Cornell University
Cornell Statistical Consulting Unit

StatNews #75

Diagnosing and Avoiding Pseudoreplication
March 2009

What is wrong with the following experimental design?

A group of elementary school teachers have developed a new, and they believe improved, 3rd grade curriculum. They wish to test the relative benefit of the new curriculum compared to the conventional one, hoping to demonstrate that it should be implemented in all the schools. For the experiment, the new curriculum will be adopted at one elementary school, and mean test scores of students from that school will be compared to mean test scores of students from a control school. The experimenters will randomly sample 20 third grade students from each school, for a total sample size of 40.

In the above design since the new method of instruction is implemented at the school level, the experimental unit is a school. The schools, **not** the students, are randomly assigned to either the new method of instruction or the control (conventional) method of instruction. The population of interest is all the schools not just these two schools.

The mistake that the experimenters have made is that they have not replicated their treatments correctly. They actually have false "replicates" or subsamples. Since the treatment is applied to the schools, the students are not replicates of the experimental unit, but subsamples of the schools. The test scores of the students within a school are not independent from each other. This error is known as **pseudoreplication**.

In an experiment, replicates, also known as experimental units, are defined as independent applications of the same treatment. Students at one school are not independent units; they are samples from a single site. In this case, experimental units are the different *schools* rather than different students. An example of a correct design for this experiment would be to assign the new curriculum to 10 schools and compare the performance of students from these schools to the performance of students from 10 control schools. In this case, the total sample size is 20 because the experiment has been replicated across 20 experimental units (schools).

Consider another example:

Suppose we are setting up an experiment to measure growth of plants. If the treatment of interest is nitrogen fertilizer applied to subplots of 10 plants, then an individual plant is a pseudoreplicate or subsample. On the other hand, if the nitrogen fertilizer treatment can be randomly assigned to each individual plant, then each individual plant is an experimental unit or replicate.

The purpose of replication in an experiment is to control for unaccounted variability and to increase the precision and confidence of the experimental findings. Treating subsamples, or pseudo replicates, as experimental unit replicates, may result in incorrect statistical inference and needlessly increase experimental costs.

For example, in an experiment that seeks to test a new variety of corn, growing one crop of the new variety does not give much information. Perhaps that specific crop had a higher yield for the new variety than a conventional corn, but under different conditions the outcome could have produced different results. The experimenters would be able to draw valid conclusions only if the experiment included several crops planted under different circumstances, so that the effects of random variability in environmental and other factors would not bias the results. To quote Ronald Fisher: "No one would now dream of testing the response to a treatment by comparing two plots, one treated and the other untreated."

Pseudoreplication can be avoided in experiments if the units of analysis (e.g. student in a class, plants) and the units of replication (or the units receiving each treatment, e.g. schools, groups of plants) are clearly defined to ensure that replicates are independent.

If you need help with your experimental design or have statistical questions, please do not hesitate to contact a statistical consultant at the Cornell Statistical Consulting Unit.

References:

1. Montgomery D., Design and Analysis of Experiments, Wiley, 2005.

2. Casella G., Statistical Design, Springer, 2008.

Author: Andrea Knezevic