Practical Stats Newsletter for July 2019

Subscribe and unsubscribe: http://practicalstats.com/news
Archive of past newsletters
http://practicalstats.com/news/archive.html

In this newsletter:
A. Practical Stats Courses and Webinars
B. How Multiple Comparison Tests Work
C. Webinars vs Newsletters

A. Practical Stats Courses and Webinars
On our online training site: http://practicalstats.teachable.com/

Our self-paced Applied Environmental Statistics course is available in two parts, each $650 USD for a 1-year access for one person. Or get both courses together (equivalent to our week-long in-person course) in a bundle for $1200 USD. See our online training site at the link above.

Our Nondetects And Data Analysis course is available online. One year's access to the materials costs $695. It covers regression, group testing, and most anything you'd want to do with data that have nondetects, all without substituting fabricated numbers like ½ the detection limit. The R scripts included provide new functionality to make data analysis easier (see our March 2019 newsletter for an example – multiple comparison methods for data with nondetects), and are a step forward from the NADA package in R.

Our next webinar is
"Which of These Things is Not Like the Others? How Multiple Comparison Tests Work"
on July 16, 2019   1:00 PM Eastern, 10 am Pacific time.
Register at:
https://attendee.gotowebinar.com/register/2808021155706915853

B. How Multiple Comparison Tests Work
This newsletter and the accompanying webinar present one of the many new topics in the 2nd edition of *Statistical Methods in Water Resources*, soon to be released by the US Geological Survey in its Techniques and Methods series of publications. When it is released you can download the book at no charge from http://practicalstats.com/info2use/books.html or from the USGS Publications website.

Analysis of Variance (ANOVA) and Kruskal-Wallis (K-W) tests determine whether any difference occurs between group means (ANOVA) or sets of percentiles (cdfs, tested by nonparametric tests like K-W). Multiple comparison tests (MCTs) follow these general tests to provide information on which groups differ from others. The modern versions of these tests set something like an overall error rate, the probability of making at least one error in the pattern of groups. For six groups, each represented by a letter, the resulting pattern might look like
A > B = C = D = E > F

with each letter standing for a group mean (ANOVA) or cdf (K-W). The pattern is determined by computing a series of two-group tests, t-tests to follow ANOVA and rank-sum tests to follow K-W. The overall error rate is set to the alpha level for the preceding general test, say 0.05, which requires that the individual two-group tests must have an alpha level smaller than that of the overall rate. The smaller the acceptable error rate alpha the harder it is to see differences between groups. The method to define the two-group (called individual or pairwise) error rate while keeping the overall rate constant defines different MCTs. If an MCT author finds a way to provide a higher pairwise error rate while keeping the overall error rate the same, this provides more power to see differences between the groups.

One of the first MCT pairwise error rate adjustments was by Bonferroni, who simply divided the overall error rate by the number of two-group comparisons being made. For k groups, there are $c = k(k-1)/2$ pairwise comparisons, so Bonferroni's adjustment divided the overall error rate by c. For k=6 groups there are c=15 pairwise comparisons, so each two-group test would compare its p-value to $0.05/15 = 0.0033$. If the p-value for the test falls below 0.0033 the two groups' means or cdfs are found to differ, and the resulting pattern has an overall error rate of 0.05. Bonferroni's adjustment does not have a great deal of power to see differences, because the adjustment produces a pairwise error rate that is quite low, making it hard to see differences.

Later MCTs tests, including Dunn's and Holm's tests, keep the overall error rate at its desired level while increasing the pairwise error rate for the two-group comparisons. These find more group differences, so fewer equal signs in the pattern schematic, because of their increased power. However, recent research in genomics requires hundreds to thousands of two-group tests to be made, making it unlikely that any of these methods would find group differences. Instead, an alternative objective was proposed by Benjamini and Hochberg (1995), the False Discovery Rate (FDR). The FDR sets error rates for 'false positives' only, a false declaration of a difference between two groups. This is a subset of the cases that an overall error rate addresses. However, it allows pairwise error rates to stay higher, closer to the overall rate, by redefining the rate of interest to be the FDR. In doing so, more differences are found, and so fewer no-differences declared. This minimizes the impact of not addressing the 'false negative' rate, because there are fewer no-differences to worry about. For many disciplines, including environmental sciences, using the FDR makes a lot of sense. It minimizes the problem area of falsely finding differences, while increasing the ability to see real differences between groups. It is the MCT recommended for use in the new edition of *Statistical Methods in Water Resources*.

Details on how the Benjamini-Hochberg (BH) adjustment works are given in our webinar "Which of These Things is Not Like the Others?". If you missed the live broadcast, you may stream the recording from our Training Site. The bottom line is that for iron concentrations at low flow defined by six land-use groups, the Bonferroni MCT found only four significant differences out of the 15 two-group tests for difference, while the BH MCT found 9 of the 15 to differ while also controlling for no more than 5% false differences. The R code and boxplots illustrating the differences are below. The webinar also tells you why, even though the two-group tests are run using alpha levels below the overall rate, the output from the tests provide p-values that the user compares to their overall error rate, say 0.05.

```
> kruskal.test(fe~group)
        Kruskal-Wallis rank sum test
data:  fe by group
Kruskal-Wallis chi-squared = 33.781, df = 5, p-value = 2.633e-06
```

```
> pairwise.wilcox.test (fe, group, p.adjust.method = "BH")
      Pairwise comparisons using Wilcoxon rank sum test
data:  fe and group
         lime_ab lime_re  lime_un  sand_ab sand_re
lime_re 0.01031 -        -        -       -
lime_un 0.00052 0.04262  -        -       -
sand_ab 1.00000 0.01635  0.00090  -       -          9 of 15 tests significant at alpha 0.05
sand_re 0.69026 0.02050  0.00052  0.45843 -
sand_un 0.02779 0.33746  0.00601* 0.07417 0.05666
```
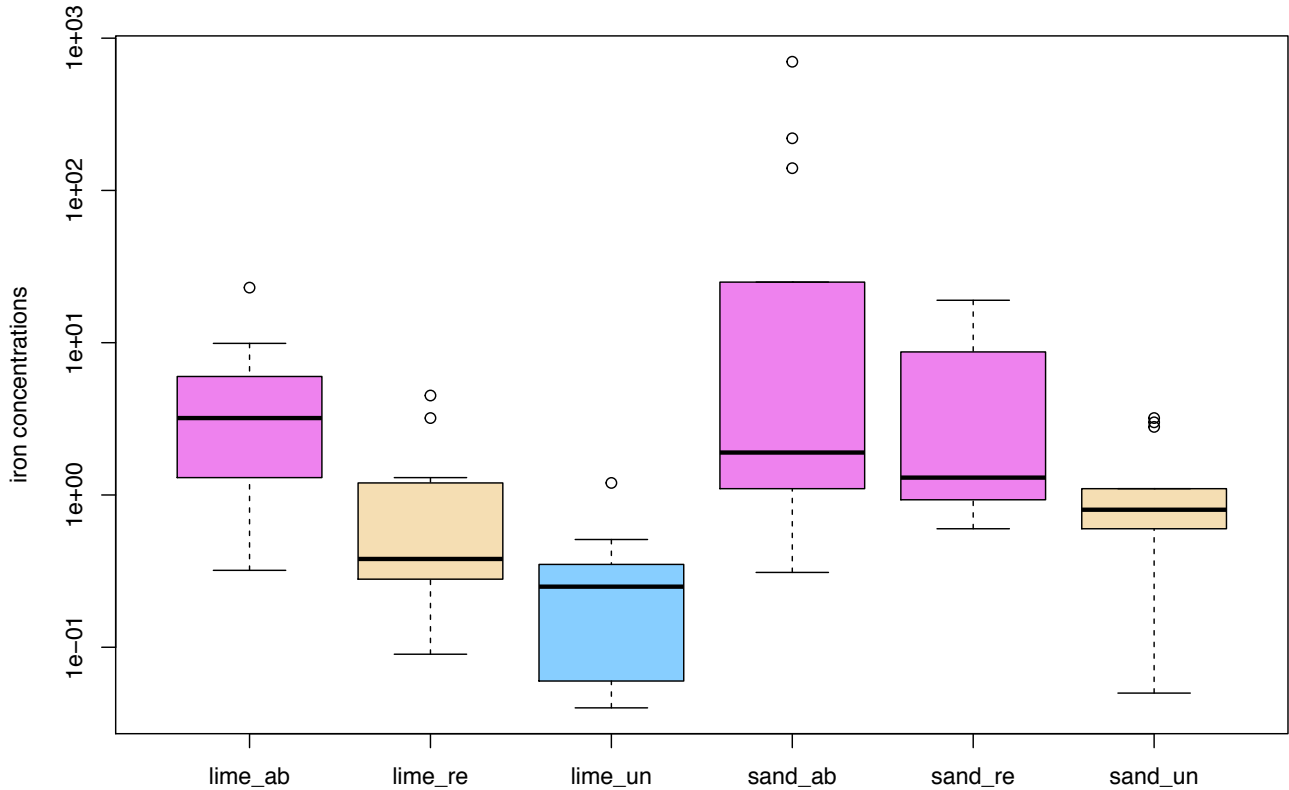

Figure 1. Boxplots of iron concentrations classed into 6 land-use and geology groups. Boxes with different colors were found to be significantly different by the BH multiple comparison test.

C. Webinars vs Newsletters

I've been doing monthly webinars this year instead of the every-two-months newsletters that we've released for the previous 12+ years. Webinars allow me to include much more information than the newsletters do. I do want to remind you that our newsletter archive is available at
http://practicalstats.com/news/archive.html
and contains all of the newsletters we've done in the past, except those whose information became out of date. If you haven't attended all of this year's webinars, the recordings are at our Training Site, http://practicalstats.teachable.com . Click on the "Show more courses" button to see more than the top three listings. The webinars have come from material in our two online courses, Applied Environmental Statistics and Nondetects And Data Analysis, which are also located there. Both the courses and

webinars can be streamed in your timezone, at any time of day you want to view them, on a computer or tablet (in theory, also on your smartphone). Let others that may be interested know about the free webinars, and the courses.

'Til next time,

Dennis Helsel
Practical Stats LLC
  -- Make sense of your data