Practical Stats Newsletter for April 2008

Subscribe and Unsubscribe: http://practicalstats.com/news

In this newsletter:
1. Upcoming Courses
2. Sample Size for Nonparametric Tests

1. Upcoming Courses

We've held off releasing this newsletter until we had firm dates for our Fall courses. We now think we do. But first, there's only a short time to sign up for our *Applied Environmental Statistics* course, coming up next month in Bordentown, NJ. It's the week before the National Water Quality Monitoring Conference in Atlantic City. Registration is online at http://www.practicalstats.com/new_classes/classes.html.

Next up is a one-day *Introduction to Practical Statistics* on Sept. 24, 2008, just before the California Groundwater Resources Association annual meeting in Costa Mesa CA. It will introduce some of the concepts in our *Applied Environmental Statistics* course, looking at them from the perspective of groundwater quality. For more information, see http://www.grac.org/stats.asp . If you know a ground water scientist who has been alienated by statistics, this workshop will begin the reconciliation.

Untangling Multivariate Relationships 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. More detail on course content is on our website. It will be held in Tampa, FL on Nov. 11-12, 2008. Registration will be online by early May.

Immediately following on Nov. 13-14 at the same location is *Nondetects And Data Analysis,* the course that illustrates methods for correctly handling data with nondetects. Registration will again be online by early May. Then you'll be able to sign up for one or both of these November courses at
http://www.practicalstats.com/new_classes/classes.html.
The following week in Tampa is the SETAC (Society for Environmental Toxicology and Chemistry) annual meeting.

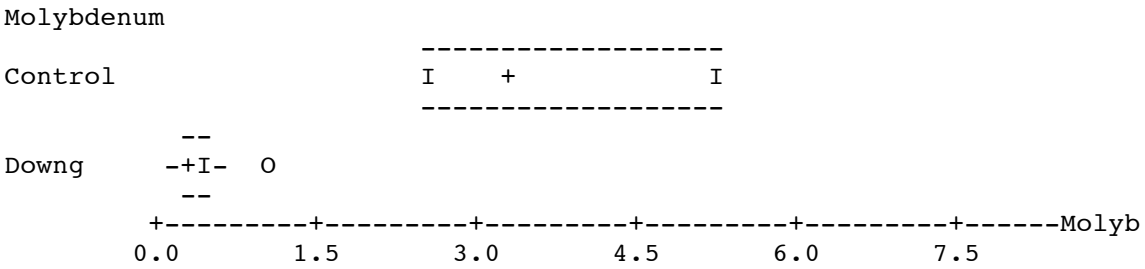
Let others know of these dates and courses.

2. Sample Size for Nonparametric Tests

Statistical software usually provides routines for determining sample sizes necessary for parametric tests such as the t-test or ANOVA. The equations for determining 'how many samples' require estimates of the standard deviation, the desired alpha level of the test, and the desired level of power. Power is the probability of rejecting the null hypothesis (finding differences) when groups are actually different. Also needed is the minimum difference between means you want to be able to detect. In addition to commercial software, you can also find parametric sample size calculators online.

Few people realize that there are also equations for determining sample size requirements for nonparametric tests such as the Mann-Whitney rank-sum test. Standard software does not generally include them. In our *Applied Environmental Statistics* course, however, we provide macros to compute them and discuss their relevance to the skewed data common in environmental sciences.

For example, the AES class uses a data set of molybdenum concentrations in ground water. A line-printer plot is given below, to avoid attaching a graph file that might be interpreted as spam (old technology does still have its uses). The control site consists of a set of only 3 observations. Another set of 13 values were measured downgradient (Downg) of a source of low-Mo water. The data are skewed. All 3 of the control site concentrations are greater than the downgradient concentrations. The variability within the two groups is radically different. The two-sample t-test finds no difference in the concentrations. The Mann-Whitney test does.



From the sample size calculator using the difference in means observed in this data, and an estimate of standard deviation from the most variable group, to achieve 90 percent power would require 13 observations in each group. This assumes that both groups have a normal distribution and the same variance, which they do not. We had only 3 observations in the control group, so our power was quite a bit less than 90 percent. It is no wonder that the two means were not declared different with the t-test. It has low power.

Power for the nonparametric test is computed differently. Here the determinant of difference is how frequently one group is greater than the other. In our data, all of the data in one group are larger than the other, a large signal. The calculations allow for different observed sample sizes rather than requiring that sample sizes be equal in each group. The achieved power is 85 percent. So it is not surprising that the Mann-Whitney test found differences in the median concentrations. For the equivalent tests (the parametric test for differences in means, the nonparametric test for differences in medians), with the same sample size and data the nonparametric test had considerably more power than the t-test.

Power calculations can be a guide to how many samples are required to achieve the possibility of seeing a signal if it is present. If we need 20 samples in each group but have only 6, it is unlikely we will find differences even though they may be present. Parametric sample size calculators have the restrictions of requiring normal distributions

and equal numbers of observations for each group of data. Nonparametric calculators can avoid these assumptions, better fitting the realities of environmental studies. Our AES course has as a main goal to discuss and use methods that fit reality -- that are practical.

Nonparametric power and sample size calculations are discussed in Noether (1987), “Sample size determination for some common nonparametric tests", Journ. of the American Statistical Assoc., v. 82, no. 398, p. 645-647. In AES we translate that article and turn it into working software.

'Til next time,

Practical Stats
-- Make sense of your data