DOWNLOAD PDF ANALYSING ECOLOGICAL DATA (STATISTICS FOR BIOLOGY AND HEALTH)

Chapter 1 : Analysing Ecological Data: Alain F Zuur, Elena N Ieno, Graham M Smith | NHBS Book Shop

This book provides a practical introduction to analyzing ecological data using real data sets. The first part gives a largely non-mathematical introduction to data exploration, univariate methods (including GAM and mixed modeling techniques), multivariate analysis, time series analysis, and spatial statistics.

Principal component analysis PCA , correspondence analysis CA , discriminant analysis DA and non-metric multidimensional scaling NMDS can be used to analyse data without explanatory variables, whereas canonical correspondence analysis CCA and re- dundancy analysis RDA use both response and explanatory variables. Most ordination techniques create linear combinations of the variables: The linear combination can be imagined as multiplying all elements in a column with a par- ticular value, followed by a summation over the columns. The idea of calculating a linear combination of variables is perhaps difficult to grasp at first. However, think of a diversity index like the total abundance. This is the sum of all variables all ctjS are one , and summarise a large number of variables with a single diversity index. The underlying idea is that the most im- portant features in the N variables are caught by the new variable Z b Obviously one component cannot represent all features of the N variables and a second com- ponent may be extracted: Most ordination techniques are designed in such a way that the first axis is more important than the second, the second more important than the third, etc. It should be noted that the variables in equa- tions They are either centred which means that the mean of each variable is subtracted or normalised centred and then divided by the standard deviation , and the implication of this is dis- cussed later in this chapter. The multiplication factors Cy are called loadings. These latter vari- ables are considered as explanatory variables, so we are modelling a cause-effect relationship. Two easy explanations PCA is one of the oldest and most commonly used ordination methods. The reason for its popularity is perhaps its simplicity. Before introducing PCA, we need to clear up some possible misconceptions. PCA cannot cope with missing values but neither can most other statistical methods , it does not require normal- ity, it is not a hypothesis test, there are no clear distinctions between response variables and explanatory variables, and it is not written as principle component analysis but as principal component analysis. There are various ways to introduce PCA. Our first approach is based on Shaw , who used the analogy with shadows to explain PCA. Imagine giving a presentation. If you put your hand in front of the overhead projector, your three- dimensional hand will be projected on a two-dimensional wall or screen. The chal- lenge is to rotate your hand such that the projection on the screen resembles the original hand as much as possible. This idea of projection and rotation brings us to a more statistical approach of introducing PCA. There is a clear negative relationship between the two variables. Panel B shows the same scatterplot, except that both variables are now mean deleted and divided by the standard deviation also called normalisation. We also used axes with the same range from Now suppose that we want to have two new axes such that the first axis represents most information, and the second axis the second most information. The di- agonal line from the upper left to the lower right in Figure Projecting the points on this new axis results in the axis with the largest vari- ance; any line with another angle with the x-axis will have a smaller variance. The additional restriction we put on the new axes is that the axes should be perpen- dicular to each other. Hence, the other line perpendicular to the first new axis in the same panel is the second axis. Panel C shows the same graph as panel B ex- cept that the new axes are presented in a more natural direction. This graph is Two easy explanations called a PCA ordination plot. In the same way as the three-dimensional hand was projected on a two-dimensional screen, we can project all the observations onto the first PCA axis, and omit the second axis Figure In applying PCA, one hopes that the variance of most components is negligible, and that the variation in the data can be described by a few independent principal components. Two technical explanations Having presented two relatively easy introductions to PCA, we feel it is neces- sary to present PCA in a mathematical context as well. Readers not familiar with matrix algebra may skip this section and continue with the illustrations of PCA in Section We discuss two mathematical derivations of PCA. PCA as an eigenvalue decomposition More details on the technical aspects of PCA can be found in Jolliffe , and only a short summary is presented below. An aspect we have not mentioned so far is that the loadings are not unique. Therefore, a restriction on the loadings is needed to make the solution unique: Although it seems rather arbitrary to set the sum of the squared loadings to 1, it does not have any important consequences. It is relatively easy to show that the loadings can be obtained by solving the following constrained optimisation problem. For the first axis, we have: The maximisation is with respect to the unknown parameters Ci. This may seem like magic to readers not used to matrix algebra, but this expression is fundamental in mathematics and is called the eigen- value equation for S, X is the eigenvalue and Ci the eigenvector. S is either the co-variance or correlation matrix, depending on whether the variables in Y are cen- tred or normalised. This process can easily be extended to get the second or higher axes. Statistical software can be used to obtain C which can be used to obtain the axes. The motivation to present the eigenvalue equation for PCA is that it justifies the iterative algorithm presented in the next paragraph, and this algorithm is used to explain RDA in the next chapter. PCA as an iterative algorithm The last approach to explain PCA is by using an iterative algorithm, which was presented in Jongman et al. The algorithm has the following steps. Normalise or centre the variables in Y. Obtain initial scores z e. For second and higher axes: Make z uncorrelated with previous axes using a regression analysis. Scale z to unit variance: Repeat steps 2 to 6 until convergence. Once the algorithm is finished, the loadings c and principal components z scores are identical to those obtained from the PCA eigenvalue equation. This data- set consists of seven morphological variables taken from approximately sparrows. The morphological variables were wingcrd, flatwing, tarsus, head, cul- men, nalopsi and weight wt ; see Section As the PCA will be based on the correlation matrix, it may be a good idea to in- spect the correlation matrix before applying PCA. For the sparrow data, the corre- lation coefficients are relatively high Table The lengths of the wing are measured in two different ways, as the wing chord wingcrd and as the flattened wing flatwing. Therefore the correla- tion between them is high, however the PCA can deal with this. Hence, for the moment we will use both variables. However, the problem is the interpretation of this graph. This is typical for morphometric data; very often the first axis represents the overall shape of the animals. A more detailed discussion on morphometric data analysis is given in Chapter The second axis seems to represent differences between wingcrd and flatwing versus culmen and nalospi. First two axes obtained by PCA for the sparrow data. They can be expressed as numbers, percentage of the total variance, or as cumula-tive percentage of the total variance Table Some software packages present the eigenvalues that come out of the numerical routines, whereas others scale them so that the total sum is equal to one. The eigenvalues in the first column of Table The first eigenvalue is 0. The second eigenvalue is 0. Eigenvalues and eigenvalues expressed as cumulative percentage. Some software packages rescale the eigenvalues so that the sum of all eigenvalues is equal to 1. These are given in the second column. Unsealed eigenvalues are in the third column. The sum of all eigenvalues is 7. Axis 1 2 3 4 Eigenvalue scaled 0. In this case, two axes are sufficient. Another option is a scree plot of the eigenvalues Figure In such a graph, the eigenvalues are plotted as vertical lines next to each other. A scree plot of the eigenvalues would then show which of the axes are important long lines and the change point elbow at which the axes become less important. In this case, one axis is sufficient. Some software packages use barplots or just lines in the scree plot. Yet, another tool is the broken stick model approach Jolliffe ; Legendre and Legendre In this case, the broken stick values of the first three axes are 0. So based on the broken stick model, only the first axis is of interest. Other tools to select the number of axes e. If the first few axes explain a low percentage, then it might be worthwhile to investigate whether there are outliers, or whether the relationships between variables are non-linear. If either occurs, then consider a transformation or accept the fact that ecological data are genuinely noisy. The problem with the PCA ordination plot is that is does not tell us whether there are differences between groups of observations e. Indeed, many authors do this. But with prior knowledge on a grouping structure in the observations avail-able, this is not recommended as other statistical techniques can and should be applied for this purpose, for example discriminant analysis or classification mod- els. In the next section, an extension of the ordination diagram is presented, which may give more visual information, namely the biplot. Eigenvalues for the sparrow data, obtained by PCA, are presented as vertical lines.

Chapter 2 : calendrierdelascience.com:Customer reviews: Analysing Ecological Data (Statistics for Biology

They offer step by step analyses of different ecological data sets using different types of regression (linear, logistic, etc.), additive models, tree models, multivariate models, factor analysis, time series, spatial analysis, and others.

Ecologists are notoriously passionate about their subject and most, if not all, use statistics to support their field and laboratory studies. Very few ecologists are statisticians and although many of us have had some statistical training as part of our first and higher degrees, few can claim a high level of statistical expertise with a great deal of confidence. As a result there is a fair number of books written to help ecologists and biologists select and use appropriate statistical techniques. These range from books examining the basis of experimental design e. Underwood to those that are written to support particular statistical programmes e. Although written with a view to supporting Brodgar their own statistical programme the authors stress that any of the techniques discussed in the book and all the accompanying example datasets housed on a website can be analysed using R a freely available and extremely powerful, if not a little intimidating, statistical library and scripting language. This is certainly true but researchers with only limited to intermediate statistical knowledge would certainly benefit from using Brodgar itself as the windows menu driven approach with its hard-coded R scripts plus a range of programmed techniques makes for a much easier introduction to the topic. The book has thirty-seven chapters loosely structured into three main areas: These are essential first steps for any researcher which should be completed before selecting which analytical techniques are available and appropriate for their field or experimental data. Are your data linear? Are they normally distributed? Are there any outliers? These are important questions and have implications for the sorts of techniques you can legitimately select for the data crunching later on. The chapter on exploration Chapter 4 illustrates a range visualisation tools that are available within the R statistical library and it is especially useful. My experience is that this is an often overlooked step in the research process so it is very positive to see it so heavily emphasized in this book. The second section introduces readers to the multitude of techniques available to the willing ecologist, starting with Linear Regression developing through to complex Generalised Linear Models GLMs and their non-parametric cousins General Additive Models GAMs and concluding with Mixed Models. Thereafter, ordination techniques e. There is a strong sense of development throughout the book and the third element of the mix introduces the reader to a sequence of case studies of the use of many of the aforementioned techniques. The datasets themselves are available for download and experimentation on the website that supports the Brodgar programme and now the book. The case studies are an excellent addition and reinforce the preceding chapters that guide readers in their selection of appropriate ways to analyze datasets from a variety of sample designs with markedly different structures. This is first statistical text I have read I have many on my shelves that takes the reader from the starting point of sample design, through data visualisation to technique selection, all supported by case studies using real field data. It does this in a clear and systematic manner and is well-written throughout. These early chapters are excellent and having read them I felt much more secure and confident about the quality and structure of my own datasets - it was like having a light switched on in a dark room. The other major strength is the ground that the book covers - it is very wide-ranging and draws a much broader range of univariate and multivariate techniques than most of its competitor texts. If you want more confidence in your own analyses or just a deeper understanding of the statistical techniques you routinely use and others you could use this book is for you. I cannot praise it enough; copies are doing the rounds in my research team as I type. One person found this helpful.

DOWNLOAD PDF ANALYSING ECOLOGICAL DATA (STATISTICS FOR BIOLOGY AND HEALTH)

Chapter 3 : Analysing Ecological Data (Statistics for Biology and Health) (May 3, edition) | Open Library

download analysing ecological data statistics for biology and health This book provides a practical introduction to analysing ecological data using real data sets collected as part of postgraduate ecological studies or research projects.

In stock Clearance price: The first part of Analysing Ecological Data gives a largely non-mathematical introduction to data exploration, univariate methods including GAM and mixed modelling techniques , multivariate analysis, time series analysis e. The second part provides 17 case studies, mainly written together with biologists who attended courses given by the first authors. The case studies include topics ranging from terrestrial ecology to marine biology. The case studies can be used as a template for your own data analysis; just try to find a case study that matches your own ecological questions and data structure, and use this as starting point for you own analysis. The book would be a suitable companion to statistics courses for both ecologists and statisticians at the introductory graduate level!. All in all, I enjoyed reading the book and marvel at the wide range of sophisticated statistical models used in modern ecology. Undergraduates, postgraduates and scientists engaged in areas of the environmental sciences and ecological research. The material presented in this book has been developed and used by the authors in teaching statistics to its intended readership. The text is divided into two parts [I have no doubt that for undergraduate students the main strength of the book will be the breadth of topics covered by the case studies â€" ranging from terrestrial ecology to marine biology. Who is it for? Upper undergraduates, postgraduates and researchers in ecology. Presentation It links ecological data, data analysis and discussion of the approaches. Would you recommend it? If you want an edited volume on different methods of ecological data analysis, then this book is worth looking through. No other book combines as many good ecological data sets with such thoughtfully written analyses. I give this book two enthusiastic thumbs up! This is an excellent, nicely presented and very readable book. I would highly recommend it to numerate researchers and students interested in environment and ecological data analysis. In summary, I can recommend the book primarily as advanced material for ecologists [Dormann, Basic and Applied Ecology, Vol.

Chapter 4 : ãƒŸã‚ºãƒŽãƒ—ãƒ- ãƒŸã‚ºãƒŽ 54ï¼ºï¼¥ï¼´ï¼´ï¼ˆã‚¼ãƒƒãƒˆï¼‰ ã‚¢ãƒ³ãƒ'ã‚¤ãƒ¤ãƒ'ãƒ³ãƒ„ ï¾

Analysing ecological data by, May 3, , Springer edition, in English.