

Practical Stats Newsletter for Summer 2006

In this newsletter:

1. Free Statistical Software: R
2. R software for nondetects
3. Getting Started with R and NADA for R

1. Free Statistical Software: R

This is our second of two newsletters on free statistical software. Unlike Dataplot (reviewed in our April newsletter) which has not been widely used, R statistical software is used throughout the world, at leading universities, and in many corporations. You'll find many textbooks on R, and contributed packages to enhance R's capabilities are actively being developed. R has its roots in statistical routines called S, authored at the former Bell Labs. S was a series of Unix commands to perform methods routinely applied by statisticians. To make it accessible to more people, S has evolved in two directions. The first is commercial software called S-Plus, where a GUI interface has been put over top of the routines. The second is R, a free implementation of the S language. R uses a UNIX-like command structure through which complex calculations can be performed. R runs on Unix systems, Macintosh and Windows. All three use the same command structure, so learning R on one is directly applicable to the others. Writing a series of R commands is roughly equivalent to writing a SAS program - not as simple as pull-down menus, but well within the capabilities of a wide range of environmental scientists. And R's breadth of procedures and price can't be beat.

2. R software for nondetects

One of the online reviews of "Nondetects And Data Analysis" by Helsel (Wiley publishers, 2005) at Amazon (Amazon's price on the book is consistently lower than elsewhere) was very complimentary, and then stated

"It would also be nice if some of the methods were demonstrated in R as well as in Minitab. Alas, we cannot have all that we want."

Well, now you can! R commands to compute descriptive statistics, perform hypothesis tests, and build regression models, all for data that include nondetect values (left-censored data), are available through the CRAN site in the contributed package "NADA". NADA for R was developed by scientists at the U.S. Geological Survey, based on the NADA book. Both parametric and nonparametric methods are included. To download the package, see below.

Example use of the package is given through exercises in a pdf file on the NADA web site, at <http://practicalstats.com/nada/nada4R.htm> . The NADA for R Exercises file is the handout for the recent 1-day workshop presented to the American Statistical Association at its 2006 meeting. In it you will find example use of all the NADA for R

functions, including cenfit, cenreg, cenmle, and cenken. The NADA for R functions parallel the Minitab functions that have been available on the NADA site since publication of the book in 2005.

3. Getting Started with R and NADA for R

To download R software, point your browser to <http://www.r-project.org> . Follow the directions there for download and installation of the base package. The default answers for all installation questions work fine, placing the software in your Programs (Applications on Mac) folder. Start the program and it will tell you how to view demos of the program, to get familiar with its use.

For documentation, there is a free "Introduction to R" available through the left-hand links on the CRAN site. Click on Manuals and look for the introduction. The FAQs available there may also be of interest. You will probably also want a book to help you start using the program. There are numerous such books on Amazon and elsewhere. One that we have found helpful is "Introductory Statistics with R" by Peter Dalgaard, but there are many others.

Contributed packages such as NADA for R can be downloaded and installed from within the R software. Under the "Packages" menu, select a nearby CRAN mirror for download. Then select "Install Packages..." and choose NADA from the alphabetized list. Its that easy. Or on the command line at the > prompt, type in
> install.packages(NADA)

If you find the pdf file of NADA exercises on the NADA web site to be helpful, but want more, look for a 2007 offering of our NADA training course. The next offering will be posted on the Practical Stats website, and announced in an upcoming newsletter. At the training course we demonstrate how and why to use NADA methods for estimating statistics, performing hypothesis tests, and regression for data with nondetects.

'Til next time,

Practical Stats
<http://www.practicalstats.com>

-- Make sense of your data