

Practical Stats Newsletter for April 2007

From Website: <http://www.practicalstats.com>

In this newsletter:

1. New courses, New website, new macros
2. Nonmetric Multidimensional Scaling
3. Untangling Multivariate Relationships course, Nov. 1-2, 2007

1. New courses, New website, new macros

Since our January newsletter there are many new things here at Practical Stats. First, our new course on multivariate statistics will be offered Nov. 1-2 in Golden, Colorado. It takes a practical look at how multivariate methods such as principal component analysis can be used for geochemical and biological/ecological data. How a suite of trace element concentrations can be related to a series of counts of different species, for example. As always, we focus on when to use different methods, and when not to, rather than derivations. Much more detail is on our website. Hope to see some of you there.

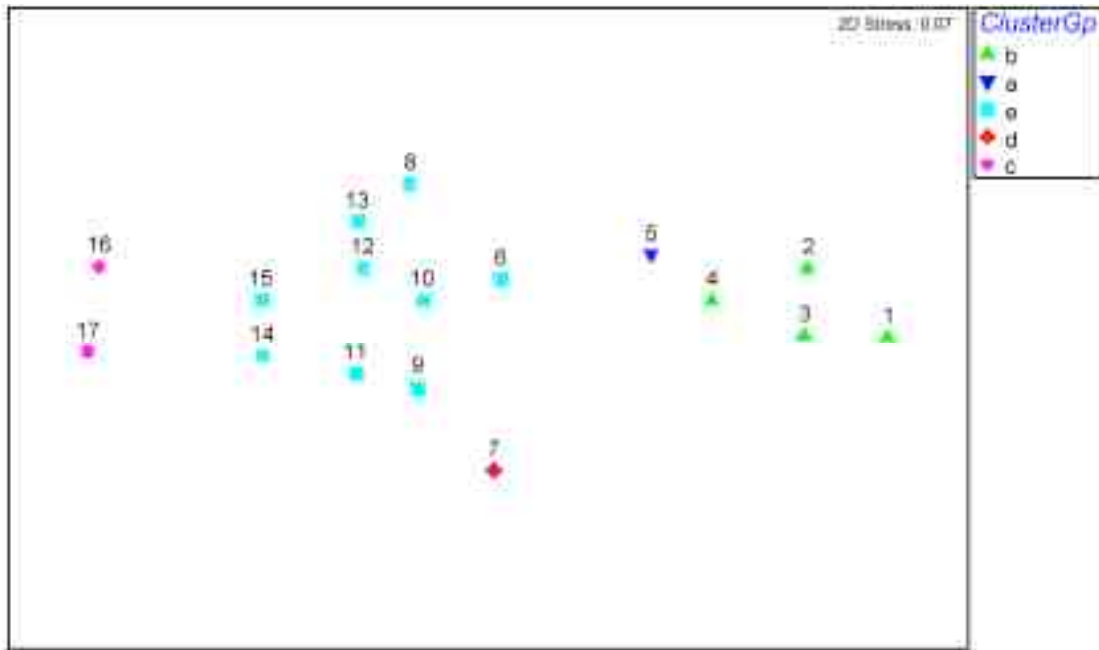
Our website has been reorganized and reformatted, with a new look. We hope it is even easier to find what you're looking for. Descriptions of our three courses and downloads of software and macros go with old standbys like the guide to 'Which method do I use?' and the evaluation of Microsoft Excel as a statistics package. Please send us any feedback or suggestions for what you'd like to see there. A convenient contact form is also on the site.

A brand new macro has been added to the NADA macros for nondetects. The kmstats macro computes Kaplan-Meier statistics for left-censored data. It does all the 'flipping' and manipulation within the macro - you simply feed it the vectors of data and censoring indicators. This macro was previously only given out in the Nondetects And Data Analysis course, but a modification of it to perform the Efron bias correction was added. Rather than contact all the past students directly, the new version of the macro is now included in the NADA for MTB (Minitab) collection on our site, and available to all for the first time. The current version of the macros is 2.3. Starting with version 2.2, all macros run in both Minitab 14 and 15. Download the latest version of the macros, if you haven't recently.

2. Nonmetric Multidimensional Scaling

Nonmetric Multidimensional Scaling (NMDS or "Nomads") is a graphical window into multidimensional data. It portrays the inter-relations between variables and groups, so that similar data or variables are plotted near each other, and dissimilar ones plotted further away. Distances are ranked distances, so that there is no continuous scale or "metric" on either the X or Y axis. The closest two observations in multidimensional space receive a distance rank of 1, and will plot closest to each other.

NMDS is most useful in picturing similarities between locations or variables, and in illustrating a gradient as points move from one part of the plot to another. Fig 1 shows a classic dataset, for example, where a gradient in light is available going from locations 1 to 17. At each location, counts of plants for about 20 different species were measured. NMDS was constructed using the patterns of species occurring at each site, and did not use the measure of light availability in constructing the plot. Yet the progression from location 1 to 17 is obvious. Also evident is that there are differences in species patterns between sites 7 and 8 that are secondary to the overall pattern, but still quite evident.



NMDS can be based on any of a number of similarity/dissimilarity (distance) measures, some more appropriate to chemicals and others to species. In our Untangling Multivariate Relationships course, we discuss distance measures such as Bray-Curtis, Euclidean and Manhattan, and when to use which. NMDS differs from a plot of principal component scores in that it is not a projection down from multivariate space, but a picture of ranked distances. This has both advantages and disadvantages when compared to principal components, and the two types of plots should be considered complimentary.

NMDS uses an iterative search, solving an optimization process to get the best picture in two or three dimensions. Several software packages compute NMDS, including SAS, Systat, NCSS, Stata and Primer. It is one of the basic tools to make multidimensional data more understandable. It is also a great tool for illustrating the results of multivariate tests such as MANOVA or its nonparametric analogues, comparing differences between groups of multidimensional data. Look for this capability in software if you expect to simultaneously deal with a collection of species or chemicals.

3. Untangling Multivariate Relationships course, Nov. 1-2, 2007

At the Colorado School of Mines campus, in Golden Colorado. See the website for online registration and details.

'Til next time,

Practical Stats

<http://www.practicalstats.com>

-- Make sense of your data