
Chapter 1 Introduction

- 1.1 Almost every biological theory or hypothesis is a model of how nature works
- 1.2 We use data to separate wrong models from (possibly) correct ones (or bad models from less bad ones)
 - 1.2.1 What kind of data do we need?
 - 1.2.2 The Signal and the Noise
 - 1.2.3 Random & representative
 - 1.2.4 How do we decide if a model fits well? What's the "best" model?
 - 1.2.5 What do I do next?
- 1.3 There are many ways to reach wrong conclusions!
- 1.4 There are also many ways to be right
- 1.5 Our philosophy in this book
 - 1.5.1 Think clearly
 - 1.5.2 Think in advance
 - 1.5.3 Think before you analyze
- 1.6 How this book is structured
- 1.7 A bit of housekeeping

Chapter 2 Things to know before proceeding

- 2.1 Samples, populations and statistical inference
- 2.2 Probability
- 2.3 Probability distributions
 - 2.3.1 Distributions for variables
 - 2.3.2 Distributions for statistics
- 2.4 Frequentist ("classical") estimation
 - 2.4.1 Methods for estimation
 - 2.4.2 Simple parameters and statistics
 - 2.4.2.1 *Center (location) of distribution*
 - 2.4.2.2 *Spread or variability*
 - 2.4.3 Sampling distribution of the mean
 - 2.4.4 Standard error of the sample mean
 - 2.4.5 Confidence intervals for population mean
 - 2.4.5.1 *Interpretation of confidence intervals for population mean*
 - 2.4.6 Standard errors and confidence intervals for other statistics
 - 2.4.7 Resampling methods for frequentist estimation
 - 2.4.7.1 *Bootstrap*
 - 2.4.7.2 *Jackknife*
- 2.5 Hypothesis testing
 - 2.5.1 Frequentist statistical hypothesis testing
 - 2.5.2 Decision errors
 - 2.5.3 One- and two-tailed tests
 - 2.5.4 Multiple hypothesis testing
 - 2.5.4.1 *Adjusting threshold levels or P-values*
 - 2.5.4.2 *False discovery rates*
 - 2.5.5 Testing hypotheses about means and variances for two populations
 - 2.5.6 Parametric tests and their assumptions
 - 2.5.6.1 *Robust parametric tests*
 - 2.5.6.2 *Randomization (permutation) tests*
 - 2.5.6.3 *Rank-based non-parametric tests*
- 2.6 Comments on frequentist inference
- 2.7 Bayesian statistical inference
 - 2.7.1 Prior knowledge and probability
 - 2.7.2 Likelihood function
 - 2.7.3 Posterior probability
 - 2.7.4 Model comparison and Bayes factors
 - 2.7.5 Final comments
- 2.8 Further Reading

Chapter 3 Sampling and Experimental Design

- 3.1 Sampling Design
 - 3.1.1 Probability Sampling
 - 3.1.1.1 *Simple Random Sampling*
 - 3.1.1.2 *Stratified Sampling*
 - 3.1.1.3 *Cluster Sampling*
 - 3.1.1.4 *Systematic Sampling*
 - 3.1.1.5 *Unequal Probability Sampling*
 - 3.1.1.6 *Adaptive Sampling*
 - 3.1.2 Sample Size for Random Sampling
 - 3.1.3 Non-probability Sampling
 - 3.1.3.1 *Convenience Sampling*
 - 3.1.3.2 *Haphazard Sampling*
 - 3.1.3.3 *Purposive Sampling*
- 3.2 Experimental Design
 - 3.2.1 Replication of Experimental Units
 - 3.2.2 Controls
 - 3.2.3 Randomization
 - 3.2.4 Independence
 - 3.2.5 Reducing Unexplained Variance
 - 3.2.6 Limitations of Manipulative Experiments
- 3.3 Sample Size for Detecting Differences - Power Analysis
 - 3.3.1 Using Power to Plan Experiments (*a priori* Power Analysis)
 - 3.3.1.1 *Sample Size Calculation (Power, σ , α , ES Known)*
 - 3.3.1.2 *Effect Size Calculation (Power, n , σ , Known)*
 - 3.3.1.3 *Sequence for Using Power Analysis to Design Experiments*
 - 3.3.2 *Post hoc* Power Calculation
 - 3.3.3 Effect Size
 - 3.3.3.1 *What if we can't Identify an Effect Size about which we feel Confident?*
 - 3.3.4 Using Power Analyses
- 3.4 Key Points
- 3.5 Further Reading

Chapter 4 Introduction to Linear Models

- 4.1 What is a linear model?
- 4.2 Components of Linear Models
 - 4.2.1 Types of Response Variables
 - 4.2.1.1 *Continuous Response Variables*
 - 4.2.1.2 *Discrete Response Variables*
 - 4.2.2 Types of Predictor Variables
 - 4.2.2.1 *Categorical (Discrete) vs. Continuous*
 - 4.2.2.2 *Fixed vs. Random*
- 4.3 Assembling our Linear Model
- 4.4 Estimation for Linear Models
 - 4.4.1 Ordinary Least Squares
 - 4.4.2 Maximum Likelihood
 - 4.4.3 Robust Estimation Methods for Linear Models
- 4.5 How Well Does a Model Fit?
 - 4.5.1 OLS Measures of Fit: Analysis of Variance
 - 4.5.2 ML Measures of Fit: Log-likelihood and Deviance
 - 4.5.3 Information Criteria
- 4.6 Assumptions for Linear Model Inference
 - 4.6.1 Assumptions for OLS
 - 4.6.1.1 *Zero Conditional Mean of Errors*
 - 4.6.1.2 *Independence of Errors*
 - 4.6.1.3 *Homogeneity of Error Variances*
 - 4.6.1.4 *Normality of Errors*
 - 4.6.1.5 *Solutions*

- 4.6.2 Assumptions for ML
 - 4.6.3 Model Diagnostics
 - 4.7 Types of Linear Models
 - 4.7.1 General Linear Models
 - 4.7.1.1 *Continuous Response and Predictor(s) – “Regression” Models*
 - 4.7.1.2 *Continuous (Normal) Response and Categorical Predictors – “ANOVA” Models*
 - 4.7.2 Generalized Linear Models
 - 4.7.2.1 *Binary Response, Continuous Predictor(s): Logistic Models*
 - 4.7.2.2 *Poisson Response, Continuous Predictor(s)*
 - 4.7.2.3 *Contingency Tables: Loglinear Model*
 - 4.7.3 Linear Mixed Models (General and Generalized)
 - 4.8 Key Points
 - 4.9 Further Reading
-

Chapter 5 Exploratory Data Analysis

- 5.1 Basic Graphical Tools
 - 5.1.1 Some Common Basic Graphs
 - 5.1.1.1 *Histogram*
 - 5.1.1.2 *Dotplot*
 - 5.1.1.3 *Boxplot*
 - 5.1.1.4 *Probability Plot*
 - 5.1.1.5 *Scatterplot*
 - 5.1.1.6 *Scatterplot Matrix (SPLOM)*
 - 5.1.2 Smoothing
 - 5.1.3 Residual Plots
 - 5.2 Outliers
 - 5.3 Am I Fitting the Right Model?
 - 5.3.1 The Underlying Probability Distribution
 - 5.3.2 Homogeneity of Variances
 - 5.3.3 Is my Linear Model “Linear”?
 - 5.4 Is it Normal to Transform Data?
 - 5.4.1 Transformations and Distributional Assumptions
 - 5.4.2 Transformations and Linearity
 - 5.4.3 Transformations and Additivity
 - 5.4.4 Do we Really Need a Data Transformation?
 - 5.5 Standardizations
 - 5.6 Missing Data
 - 5.6.1 Missing Data Mechanisms
 - 5.6.2 Detecting Missing Data
 - 5.6.3 Methods for Missing Data
 - 5.6.3.1 *Deletions*
 - 5.6.3.2 *Single Imputation*
 - 5.6.3.3 *Multiple Imputation*
 - 5.7 Key Points
 - 5.8 Further Reading
-

Chapter 6 Simple Linear Models with One Predictor

- 6.1 Linear Model for a Single Continuous Predictor - Linear Regression
 - 6.1.1 Linear Model for a Continuous Predictor (Linear Regression Model)
 - 6.1.2 Model Parameters
 - 6.1.2.1 *Regression Slope*
 - 6.1.2.2 *Intercept*
 - 6.1.2.3 *Predicted Values*
 - 6.1.2.4 *Error Terms and Their Variance*
 - 6.1.2.5 *Standardized Coefficients*
 - 6.1.3 Inference for Parameters
 - 6.1.3.1 *Standard Errors and Confidence Intervals*
 - 6.1.3.2 *Statistical Hypotheses*

- 6.1.4 Inference for Predicted Values
- 6.1.5 Model Comparison and the Analysis of Variance
- 6.1.6 Regression Through the Origin
- 6.1.7 Regression with X Random
- 6.2 Linear model for a Single Categorical Predictor (Factor)
 - 6.2.1.1 *Completely Randomized (Experimental) Designs*
 - 6.2.1.2 *Observational (Non-experimental) Designs*
- 6.2.2 Linear Model for a Categorical Predictor
 - 6.2.2.1 *Linear Effects Model*
 - 6.2.2.2 *Means Model*
 - 6.2.2.3 *Regression (Dummy Variable) Model*
- 6.2.3 Model Parameters
 - 6.2.3.1 *Predicted Values*
 - 6.2.3.2 *Error Terms and Their Variance*
- 6.2.4 Inference for Parameters
 - 6.2.4.1 *Standard Errors and Confidence Intervals*
 - 6.2.4.2 *Hypothesis Tests*
- 6.2.5 Model Comparison and the Analysis of Variance
- 6.2.6 Unequal Sample Sizes (Unbalanced Designs)
- 6.2.7 Specific Comparisons of Group Means
 - 6.2.7.1 *Planned Comparisons or Contrasts*
 - 6.2.7.2 *Unplanned Pairwise Comparisons*
- 6.3 Predictor effects
 - 6.3.1 Continuous Predictor (Regression) Models
 - 6.3.2 Categorical Predictor Models
- 6.4 Assumptions
 - 6.4.1 Zero Conditional Mean
 - 6.4.2 Independence
 - 6.4.3 Variance Homogeneity
 - 6.4.4 Normality
- 6.5 Model Diagnostics
 - 6.5.1 Residuals
 - 6.5.2 Leverage
 - 6.5.3 Influence Measures
 - 6.5.4 Diagnostic Plots
 - 6.5.4.1 *Scatterplots*
 - 6.5.4.2 *Boxplots*
 - 6.5.4.3 *Residual Plots*
 - 6.5.5 Transformations
- 6.6 Robust Linear Models
 - 6.6.1 Rank-based (“Nonparametric”) Methods
 - 6.6.1.1 *Continuous Predictor (Regression)*
 - 6.6.1.2 *Categorical Predictor*
 - 6.6.2 Generalized (Weighted) Least Squares
 - 6.6.3 Other Robust Methods
 - 6.6.3.1 *Categorical Predictors - Handling Heterogeneous Variances*
 - 6.6.4 Resampling and Permutation Methods
- 6.7 Power of Single Predictor Linear Models
 - 6.7.1 Regression Models
 - 6.7.2 Categorical Predictor Models
- 6.8 Key points
- 6.9 Further Reading

Chapter 7 Linear Models for Factorial (Crossed) Designs

- 7.1 Two Factor Fully Factorial (Crossed) Designs
 - 7.1.1 Completely Randomized (Experimental) Designs
 - 7.1.2 Observational (Non-experimental) Designs
 - 7.1.3 Designs that Combine Completely Randomized Factors with Non-randomized (Observational) Factors

- 7.1.4 The Factorial Linear Effects Model
- 7.1.5 Model Parameters
 - 7.1.5.1 *Predicted values*
 - 7.1.5.2 *Error terms and their variance*
- 7.1.6 Inference for Parameters
 - 7.1.6.1 *Standard errors and confidence intervals*
- 7.1.7 Model Comparison and Analysis of Variance
 - 7.1.7.1 *Balanced Designs*
 - 7.1.7.2 *Unbalanced Designs*
- 7.1.8 More on Main Effects and Interactions
- 7.1.9 Interactions and Transformations
- 7.1.10 Specific Comparisons of Marginal Means
- 7.1.11 Interpreting Interactions
 - 7.1.11.1 *Graphs*
 - 7.1.11.2 *Simple Main Effects*
 - 7.1.11.3 *Treatment–Contrast and Contrast–Contrast Interactions*
- 7.1.12 Predictor Effects
- 7.1.13 Assumptions
- 7.1.14 Robust Factorial ANOVAs
- 7.2 Complex Factorial Designs
 - 7.2.1 Missing Cells
 - 7.2.2 Fractional Factorial Designs
- 7.3 Power and Sample Size in Factorial Designs
- 7.4 Key points
- 7.5 Further Reading

Chapter 8 Multiple Regression Models

- 8.1 Linear Model for Multiple Continuous Predictors – Multiple Regression
 - 8.1.1 The Multiple Linear Regression Model
 - 8.1.2 Model Parameters
 - 8.1.1.1 *Intercept and Partial Regression Slopes*
 - 8.1.1.2 *Predicted Values*
 - 8.1.1.3 *Error Terms and Their Variance*
 - 8.1.1.4 *Standardized Partial Regression Slopes*
 - 8.1.3 Inference for Parameters
 - 8.1.1.5 *Standard Errors and Confidence Intervals*
 - 8.1.1.6 *Statistical Hypotheses*
 - 8.1.2 Model Comparison and Analysis of Variance
 - 8.1.3 Assumptions of Multiple Linear Regression Models
 - 8.1.4 Model Diagnostics
 - 8.1.4.1 *Leverage*
 - 8.1.4.2 *Residuals*
 - 8.1.4.3 *Influence*
 - 8.1.5 Diagnostic Graphics
 - 8.1.5.1 *Scatterplots*
 - 8.1.5.2 *Residual Plots*
 - 8.1.6 Transformations
 - 8.1.7 Collinearity
 - 8.1.7.1 *Detecting Collinearity*
 - 8.1.7.2 *Dealing with Collinearity*
 - 8.1.8 Interactions in Multiple Regression
 - 8.1.8.1 *Probing Interactions*
 - 8.1.9 Regression Models with Polynomial Terms
 - 8.1.10 Other Issues in Multiple Linear Regression
 - 8.1.10.1 *Regression Through the Origin*
 - 8.1.10.2 *Weighted (generalized) least squares*
 - 8.1.10.3 *X random (Model II regression)*
 - 8.1.10.4 *Robust Regression*
 - 8.1.10.5 *Missing Data*

- 8.1.10.6 *Power of Tests*
- 8.1.11 Categorical Predictors in Multiple Regression Models
- 8.2 Analysis of Covariance
 - 8.2.1 Linear Models for Simple Analyses of Covariance
 - Predicted Values and Residuals*
 - 8.2.2 Model Comparison and the Analysis of (Co)variance
 - 8.2.3 Assumptions of ANCOVA Models
 - 8.2.4 Homogeneous Within-group Regression Slopes
 - 8.2.4.1 *Evaluating Within-group Regression Slopes*
 - 8.2.4.2 *Dealing with Heterogeneous Within-group Regression Slopes*
 - 8.2.5 Robust ANCOVA
 - 8.2.6 Unequal Sample Sizes (Unbalanced Designs)
 - 8.2.7 Specific Comparisons of Adjusted Means
 - 8.2.7.1 *Planned comparisons*
 - 8.2.7.2 *Unplanned Comparisons*
 - 8.2.8 Factorial Designs
 - 8.2.9 Designs with Two or More Covariates
- 8.3 Key points
- 8.4 Further reading

Chapter 9 Predictor Importance and Model Selection in Multiple Regression Models

- 9.1 Relative Predictor Importance
 - 9.1.1 Single Model Methods
 - 9.1.1.1 *Standardized Partial Regression Slopes*
 - 9.1.1.2 *Tests on Partial Regression Slopes*
 - 9.1.2 Multiple Model Methods
 - 9.1.2.1 *Change in Explained Variation*
 - 9.1.2.2 *LMG and Hierarchical Partitioning*
 - 9.1.2.3 *Proportional Marginal Variance Decomposition (PVMD)*
 - 9.1.3 Recommendations of Relative Importance
- 9.2 Model Selection
 - 9.2.1 Model Selection Criteria
 - 9.2.1.1 *Comparisons to Full Model*
 - 9.2.1.2 *Information Criteria*
 - 9.2.2 Traditional Stepwise Selection
 - 9.2.3 All Subsets and Information Criteria
 - 9.2.4 Model Averaging
 - 9.2.5 Model Validation
- 9.3 Regression Trees
 - 9.3.1 Standard Regression Trees
 - 9.3.2 Bagging and Boosted Regression Trees
- 9.4 Key points
- 9.5 Further reading

Chapter 10 Random Factors in Factorial and Nested Designs

- 10.1 Fixed vs. Random Effects and Mixed Models
 - 10.1.1 Designs Applicable to Mixed Models
 - 10.1.1.1 *Single Random Factor Designs*
 - 10.1.1.2 *Nested or Hierarchical Designs*
 - 10.1.1.3 *Crossed and Block Designs*
 - 10.1.1.4 *Split-plot Designs*
 - 10.1.1.5 *Repeated Measures (RM) and Longitudinal Designs*
- 10.2 Fitting Linear Models with Fixed and Random Factors
 - 10.2.1 Traditional OLS “ANOVA” Models Approach
 - 10.2.1.1 *Estimation and Tests*
 - 10.2.1.2 *Assumptions and Diagnostics*
 - 10.2.1.3 *Overview*
 - 10.2.2 Linear Mixed Effect (or Multilevel) Models Approach

- 10.2.2.1 *Estimation and Tests*
 - 10.2.2.2 *Assumptions and Diagnostics*
 - 10.2.2.3 *Overview*
 - 10.2.3 Modeling Strategies
- 10.3 Simple Random Factor Designs
 - 10.3.1 Traditional OLS Approach
 - 10.3.2 Linear Mixed effects (Multilevel) Models
- 10.4 Multilevel Regressions
- 10.5 Nested (Hierarchical) Designs
 - 10.5.1 Two-level Nested Designs
 - 10.5.1.1 *OLS Analysis*
 - 10.5.1.2 *Mixed Effects Model Analysis*
 - 10.5.1.3 *Pooling and Model Selection in Nested Analyses*
 - 10.5.2 More Complex Nested Designs
 - 10.5.3 Sample Size and Nested Designs
- 10.6 Factorial Mixed Designs
 - 10.6.1 Types of Factorial Mixed Designs
 - 10.6.1.1 *General Factorial Mixed Designs*
 - 10.6.1.2 *Randomized Block Designs*
 - 10.6.2 Analysis of Crossed Designs with One Fixed and One Random Factor
 - 10.6.2.1 *OLS Analysis*
 - 10.6.2.2 *Linear Mixed Effects (Multilevel) Model*
 - 10.6.3 Crossed Designs with Two or More Fixed Factors and One Random Factor
 - 10.6.4 Design and Analysis Issues with Crossed Mixed Designs and Their Models
 - 10.6.4.1 *Number of Random Factor Groups*
 - 10.6.4.2 *Issues with Multiple Random Factors*
 - 10.6.4.3 *Efficiency of Blocking*
 - 10.6.4.4 *Missing Values in CB Designs*
 - 10.6.4.5 *Incomplete Block and Latin Square Designs*
- 10.7 Key points
- 10.8 Further reading

Chapter 11 Split-plot (Split-unit) Designs: Partly Nested Models

- 11.1 Three-Factor Split-Plot Designs
 - 11.1.1 Analysis for Three-Factor Split-Plot Designs
 - 11.1.1.1 *Linear Models for Split-Plot Designs*
 - 11.1.1.2 *ANOVA and Estimates of Effects*
 - 11.1.1.3 *Null Hypotheses*
 - 11.1.1.4 *Split-Plots with Subplot Replication*
 - 11.1.2 Assumptions
 - 11.1.3 Unbalanced Split-Plot Designs
 - 11.1.4 Model Building
- 11.2 More Complex Designs
 - 11.2.1 Additional Between-Plots Factors
 - 11.2.2 Additional Within-Plots Factors
 - 11.2.3 Including Continuous Covariates
- 11.3 Key points
- 11.4 Further reading

Chapter 12 Repeated measures designs

- 12.1 Simple Repeated Measures Designs
 - 12.1.1 Analysis of Simple Repeated Measures Designs
 - 12.1.1.1 *Linear Models for Simple Repeated Measures Designs*
 - 12.1.2 Assumptions for Simple Repeated Measures Models
 - 12.1.2.1 *Independence and Covariance Structures*
 - 12.1.3 Missing Observations
- 12.2 More Complex Repeated Measures Designs
 - 12.2.1 One Between-subjects Factor

- 12.2.2 Two or More Between-subjects Factors
- 12.2.3 Two or More Within-subjects Factors
- 12.2.4 Model Selection in Complex Repeated Measures Designs
- 12.3 Key Points
- 12.4 Further reading

Chapter 13 Generalized Linear Models for Categorical Responses

- 13.1 Logistic Regression
 - 13.1.1 Binary Response with a Single Continuous Predictor – Simple Logistic Regression
 - 13.1.2 Categorical Predictors in GLMs
 - 13.1.3 Binary Response with Multiple Predictors – Multiple Logistic Regression
 - 13.1.4 Nominal and Ordinal Multinomial Response Variables
 - 13.1.5 Proportion Response
- 13.2 Count Responses – Poisson Regression
- 13.3 Goodness-of-fit for GLMs
- 13.4 Inference for Parameters in GLMs
- 13.5 Assumptions and Diagnostics for Binomial and Poisson GLMs
- 13.6 Over-dispersed Data
 - 13.6.1 Identifying Overdispersion
 - 13.6.2 Correcting for Overdispersion
 - 13.6.2.1 *Quasi-likelihood (Quasi-Poisson) Models*
 - 13.6.2.2 *Negative Binomial Models*
 - 13.6.2.3 *Including Observation-level Random Effects*
 - 13.6.3 Too Many Zeros (Zero-inflated Data)
 - 13.6.4 Zero-truncated Data
 - 13.6.5 Binomial Overdispersion
- 13.7 Contingency Tables
 - 13.7.1 Two-way Tables
 - 13.7.1.1 *Test for Independence*
 - 13.7.1.2 *Odds and Odds Ratios*
 - 13.7.1.3 *Residuals*
 - 13.7.1.4 *Small Sample Sizes*
 - 13.7.1.5 *Loglinear Models*
 - 13.7.2 Three-way and Higher Tables
 - 13.7.2.1 *Three-way Interaction*
 - 13.7.2.2 *Conditional (In)dependence*
 - 13.7.2.3 *Joint Independence*
 - 13.7.2.4 *Marginal Independence*
 - 13.7.2.5 *Complete Independence*
 - 13.7.2.6 *Hierarchical Loglinear Modeling*
 - 13.7.3 More Complex Tables
 - 13.7.4 Loglinear Versus Logistic Models for Tables
- 13.8 Generalized Linear Mixed Models (GLMMs)
- 13.9 Generalized Additive Models
- 13.10 Key Points
- 13.11 Further Reading

Chapter 14 Introduction to Multivariate Analyses

- 14.1 Multivariate Data
- 14.1 Distributions and Associations
- 14.2 Linear Combinations, Eigenvectors, and Eigenvalues
 - 14.2.1 Linear Combinations of Variables
 - 14.2.2 Eigenvalues
 - 14.2.3 Eigenvectors
 - 14.2.4 Derivation of Components
- 14.3 Multivariate Distance and Dissimilarity Measures
 - 14.3.1 Dissimilarity Indices for Continuous and Count Variables
 - 14.3.1.1 *Metric Measures*

- 14.3.1.2 *Semimetric*
- 14.3.2 Dissimilarity Indices for Dichotomous (Binary) Variables
- 14.3.3 General Dissimilarity Indices for Mixed Variables
- 14.3.4 Choosing Dissimilarity Indices
- 14.4 Data Transformation and Standardization
- 14.5 Standardization, Association, and Dissimilarity
- 14.6 Screening Multivariate Data Sets
- 14.7 Introduction to Multivariate Analyses
- 14.8 Key points
- 14.9 Further reading

Chapter 15 Multivariate analyses based on eigenanalyses

- 15.1 Principal components analysis
 - 15.1.1 Deriving components
 - 15.1.1.1 *Axis rotation*
 - 15.1.1.2 *Decomposing an association matrix*
 - 15.1.2 Interpreting the components
 - 15.1.3 How many components to retain?
 - 15.1.3.1 *Eigenvalue equals one rule*
 - 15.1.3.2 *Scree diagram*
 - 15.1.3.3 *Broken stick criterion*
 - 15.1.3.4 *Tests of eigenvalue equality*
 - 15.1.3.5 *Other methods*
 - 15.1.4 Which association matrix to use?
 - 15.1.5 Simplifying component structure
 - 15.1.6 PCA assumptions and “fit”
 - 15.1.6.1 *Assumptions*
 - 15.1.6.2 *PCA fit: residuals*
 - 15.1.7 Ordination and biplots for PCA
 - 15.1.8 Principal components regression (PCR)
 - 15.1.9 Factor analysis
- 15.2 Correspondence analysis
 - 15.2.1 Deriving the axes
 - 15.2.2 Ordination and biplots for CA
 - 15.2.3 Reciprocal averaging
- 15.3 Use of PCA and CA with ecological abundance (count) data
- 15.4 Constrained (canonical) multivariate analysis
 - 15.4.1 Redundancy analysis (RDA)
 - 15.4.2 Canonical correspondence analysis
- 15.5 Linear discriminant function analysis
 - 15.5.1 Deriving discriminant functions
 - 15.5.2 Classification and prediction
 - 15.5.3 Assumptions of discriminant function analysis
 - 15.5.4 Multivariate analysis of variance (MANOVA)
- 15.6 Key Points
- 15.7 Further reading

Chapter 16 Multivariate analyses based on (dis)similarities or distances

- 16.1 Multidimensional scaling or ordination
 - 16.1.1 Classical (metric) scaling – principal coordinates analysis (PCoA)
 - 16.1.2 Nonmetric (enhanced) multidimensional scaling
 - 16.1.2.1 *Deriving the ordination*
 - 16.1.2.2 *Interpretation of ordination plot*
- 16.2 Cluster analysis
 - 16.1.1 Agglomerative hierarchical clustering
 - 16.2.1 Divisive hierarchical clustering
 - 16.2.2 Non-hierarchical clustering
- 16.3 Analyses based on dissimilarities

- 16.3.1 Contributions of original variables to ordination
- 16.3.2 Relating dissimilarities to other variables
 - 16.3.2.1 *Mantel test*
 - 16.3.2.2 *Bio-env*
 - 16.3.2.3 *Matrix regression*
 - 16.3.2.4 *Analysis of similarities*
 - 16.3.2.5 *Multi-response permutation procedures*
 - 16.3.2.6 *Comparing dispersions*
- 16.1.2 Multivariate linear models
 - 16.3.2.7 *Distance-based redundancy analysis*
 - 16.3.2.8 *Permutational multivariate analysis of variance (PERMANOVA)*
 - 16.3.2.9 *MV-ABUND*
- 16.4 Key Points
- 16.5 Further reading

Chapter 17 Telling stories with data

- 17.1 Research doesn't exist until you tell someone
 - 17.1.1 Telling better stories: the importance of narrative
- 17.2 Summarizing data analyses
 - 17.2.1 Linear models
 - 17.2.1.1 *Continuous predictors*
 - 17.2.1.2 *Categorical predictors*
 - 17.1.1 Other analyses
- 17.3 Visualizing data
 - 17.1.2 Just show us the numbers!
 - 17.1.3 Tables
- 17.4 Graphical summaries of the data
 - 17.4.1 Some basic principles for visualizing data
 - 17.4.1.1 *Focus attention where you want it*
 - 17.4.1.2 *Eliminate clutter*
 - 17.4.1.3 *White (blank) space is good*
 - 17.4.1.4 *The nitty-gritty: scales, ticks, labels and legends*
 - 17.4.2 An appropriate visual display
 - 17.4.2.1 *Bar graph*
 - 17.4.2.2 *Line graph*
 - 17.4.2.3 *Scatterplots*
 - 17.4.2.4 *Pie charts*
 - 17.4.2.5 *Slope charts*
 - 17.4.2.6 *The challenge of -omics and other big data*
- 17.5 ~~Error bars~~: visualizing variation and precision
 - 17.5.1 Possible solutions
- 17.6 Horses for courses: what you present depends on who's listening
 - 17.6.1 Know your audience as well as possible
 - 17.6.1.1 *Talks*
 - 17.6.1.2 *Conferences and seminars*
 - 17.6.1.3 *Talks with handouts (lecture, briefing, etc.)*
 - 17.6.1.4 *Written stories (papers, theses, etc.)*
- 17.7 Software and other sources
- 17.8 Key points
- 17.9 Further reading

Chapter 18 References