On the Importance of Independence…

Donald W. Zimmerman of Carleton University recently published an article that should help motivate data analysts to heed an old rule in using the Student \( t \) test. This basic test for differences assumes that the data are a random sample of independent observations. What happens when we ignore this rule and use the \( t \)-test on data that violate this assumption?

Zimmerman computed the actual Type I error (rejecting the null hypothesis when it is actually true) under various scenarios of seemingly small levels of nonindependence (as indicated by the Pearson correlation between two samples). His results, for example, show that with sample size of \( N=100 \) and with two such samples correlated at a seemingly ignorable 0.02, the actual chance of Type I error is 0.258 rather than the assumed 0.05 level. The article documents how this inflation of the Type I error balloons upon the increase in the sample size, or the correlation between the samples, or both factors. For example, if we had assumed \( N=200 \) and correlation=0.04, then the actual probability of a Type I error would catapult to 0.519, which far exceeds the assumed 0.05 level.

So Zimmerman warns analysts to check the effect size. “If a sampling procedure is not strictly random and a slight undetected correlation between successive observations exists, then rejection of \( H_0 \) can occur with high probability even though the effect size is zero…”

This study has relevance for institutional researchers because they usually have semester and full-term data at their disposal for various statistical analyses. If the number of cases in each data set seems small, some analysts may erroneously pool the datasets of students from successive semesters or academic years and proceed to use the standard \( t \)-test (or the \( F \)-test in the analysis of variance) without adjusting for the loss of independence (notably from the presence of many of the same students in both datasets). Here is one situation where larger samples incur a larger risk of an erroneous conclusion (because of the misleading nominal significance level) than smaller samples. To many analysts the disadvantage of a large sample is counterintuitive, and this only makes it harder to avoid this pitfall.

Interested parties may find Zimmerman’s article, “A Warning About Statistical Significance Tests Performed on Large Samples of Nonindependent Observations,” in the following journal: *Perceptual and Motor Skills*, 2002, 94, 259-263.
