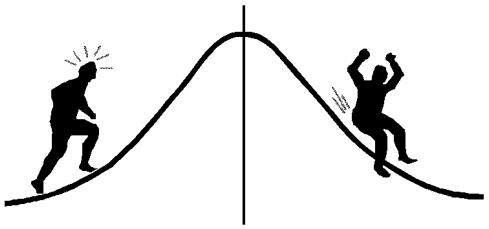

# BIOMETRICS INFORMATION

(You're 95% likely to need this information)

PAMPHLET NO. # 2 DATE: February 12, 1988

SUBJECT: The importance of replication in Analysis of Variance

"Replication assures us the means of computing experimental error" (Steel and Torrie pg. 135).

A specific treatment is said to be replicated if it is applied to more than one experimental unit within an experiment. There are many ways this can be accomplished by the experimental design, but regardless of how it is done, it is ESSENTIAL. Without some estimate of experimental error there is no "objective" way to decide if differences between means (or rather values in an unreplicated experiment) are due to the treatments or are due to inherent differences between the experimental units before any treatments were applied.

Consider the following simple one-way analysis of variance. Six different fertilizer treatments are applied to groups of five seedlings:

| FERTILIZER: | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | 1 | 6 | 11 | 16 | 21 | 26 |
| | 2 | 7 | 12 | 17 | 22 | 27 |
| SEEDLINGS: | 3 | 8 | 13 | 18 | 23 | 28 |
| | 4 | 9 | 14 | 19 | 24 | 29 |
| | 5 | 10 | 15 | 20 | 25 | 30 |

In this design, seedlings are said to be nested within treatment because not all seedlings appear for each treatment (The seedlings were numbered from 1 to 30 to indicate that each is a unique individual). To test for the significance of fertilizer, an F-test is performed, using the ratio of the mean-square for FERTILIZER to the mean-square for SEEDLINGS. The mean-square for FERTILIZER has 6-1=5 degrees of freedom. The mean-square for SEEDLINGS has 6x(5-1)=24 degrees of freedom. Everything is straight forward, and the analysis goes without a hitch.

This simple experiment replicated each treatment over five seedlings. What would have happened without replication? Why not apply each treatment to only one seedling?

If this experiment had been conducted without replication, it would be impossible to perform an F-test because the mean-square for SEEDLINGS would have 6x(1-1)=0 degrees of freedom! With 0 degrees of freedom, no estimate of experimental error would be available, and analysis of variance could not be carried out.

With more complex designs, it is sometimes possible to use high-order interaction terms to estimate error in an unreplicated experiment, but it is always wise to replicate if at all possible.

Another important reason to include replication in an experiment is to increase the scope of inference by the selection and use of more variable experimental units. For instance, suppose you want to know if one stock type will grow faster than another. You also know that climate affects growth rate. Obviously, if you set up trials of the two stock types on plots of land with a range of different climates, the conclusions of the trial will be fairly general. If, on the other hand, you had only used plots from a dry climate, it would be difficult to justify extending those conclusions to plots in wet climates.

CONTACT: John Thornton
387-3021