Chapter5 - Multidimensional Semiquantitative Data

Just to review from before, there are 3 categories of data that depend on their relative precision:

- Quantitative data are the most precise. They can be plotted on continuous axes. Examples include counts, growth, density etc.
- Semi-quantitative data are less precise and are generally ordinal ranked categories. An example would be substrate classes (e.g., fine sand to boulder).
- Qualitative are the least precise and are non-ordered categories. For example, color (although Joey suggested purple is the worst so should be ranked last).

This chapter focuses on semi-quantitative data and non-parametric statistical techniques to analyze them. Non-parametric tests are distribution free (i.e., they don't assume any underlying normal distribution) and many discussed in this chapter are ranking tests.

Non-parametric tests should be used if the any of following apply:

- if one or more descriptor is semi-quantitative
- if the purpose is to evidence monotonic rather than linear relationships
- one or more descriptor is not normally distributed
- the number of observations is small.

Monotonic refers to a function where the slope is always negative or positive. Consider a correlation between two descriptors A and B. If A is a set of integers 1-5 and B is the series (1,5,10,25,50), this function will not be linear. However, if the ranks are used (as in a non-parametric correlation test), series B becomes 1-5 and there is a perfect correlation.

5.1 Summarizes and compares methods for analyzing the 3 data types *Table 5.1 is a good diagnostic for what tests to use in what situations.*

Key take home points from this section include:

- Don't be afraid to collect data with varying precision (there's tests for everything!)
- These categories are not rigid. For instance, you can treat quantitative data as semi-quantitative data in non-parametric tests OR you can treat semi-quantitative or qualitative data as quantitative as dummy variables in a regression. So you are never stuck with a certain type of analysis (in most cases..)

5.2 One dimensional non-parametric methods There are two main families of tests. Those for independent samples and those for related samples. Independent samples test for differences in means or medians between two groups. Related samples have observations that are linked in some way, e.g., sites measured at different times, different lakes measured at several depths etc.

A brief overview of some of the semi-quantitative methods

If k (# of groups) =2 independent samples: For quantitative data this would be a t-test. The semi-quantitative equivalent is the Mann-Whitney. The M-W test takes data from both groups, ranks them into a single series, then tests the hypothesis that the MEDIANS are the same

K $>$ 2 independent samples:

One-way ANOVA for quantitative data. Semi-quantitative version is the Kruskall-Wallis test. It?s the same idea as a Mann-Whitney except applied to multiple groups. K-W tests the null hypothesis that the sum of the ranks in each group are the same ? or that they are drawn from identical populations.

K =2 for related samples: In quantitative land this is a paired t-test. Two versions for semi-quantitative data: The sign test: Looks at the difference between each set of paired observations then codes it as a + or -. Tests the null hypothesis that the number of pairs with each sign are equal. Wilcoxon matched pairs test: More powerful than the sign test. Can be used if the magnitude of the difference between the matched pairs is known (e.g., if its quantitative but non-normal). Same idea as the sign test except the differences themselves are ranked and the null hypothesis that the sum of ranks are equal for each group.

K > 2 for related samples: Quantitative would be a two-way ANOVA Semi-quantitative is Friedmans* two-way ANOVA by ranks For a matrix of related descriptors (e.g., Chl-a measured in different lakes at 4 depths), each column is ranked, then the ranks values in each row are added together. Null hypothesis is that the rank totals of rows are equal.

- Fun trivia: Milton Friedman, a famous economist who many view as the father of modern conservative politics, developed this test. I plan on boycotting it.