Interpreting Experiment Results

Donald M. Marshal
Department of Animal and Range Sciences

A typical experimental format involves evaluating the response caused by application of alternative treatments to experimental subjects (animals, carcasses, pens, pastures, etc.). The effect of a given treatment might be evaluated by comparison to a control group or to one or more other treatment groups. However, a problem with animal research (and other types as well) is that variation not due to treatments often exists among experimental subjects. Statistical procedures can be useful to determine the extent to which observed variation is due to treatment effects versus other factors.

For example, suppose that animals receiving Diet A grow faster than animals receiving Diet B. Was the observed difference in growth rates actually due to dietary differences or to other factors (e.g., genetics, age, sex, measurement error, etc.) or some of each? Statistical analyses evaluate the amount of variation between treatment groups relative to the amount of variation within treatment groups. In addition, variation caused by factors other than treatments can sometimes be eliminated by statistical analyses. A brief discussion of some of the more common statistics encountered in animal research follow.

Average or Mean. These two terms are used interchangeably. We often compare mean values of treatment groups for variables of interest. In some studies, least-squares means are reported rather than the raw means. In so-called "balanced" studies, least-squares means are often the same as raw means. However, when experimental subjects are distributed across treatment groups in an uneven or biased manner, then adjustments to the means are needed to account for the bias. Appropriate adjustments are made by the procedure of least squares.

Correlation Coefficient. This statistic is a measure of the degree of association between two variables and can range from -1 to +1. A strong positive correlation (close to +1) indicates that high values of one variable tend to occur more often than not in combination with high values of the other variable. Similarly, low values of one variable tend to be associated with low values of the other variable. In humans, for example, we generally expect a rather strong, positive correlation between height and weight. Taller individuals tend to be heavier, whereas shorter individuals tend to weigh less, on average. A strong negative correlation (near -1) indicates that high values of one trait tend to be associated with low values of the other trait. A correlation coefficient near zero indicates that the two variables are largely independent of one another.

Regression Coefficient. This statistic indicates the average change in variable $Y$ for each one unit increase in variable $X$. In its simplest form (i.e., linear regression), the regression coefficient is simply the slope of a straight line. A regression equation can be used to predict the value of the dependent variable ($Y$) for a given value of the independent variable ($X$). A more complicated procedure, known as multiple regression, can be used to derive an equation which uses several independent variables to predict a single dependent variable. An example is the USDA beef cutability equation, in which % cutability is predicted from carcass weight, external fat thickness, KPH fat, and rib-eye area.

---

1Professor.
Variance. This is a measure of variation of a variable (trait). Its unit is the square of the unit of measurement (e.g., lb²).

Standard Deviation. This is also a measure of variation calculated as the square root of the variance. Thus, its units are the same as the original trait.

Coefficient of Variation (C.V.). The C.V. is calculated as the standard deviation divided by the mean for a particular variable or trait. Dividing by the mean removes the effects of scale and units from the variable, which allows a comparison of the relative variation between two traits. The variance or standard deviation of different traits cannot be directly compared, but it might be appropriate to compare their C.V.'s.

Standard Error. Data presented in an experiment are normally based on a sample of experimental subjects drawn from some larger population. Hence, a statistic (parameter) calculated from the sample group is only an estimate of that parameter's value in the entire population. A value known as a standard error is often calculated for parameter estimates such as the mean, correlation, or regression coefficient. The standard error is an indication of the possible error associated with such estimates. It is calculated as a ± value (deviation).

The magnitude of the standard error depends on the animal-to-animal variation and on the number of animals in the sample from which the parameter was estimated. As sample size increases, a larger proportion of the whole population is included, and the likelihood is increased that the parameter estimated from the sample will closely approximate the overall population parameter. The standard error decreases as sample size increases.

Probability Value or Statistical Significance (P-value). Statistical comparisons will often be accompanied by a probability (P) value. Suppose, for example, a research paper indicated "calves receiving Diet A gained .35 lb per day more (P = .05), on average, than calves receiving Diet B." For practical purposes, we can interpret this statement to mean that the probability of attaining a difference of at least .35 lb/day for reasons other than dietary effect is about 5%. Such a difference may be said to be statistically significant at the .05 level of probability.

A difference larger than .35 lb/day in the example above would have resulted in a smaller P-value. A smaller P-value reflects increased confidence that there is a true underlying effect of the treatment. When differences between treatment means are relatively small—compared to differences between animals receiving the same treatment—then the P-value will be higher and we cannot confidently conclude that there was a true treatment effect. The size of difference required to achieve a given P-value varies between traits and studies. All other factors being equal, as sample size increases, a smaller treatment difference is required to achieve a given level of statistical significance.