Practical Stats January 2007 Newsletter

In this newsletter:
1. 2007 Courses
2. Equivalence Tests
3. New course: Unraveling Multivariate Relationships

Well, the snow here in Denver has delayed many things during the past month, including our newsletter. We've had a total well over 4 feet, coming on most every weekend for the last five. More snow due tomorrow. But our course schedule is shaping up, including our newest course, Unraveling Multivariate Relationships.

1. 2007 Courses
We have two "open enrollment" courses scheduled, with an offering of our new course planned for this fall. More information on course contents, including how to register, are always available at http://www.practicalstats.com

Nondetects And Data Analysis March 12, 2007 Golden, Colorado
We're offering NADA in a one-day format this year, at a reduced cost. We hope that makes it more enticing for more people. NADA provides an overview of methods for computing descriptive statistics, testing hypotheses, and computing regression models with values below detection limits. Numerous studies have shown that quick and dirty methods such as substituting one-half the detection limit do not work well. NADA presents what else is possible. Software and (optionally) a textbook are included. Given the weather for the past month, the skiing here will likely still be great the weekend prior to class!

Applied Environmental Statistics ~~June 11-14, 2007 Austin, Texas~~ [cancelled]
AES is our flagship survey of applied statistics. We cover methods for how to "make sense of your data", including testing for differences in groups, how to build a good regression model, and trend analysis. A full outline is available on our web site.

Unraveling Multivariate Relationships Fall, 2007 Place TBD
UMR is our newest course. A description is given below in #3. Our goals for this course are the same as for our other two:
* Make each method understandable
* Inform when to use, and when not to use, each method
* Provide insight on how each method works
* Motivation, communication, not derivation.

We hope to see you at one of our courses this year.

2. Equivalence Tests

We teach special topics when requested as part of our AES course. Teaching something different helps us to stay fresh, and keeps us in touch with newer topics of interest to the environmental community. One topic steadily requested of late is equivalence testing. Equivalence tests are a standard procedure in drug testing, and are being increasingly recognized in environmental applications. For example, when comparing two groups of paired data to determine whether they are the same or different, the standard test is the paired t-test (Helsel and Hirsch, 2002). The null hypothesis for that test is that the mean of differences between pairs of data, $D_i = x_i - y_i$, is zero. If plotted on a number line, the true value for D occupies a razor-thin line right at zero when the null hypothesis is true. Every other value for D is part of the alternative hypothesis.

The first argument against using traditional tests such as the paired t-test is that we don't necessarily expect the differences between pairs to be 0, but merely small. So the first reason to use an equivalence test is to match the test to what we know to be reality. We don't want to test an unrealistic situation. A second reason is encountered when there is a lot of data. As the numbers of observations increase, smaller and smaller differences can be detected as nonzero. The null hypothesis might be rejected when the true difference is very small, but not zero. In practice these small differences are often not of concern, and the two groups of data should be considered equivalent for practical purposes. It would be preferable to test whether the true difference is small, and therefore the two groups equivalent, or not. This is the test of equivalence.

The null hypothesis for an equivalence test occupies the range of values for the difference that is considered small by the investigator. Rather than a razor thin line for the null hypothesis, an equivalence interval is established. Differences within this interval result in the null hypothesis not being rejected, and the two groups are considered equivalent. Differences outside the interval result in rejection of the null hypothesis, and the two groups are declared "not equivalent".

To test for equivalence, the scientist must first set the boundaries of the equivalence interval. Boundaries might be in the units of measurement (plus and minus 5), or in percent. They need not be the same distance out from zero in both a positive and negative direction, though they often are. Once those boundaries are set, two hypothesis tests are performed, one at each side of the interval. If either of these tests is rejected, the difference is beyond the interval boundary, and equivalence is rejected. If neither test rejects the null hypothesis, the group difference lies inside the equivalence interval, and equivalence of the two groups is not rejected.

If few data are collected, equivalence will be difficult to reject. If equivalence means that "water quality is good" or "contamination is low", the difficulty in rejecting this assumption of equivalence may be of concern from a human or ecological health standpoint. With few data, contamination might not be detected, and the assumption of equivalence might be misunderstood to state that everything is OK, when in reality there is just little data to show it is not. To avoid merely assuming that "all is OK", the equivalence test may be set up in the reverse direction, so that the null hypothesis is inequivalence, all is "not OK", so that the two groups are assumed to be not equivalent. The burden of proof requires collecting enough data to demonstrate that all is OK.

Equivalence tests can be run for any estimation problem. In AES we also show how they might be used in trend analysis, to test for whether measurements stay equivalent over time. If not, there is a trend. But small rates of change are not of concern.
If you'd like more information on equivalence tests, a major advocate for them in environmental studies is Graham McBride of New Zealand's NIWA. Philip Dixon of Iowa State Univ. has applied them to trend investigation. Try Googling (not in my spell checker!) their names to find more information. Or come to our AES course in June.

3. New Course: Unraveling Multivariate Relationships
Environmental measurements often are simultaneously recorded on multiple scales. Multiple chemicals are analyzed. Multiple measures of ecosystem health are recorded. If statistics are computed on each, one by one, the interrelations among variables won't be seen. Valuable information is lost. Multivariate methods resolve what at first may look like noise into recognizable patterns, providing new insight into the field of study. However, these methods are daunting to many scientists, with acronyms like PCA, CCA or ANOSIM, and with unfamiliar terms like varimax rotation and detrended correspondence. The learning curve seems steep, and useful procedures go unused.

UMR covers the multivariate methods of primary interest to environmental science, focusing on what each method is designed to do, when to use them, and when not to. Methods for simplifying data are contrasted with those for establishing connections between and among explanatory and response variables. Capabilities of various software packages are reviewed. Example data sets are analyzed by each student in class. By the end, the choice of which method to use, and how to use it, simply makes sense.

Methods illustrated:
Graphs for Multivariate Data
Principal Component Analysis
Factor Analysis
Correspondence Analysis (several types)
Discriminant Analysis
Cluster Analysis
Canonical Correlation
Nonmetric Multidimensional Scaling
Nonparametric similarity between sets of ecological and chemical measures

'Til next time,

Practical Stats
http://www.practicalstats.com

-- Make sense of your data