Contents

Preface *ix*
Acknowledgements *xi*

1 Introduction *1*
1.1 The nature of paleontological data *1*
1.2 Advantages and pitfalls of paleontological data analysis *5*
1.3 Software *7*
References *8*

2 Statistical concepts *9*
2.1 The population and the sample *9*
2.2 The frequency distribution of the population *9*
2.3 The normal distribution *11*
2.4 Cumulative probability *12*
2.5 The statistical sample, estimation of distribution parameters *14*
2.6 Null hypothesis significance testing *16*
2.7 Bayesian inference *20*
2.8 Exploratory data analysis *22*
References *22*

3 Introduction to data visualization *24*
3.1 Graphic design principles *24*
3.2 Line charts *25*
3.3 Scatter plots *26*
3.4 Histograms *26*
3.5 Bar chart, box, and violin plots *29*
3.6 Normal probability plot *29*
3.7 Pie charts *31*
3.8 Ternary plots *32*
3.9 Heat maps, 3D plots, and Geographic Information System *33*
3.10 Plotting with R and Python *33*
References *37*

4 Univariate and bivariate statistical methods *38*
4.1 Parameter estimation and confidence intervals *38*
4.2 Testing for distribution *40*
4.3 Two-sample tests *43*
4.4 Multiple-sample tests *52*
4.5 Correlation *58*
4.6 Bivariate linear regression *64*
4.7 Generalized linear models *70*
4.8 Polynomial and nonlinear regression *73*
4.9 Mixture analysis *74*
4.10 Counts and contingency tables *76*
References *78*

5 Introduction to multivariate data analysis *81*
5.1 Multivariate distributions *82*
5.2 Parametric multivariate tests – Hotelling's T^2 *82*
5.3 Nonparametric multivariate tests – permutation test *85*
5.4 Hierarchical cluster analysis *86*
5.5 K-means and k-medoids cluster analysis *92*
References *94*

6 Morphometrics *96*
6.1 The allometric equation *97*
6.2 Principal components analysis *101*
6.3 Multivariate allometry *108*
6.4 Linear discriminant analysis *112*
6.5 Multivariate analysis of variance *116*
6.6 Fourier shape analysis in polar coordinates *116*
6.7 Elliptic Fourier analysis *119*
6.8 *Hangle* Fourier analysis *122*
6.9 Eigenshape analysis *123*
6.10 Landmarks and size measures *125*
6.11 Procrustes fitting *127*
6.12 PCA of landmark data *130*
6.13 Thin-plate spline deformations *132*
6.14 Principal and partial warps *136*
6.15 Relative warps *139*
6.16 Regression of warp scores *141*
6.17 Common allometric component analysis *142*
6.18 Landmarks in 3D *143*
6.19 Disparity measures *144*
6.20 Morphogroup identification with machine learning *146*
6.21 Case study: the ontogeny of a Silurian trilobite *153*
References *157*

7 Directional and spatial data analysis *162*
7.1 Analysis of directions and orientations in 2D *162*
7.2 Analysis of directions and orientations in 3D *164*
7.3 Spatial point pattern analysis *166*
References *173*

8 Analysis of tomographic and 3D-scan data *174*
8.1 The technology of x-ray tomography *174*
8.2 Processing of volume data *175*
8.3 Functional morphology with 3D data *180*
References *182*

9 Estimating paleobiodiversity *184*
9.1 Species richness estimation *185*
9.2 Rarefaction and related methods *187*
9.3 Diversity curves, origination, and extinction rates *192*
9.4 Abundance-based biodiversity indices *196*
9.5 Taxonomic distinctness *202*
9.6 Comparison of diversity indices *207*
9.7 Abundance models *208*
References *212*

10 Paleoecology and paleobiogeography *216*
10.1 Paleobiogeography *216*
10.2 Paleoecology *217*
10.3 Association similarity indices for presence-absence data *219*
10.4 Association similarity indices for abundance data *223*
10.5 ANOSIM and PerMANOVA *228*
10.6 Principal coordinates analysis *229*
10.7 Non-metric multidimensional scaling *232*
10.8 Correspondence analysis *236*
10.9 Detrended correspondence analysis *240*
10.10 Seriation *242*
10.11 Nonlinear dimensionality reduction *245*
10.12 Canonical correspondence analysis *248*
10.13 Indicator species *251*
10.14 Network analysis *252*
10.15 Size-frequency and survivorship curves *254*
10.16 Case study: Devonian paleobiogeography *256*
References *259*

11 Calibration – estimating paleoenvironments *263*
11.1 Modern analog technique *263*
11.2 Weighted averaging *265*
11.3 Weighted averaging partial least squares *267*

11.4 Which calibration method? *269*
11.5 Case study: Late Holocene temperature inferred from chironomids *271*
References *271*

12 Time series analysis *273*
12.1 Spectral analysis *274*
12.2 Wavelet analysis *282*
12.3 Autocorrelation *284*
12.4 Cross-correlation *287*
12.5 Runs test *290*
12.6 Time Series Trends and Regression *291*
12.7 Smoothing and filtering *293*
References *297*

13 Quantitative biostratigraphy *299*
13.1 Zonation of a single section *299*
13.2 Confidence intervals on stratigraphic ranges *301*
13.3 Regional and global biostratigraphic correlation *304*
13.4 Age models *330*
References *335*

14 Phylogenetic analysis *338*
14.1 A dictionary of cladistics *338*
14.2 Parsimony analysis *339*
14.3 Characters *341*
14.4 Algorithms for Parsimony Analysis *342*
14.5 Character state reconstruction *347*
14.6 Evaluation of characters and trees *348*
14.7 Case study: the systematics of heterosporous ferns *355*
14.8 Other methods for phylogenetic analysis *359*
14.9 Phylogenetic Comparative Methods *362*
References *368*

Index *371*