Data Analysis of Pre-Post Study Designs

Statnews #79

Cornell Statistical Consulting Unit

Last updated October 2020

Introduction

Pretest-posttest study designs are widely used across a range of scientific disciplines, principally for comparing groups and/or measuring change resulting from experimental treatments. The definitive characteristic of this study design is that (at least) two measurements are made on the same experimental unit: the pretest measurement made prior to the administration of a treatment or intervention and the posttest measurement made at a later time point. Pretest-posttest studies have been pervasive for many years, however many researchers are still unclear on the statistical methods most appropriate for analyzing such data.

Example

A typical example of a pretest-posttest design is be the following. A researcher hypothesizes that literacy skills can be improved by a new reading curriculum. Prior to the delivery of the curriculum, proctors administer a standardized literacy assessment at the start of the school year to 30 students in a school district. Of the 30 students, 15 are then randomly chosen to receive the new curriculum, while the 15 other students receive their district's current curriculum. At the end of the school year, all 30 students take the same standardized literacy assessment again.

There are several common approaches employed to analyze such data, four of which are presented below.

ANOVA on change scores

The change score can, i.e. the posttest score minus the pretest score, can be used as the dependent variable in an ANOVA that compares two or more groups. This is a tempting choice in that it reduces the problem from a multivariate one (two or more measurements) to a univariate one, with this new variable easily interpreted as net gain or loss. Change scores are often criticized as being less reliable than the raw scores, but a more serious problem with this approach is that change scores can be misleading when regression toward the mean is present. Regression toward the mean is the tendency of individuals with extreme prettest scores to have posttest scores that are closer to the average posttest score. This is equivalent to stating that the prettest score is negatively correlated with the change score. If an independent variable, such as the binary variable indicating treatment group, is correlated with the pretest score, then a spurious negative correlation can occur between the independent variable and the change score.

Repeated measures ANOVA

Using the data as a mixed factorial design with one between-subjects factor (treatment group) and one within-subjects factor (pretest-posttest) is also a common approach. As explained in a previous newsletter (http://cscu.cornell.edu/news/archive.php), such data can be analyzed using two different statistical methods. However, care must be taken when analyzing prettest-posttest data with this method, since it is the *interaction* between the treatment factor and pretest-posttest factor that describes the difference between treatment groups in their change over time. This interaction term is equivalent to the treatment main effect within a one-way ANOVA on change scores. Hence this method is subject to all of the previous deficiencies of the ANOVA method.

Linear model for the posttest score

A linear model with the posttest score or change score as the outcome variable and pretest scores included as a covariate is generally the preferred method for analyzing pre-post design data as it eliminates systematic bias and reduces error variance. This method implicitly takes into account regression toward the mean. It has been shown that the outcome variable can be either the raw posttest scores or change scores, as they will yield exactly equivalent results for the treatment effect. Researchers must keep in mind that though this is likely the appropriate analytic choice, the usual assumptions that underlie a linear regression model still apply: randomization, homogeneity of regression slopes, pretest measurement reliably, and a linear relationship between pretest and posttest scores. Modifications to this approach are available if violations of these assumptions are detected (e.g. nonparametric rank-transformations, interaction term between pre-test and treatment).

Linear mixed model

A linear mixed model is another appropriate choice for the modeling of prettest-posttest data. In the typical prettest-posttest design, each individual has two measurements. The linear mixed model uses these two measurements as the outcome variable and includes “time” (before vs. after) as an independent variable. An interaction term between time and the treatment indicator can be used to compare the treatment groups based on the average change in the dependent variable. Correlation between the two measurements on each individual is accounted for by including random effects for each individual.

If you need assistance with a Pretest-Posttest analysis problem, do not hesitate to contact a statistical consultant at the Cornell Statistical Consulting Unit.

References

Bonate, P.L. Analysis of Pretest-Posttest Designs. Chapman & Hall, 2000.

Dimitrov, D.M. & Rumrill, P.D. Pretest-Posttest Designs and Measurement of Change. Work 20, 159-165, IOS Press, 2003.

Bland, J.M. & Altman, D.G. (1994). Statistic Notes: Regression towards the mean. British Medical Journal, 308, (6942): 1499.