing 32
3.1.1 Classical statistical hypothesis testing 32
3.1.2 Associated probability and Type I error 34
3.1.3 Hypothesis tests for a single population 35
3.1.4 One- and two-tailed tests 37
3.1.5 Hypotheses for two populations 37
3.1.6 Parametric tests and their assumptions 39
3.2 Decision errors 42
3.2.1 Type I and II errors 42
3.2.2 Asymmetry and scalable decision criteria 44
3.3 Other testing methods 45
3.3.1 Robust parametric tests 45
3.3.2 Randomization (permutation) tests 45
3.3.3 Rank-based non-parametric tests 46
3.4 Multiple testing 48
3.4.1 The problem 48
3.4.2 Adjusting significance levels and/or *P* values 49
3.5 Combining results from statistical tests 50
3.5.1 Combining *P* values 50
3.5.2 Meta-analysis 50
3.6 Critique of statistical hypothesis testing 51
3.6.1 Dependence on sample size and stopping rules 51
3.6.2 Sample space – relevance of data not observed 52
3.6.3 *P* values as measure of evidence 53
3.6.4 Null hypothesis always false 53
3.6.5 Arbitrary significance levels 53
3.6.6 Alternatives to statistical hypothesis testing 53
3.7 Bayesian hypothesis testing 54

4 Graphical exploration of data 58

4.1 Exploratory data analysis 58
4.1.1 Exploring samples 58
4.2 Analysis with graphs 62
4.2.1 Assumptions of parametric linear models 62
4.3 Transforming data 64
4.3.1 Transformations and distributional assumptions 65
4.3.2 Transformations and linearity 67
4.3.3 Transformations and additivity 67
4.4 Standardizations 67
4.5 Outliers 68
4.6 Censored and missing data 68
4.6.1 Missing data 68
4.6.2 Censored (truncated) data 69
4.7 General issues and hints for analysis 71
4.7.1 General issues 71

5 | Correlation and regression 72

5.1 Correlation analysis 72
- 5.1.1 Parametric correlation model 72
- 5.1.2 Robust correlation 76
- 5.1.3 Parametric and non-parametric confidence regions 76

5.2 Linear models 77

5.3 Linear regression analysis 78
- 5.3.1 Simple (bivariate) linear regression 78
- 5.3.2 Linear model for regression 80
- 5.3.3 Estimating model parameters 85
- 5.3.4 Analysis of variance 88
- 5.3.5 Null hypotheses in regression 89
- 5.3.6 Comparing regression models 90
- 5.3.7 Variance explained 91
- 5.3.8 Assumptions of regression analysis 92
- 5.3.9 Regression diagnostics 94
- 5.3.10 Diagnostic graphics 96
- 5.3.11 Transformations 98
- 5.3.12 Regression through the origin 98
- 5.3.13 Weighted least squares 99
- 5.3.14 X random (Model II regression) 100
- 5.3.15 Robust regression 104

5.4 Relationship between regression and correlation 106

5.5 Smoothing 107
- 5.5.1 Running means 107
- 5.5.2 LO(W)ESS 107
- 5.5.3 Splines 108
- 5.5.4 Kernels 108
- 5.5.5 Other issues 109

5.6 Power of tests in correlation and regression 109

5.7 General issues and hints for analysis 110
- 5.7.1 General issues 110
- 5.7.2 Hints for analysis 110

6 | Multiple and complex regression 111

6.1 Multiple linear regression analysis 111
- 6.1.1 Multiple linear regression model 114
- 6.1.2 Estimating model parameters 119
- 6.1.3 Analysis of variance 119
- 6.1.4 Null hypotheses and model comparisons 121
- 6.1.5 Variance explained 122
- 6.1.6 Which predictors are important? 122
- 6.1.7 Assumptions of multiple regression 124
- 6.1.8 Regression diagnostics 125
- 6.1.9 Diagnostic graphics 125
- 6.1.10 Transformations 127
- 6.1.11 Collinearity 127

6.1.12 Interactions in multiple regression 130
6.1.13 Polynomial regression 133
6.1.14 Indicator (dummy) variables 135
6.1.15 Finding the "best" regression model 137
6.1.16 Hierarchical partitioning 141
6.1.17 Other issues in multiple linear regression 142
6.2 Regression trees 143
6.3 Path analysis and structural equation modeling 145
6.4 Nonlinear models 150
6.5 Smoothing and response surfaces 152
6.6 General issues and hints for analysis 153
6.6.1 General issues 153
6.6.2 Hints for analysis 154

7 Design and power analysis 155

7.1 Sampling 155
7.1.1 Sampling designs 155
7.1.2 Size of sample 157
7.2 Experimental design 157
7.2.1 Replication 158
7.2.2 Controls 160
7.2.3 Randomization 161
7.2.4 Independence 163
7.2.5 Reducing unexplained variance 164
7.3 Power analysis 164
7.3.1 Using power to plan experiments (*a priori* power analysis) 166
7.3.2 *Post hoc* power calculation 168
7.3.3 The effect size 168
7.3.4 Using power analyses 170
7.4 General issues and hints for analysis 171
7.4.1 General issues 171
7.4.2 Hints for analysis 172

8 Comparing groups or treatments – analysis of variance 173

8.1 Single factor (one way) designs 173
8.1.1 Types of predictor variables (factors) 176
8.1.2 Linear model for single factor analyses 178
8.1.3 Analysis of variance 184
8.1.4 Null hypotheses 186
8.1.5 Comparing ANOVA models 187
8.1.6 Unequal sample sizes (unbalanced designs) 187
8.2 Factor effects 188
8.2.1 Random effects: variance components 188
8.2.2 Fixed effects 190
8.3 Assumptions 191
8.3.1 Normality 192
8.3.2 Variance homogeneity 193
8.3.3 Independence 193

8.4 ANOVA diagnostics 194
8.5 Robust ANOVA 195
8.5.1 Tests with heterogeneous variances 195
8.5.2 Rank-based ("non-parametric") tests 195
8.5.3 Randomization tests 196
8.6 Specific comparisons of means 196
8.6.1 Planned comparisons or contrasts 197
8.6.2 Unplanned pairwise comparisons 199
8.6.3 Specific contrasts versus unplanned pairwise comparisons 201
8.7 Tests for trends 202
8.8 Testing equality of group variances 203
8.9 Power of single factor ANOVA 204
8.10 General issues and hints for analysis 206
8.10.1 General issues 206
8.10.2 Hints for analysis 206

9 | Multifactor analysis of variance 208

9.1 Nested (hierarchical) designs 208
9.1.1 Linear models for nested analyses 210
9.1.2 Analysis of variance 214
9.1.3 Null hypotheses 215
9.1.4 Unequal sample sizes (unbalanced designs) 216
9.1.5 Comparing ANOVA models 216
9.1.6 Factor effects in nested models 216
9.1.7 Assumptions for nested models 218
9.1.8 Specific comparisons for nested designs 219
9.1.9 More complex designs 219
9.1.10 Design and power 219
9.2 Factorial designs 221
9.2.1 Linear models for factorial designs 225
9.2.2 Analysis of variance 230
9.2.3 Null hypotheses 232
9.2.4 What are main effects and interactions really measuring? 237
9.2.5 Comparing ANOVA models 241
9.2.6 Unbalanced designs 241
9.2.7 Factor effects 247
9.2.8 Assumptions 249
9.2.9 Robust factorial ANOVAs 250
9.2.10 Specific comparisons on main effects 250
9.2.11 Interpreting interactions 251
9.2.12 More complex designs 255
9.2.13 Power and design in factorial ANOVA 259
9.3 Pooling in multifactor designs 260
9.4 Relationship between factorial and nested designs 261
9.5 General issues and hints for analysis 261
9.5.1 General issues 261
9.5.2 Hints for analysis 261

10 Randomized blocks and simple repeated measures: unreplicated two factor designs 262

10.1 Unreplicated two factor experimental designs 262
10.1.1 Randomized complete block (RCB) designs 262
10.1.2 Repeated measures (RM) designs 265
10.2 Analyzing RCB and RM designs 268
10.2.1 Linear models for RCB and RM analyses 268
10.2.2 Analysis of variance 272
10.2.3 Null hypotheses 273
10.2.4 Comparing ANOVA models 274
10.3 Interactions in RCB and RM models 274
10.3.1 Importance of treatment by block interactions 274
10.3.2 Checks for interaction in unreplicated designs 277
10.4 Assumptions 280
10.4.1 Normality, independence of errors 280
10.4.2 Variances and covariances – sphericity 280
10.4.3 Recommended strategy 284
10.5 Robust RCB and RM analyses 284
10.6 Specific comparisons 285
10.7 Efficiency of blocking (to block or not to block?) 285
10.8 Time as a blocking factor 287
10.9 Analysis of unbalanced RCB designs 287
10.10 Power of RCB or simple RM designs 289
10.11 More complex block designs 290
10.11.1 Factorial randomized block designs 290
10.11.2 Incomplete block designs 292
10.11.3 Latin square designs 292
10.11.4 Crossover designs 296
10.12 Generalized randomized block designs 298
10.13 RCB and RM designs and statistical software 298
10.14 General issues and hints for analysis 299
10.14.1 General issues 299
10.14.2 Hints for analysis 300

11 Split-plot and repeated measures designs: partly nested analyses of variance 301

11.1 Partly nested designs 301
11.1.1 Split-plot designs 301
11.1.2 Repeated measures designs 305
11.1.3 Reasons for using these designs 309
11.2 Analyzing partly nested designs 309
11.2.1 Linear models for partly nested analyses 310
11.2.2 Analysis of variance 313
11.2.3 Null hypotheses 315
11.2.4 Comparing ANOVA models 318
11.3 Assumptions 318
11.3.1 Between plots/subjects 318
11.3.2 Within plots/subjects and multisample sphericity 318

11.4 Robust partly nested analyses 320
11.5 Specific comparisons 320
11.5.1 Main effects 320
11.5.2 Interactions 321
11.5.3 Profile (i.e. trend) analysis 321
11.6 Analysis of unbalanced partly nested designs 322
11.7 Power for partly nested designs 323
11.8 More complex designs 323
11.8.1 Additional between-plots/subjects factors 324
11.8.2 Additional within-plots/subjects factors 329
11.8.3 Additional between-plots/subjects and within-plots/subjects factors 332
11.8.4 General comments about complex designs 335
11.9 Partly nested designs and statistical software 335
11.10 General issues and hints for analysis 337
11.10.1 General issues 337
11.10.2 Hints for individual analyses 337

12 Analyses of covariance 339

12.1 Single factor analysis of covariance (ANCOVA) 339
12.1.1 Linear models for analysis of covariance 342
12.1.2 Analysis of (co)variance 347
12.1.3 Null hypotheses 347
12.1.4 Comparing ANCOVA models 348
12.2 Assumptions of ANCOVA 348
12.2.1 Linearity 348
12.2.2 Covariate values similar across groups 349
12.2.3 Fixed covariate (X) 349
12.3 Homogeneous slopes 349
12.3.1 Testing for homogeneous within-group regression slopes 349
12.3.2 Dealing with heterogeneous within-group regression slopes 350
12.3.3 Comparing regression lines 352
12.4 Robust ANCOVA 352
12.5 Unequal sample sizes (unbalanced designs) 353
12.6 Specific comparisons of adjusted means 353
12.6.1 Planned contrasts 353
12.6.2 Unplanned comparisons 353
12.7 More complex designs 353
12.7.1 Designs with two or more covariates 353
12.7.2 Factorial designs 354
12.7.3 Nested designs with one covariate 355
12.7.4 Partly nested models with one covariate 356
12.8 General issues and hints for analysis 357
12.8.1 General issues 357
12.8.2 Hints for analysis 358

13 Generalized linear models and logistic regression 359

13.1 Generalized linear models 359
13.2 Logistic regression 360
13.2.1 Simple logistic regression 360
13.2.2 Multiple logistic regression 365
13.2.3 Categorical predictors 368
13.2.4 Assumptions of logistic regression 368
13.2.5 Goodness-of-fit and residuals 368
13.2.6 Model diagnostics 370
13.2.7 Model selection 370
13.2.8 Software for logistic regression 371
13.3 Poisson regression 371
13.4 Generalized additive models 372
13.5 Models for correlated data 375
13.5.1 Multi-level (random effects) models 376
13.5.2 Generalized estimating equations 377
13.6 General issues and hints for analysis 378
13.6.1 General issues 378
13.6.2 Hints for analysis 379

14 Analyzing frequencies 380

14.1 Single variable goodness-of-fit tests 381
14.2 Contingency tables 381
14.2.1 Two way tables 381
14.2.2 Three way tables 388
14.3 Log-linear models 393
14.3.1 Two way tables 394
14.3.2 Log-linear models for three way tables 395
14.3.3 More complex tables 400
14.4 General issues and hints for analysis 400
14.4.1 General issues 400
14.4.2 Hints for analysis 400

15 Introduction to multivariate analyses 401

15.1 Multivariate data 401
15.2 Distributions and associations 402
15.3 Linear combinations, eigenvectors and eigenvalues 405
15.3.1 Linear combinations of variables 405
15.3.2 Eigenvalues 405
15.3.3 Eigenvectors 406
15.3.4 Derivation of components 409
15.4 Multivariate distance and dissimilarity measures 409
15.4.1 Dissimilarity measures for continuous variables 412
15.4.2 Dissimilarity measures for dichotomous (binary) variables 413
15.4.3 General dissimilarity measures for mixed variables 413
15.4.4 Comparison of dissimilarity measures 414
15.5 Comparing distance and/or dissimilarity matrices 414

15.6 Data standardization 415
15.7 Standardization, association and dissimilarity 417
15.8 Multivariate graphics 417
15.9 Screening multivariate data sets 418
15.9.1 Multivariate outliers 419
15.9.2 Missing observations 419
15.10 General issues and hints for analysis 423
15.10.1 General issues 423
15.10.2 Hints for analysis 424

16 | Multivariate analysis of variance and discriminant analysis 425

16.1 Multivariate analysis of variance (MANOVA) 425
16.1.1 Single factor MANOVA 426
16.1.2 Specific comparisons 432
16.1.3 Relative importance of each response variable 432
16.1.4 Assumptions of MANOVA 433
16.1.5 Robust MANOVA 434
16.1.6 More complex designs 434
16.2 Discriminant function analysis 435
16.2.1 Description and hypothesis testing 437
16.2.2 Classification and prediction 439
16.2.3 Assumptions of discriminant function analysis 441
16.2.4 More complex designs 441
16.3 MANOVA vs discriminant function analysis 441
16.4 General issues and hints for analysis 441
16.4.1 General issues 441
16.4.2 Hints for analysis 441

17 | Principal components and correspondence analysis 443

17.1 Principal components analysis 443
17.1.1 Deriving components 447
17.1.2 Which association matrix to use? 450
17.1.3 Interpreting the components 451
17.1.4 Rotation of components 451
17.1.5 How many components to retain? 452
17.1.6 Assumptions 453
17.1.7 Robust PCA 454
17.1.8 Graphical representations 454
17.1.9 Other uses of components 456
17.2 Factor analysis 458
17.3 Correspondence analysis 459
17.3.1 Mechanics 459
17.3.2 Scaling and joint plots 461
17.3.3 Reciprocal averaging 462
17.3.4 Use of CA with ecological data 462
17.3.5 Detrending 463
17.4 Canonical correlation analysis 463

17.5 Redundancy analysis 466
17.6 *Canonical correspondence analysis* 467
17.7 Constrained and partial "ordination" 468
17.8 General issues and hints for analysis 471
17.8.1 General issues 471
17.8.2 Hints for analysis 471

18 | Multidimensional scaling and cluster analysis 473

18.1 Multidimensional scaling 473
18.1.1 Classical scaling – principal coordinates analysis (PCoA) 474
18.1.2 Enhanced multidimensional scaling 476
18.1.3 Dissimilarities and testing hypotheses about groups of objects 482
18.1.4 Relating MDS to original variables 487
18.1.5 Relating MDS to covariates 487
18.2 Classification 488
18.2.1 Cluster analysis 488
18.3 Scaling (ordination) and clustering for biological data 491
18.4 General issues and hints for analysis 493
18.4.1 General issues 493
18.4.2 Hints for analysis 493

19 | Presentation of results 494

19.1 Presentation of analyses 494
19.1.1 Linear models 494
19.1.2 Other analyses 497
19.2 Layout of tables 497
19.3 Displaying summaries of the data 498
19.3.1 Bar graph 500
19.3.2 Line graph (category plot) 502
19.3.3 Scatterplots 502
19.3.4 Pie charts 503
19.4 Error bars 504
19.4.1 Alternative approaches 506
19.5 Oral presentations 507
19.5.1 Slides, computers, or overheads? 507
19.5.2 Graphics packages 508
19.5.3 Working with color 508
19.5.4 Scanned images 509
19.5.5 Information content 509
19.6 General issues and hints 510

References 511

Index 527