**Analyzing Nutrient Simulation Experiments with a one-way ANOVA**

**Author:** Rebecca L North, University of Missouri-Columbia, USA

**Course Information**

Department: School of Natural Resources

Level: **Upper Undergraduate**

Course type: **Both**

Students: **Majors**

Number of Students: 10

**Focus:** In addition to learning programming/how to use R, students will determine if there are significant differences between experimental treatments in nutrient simulation experiments that they designed and carried out.

**Overview:** In this lesson, we will provide an overview of parametric and non-parametric statistics and when and where they should be applied. We will discuss what tests are appropriate to analyze the data from the students’ nutrient simulation experiments (NSEs). We will have a hands-on tutorial in R and introduce the students to swirl using the lesson “RprogrammingE”. As a group, we will complete the lesson “NorthNSEANOVA” using a dummy dataset. Then the students will individually work through the 1-way ANOVA using their own datasets.

**Learning objectives:** By the end of this lab, students should be able to understand how and why they are applying 1-way ANOVAs to analyze their individual nutrient addition experiments. They will be comfortable loading datasets, executing commands in R and interpreting and reporting statistical outputs/results. They should each have an F value and p values associated with their experiments from the 1-way ANOVA and post hoc Tukey tests. They should know if their experimental treatments resulted in significantly different chlorophyll a concentrations.

**Lesson sequence:** Provide a numbered, ordered list of the activities within your swirl lesson. This list can be taken from step 4 in your initial lesson design, with any modifications that were introduced.

1. Introduction to ANOVA (ANalaysis Of VAriance) including identification of continuous and categorical variables.
2. Loading, plotting, and analyzing a Dummy dataset.
3. Implementing a one-way ANOVA using the aov function.
4. Interpreting p values and deciding whether to accept or reject null hypotheses.
5. Overview and testing of ANOVA assumptions with the Bartlett test of homogeneity and the Shapiro-Wilk test to examine normality of residuals.
6. Overview of data transformations.
7. Interpreting an ANOVA table and applying a post-hoc test (TukeyHSD) to examine pairwise differences in the continuous variable.
8. Interpretation and reporting of statistical results.
9. In class loading of individual datasets and repeat of the swirl lesson.

**Pre-lesson activities:** Students conducted a nutrient simulation experiment by collecting cubitainers of water from a local lake/reservoir. Three cubitainers had nothing added to them and served as controls, in triplicate the students amended the remaining cubitainers with nitrogen (N), phosphorus (P) or both (N+P). Experimental cubitainers were incubated for 9 days and then filtered for chlorophyll a (chla) concentrations (a proxy for phytoplankton biomass). The highest chla concentrations indicated which nutrient was limiting to phytoplankton growth. In the 4 hour lab period, we provide an overview lecture on statistics and introduce them to RStudio and swirl using RprogrammingE before implementing this swirl lesson.

**Post-lesson activities:** After the swirl lesson, students load their own experimental data (in place of the Dummydata) and repeat the lesson to analyze their own experimental results. Final lab report should include a figure of the experimental results, and results of the one-way ANOVA and Tukey tests (F statistic, and p values using proper nomenclature). They use these results to assess what the limiting nutrient was in their study lake/reservoir.

**Implementation notes:** It takes much longer for students to work through this lesson than I initially thought. I took a 4 hour lab period to provide a background lecture (~30 mins), work through an initial introductory swirl lesson (~50 mins), this lesson (~70 mins), and then having them repeat it with their own data (~60 mins).