Letter to the Editor

Horgan, G. W. (1999), "Use of Spreadsheets for Demonstrating Experimental Power and Variability," *Journal of Statistics Education*, 7(1).

The second paragraph of Section 2, "Spreadsheet for the Power of the Two-Sample t-Test," requires clarification. Power can be illustrated through simulation, but in a completely different way than is illustrated here. In fact, the simulation in Figure 1 carries out a simulated experiment once. The fact that one "can do another experiment by pressing the F9 key" doesn't help at all to clarify the concept of the power of a test; a "long run" property should be demonstrated.

The p-value in the spreadsheet in Figure 1 (in fact, the only result in boldface) is inappropriate, because p-values relate to the level of significance, whereas the intention of the paper is to demonstrate power. When the "Truth" is as given in Figure 1 (and, in fact, the main point of the simulation is that the reader should play with situations where the control mean differs from the treated mean), then the t-statistic follows a noncentral t-distribution. Excel computes probabilities for the usual (central) t-distribution and produces p-values on the basis of that computation.

The picture with separate confidence intervals for Treated and Control doesn't tell much in a context where comparison is the research question. What does the reader learn from seeing overlapping confidence intervals (apart from the fact that the level of confidence isn't even mentioned)? Can treatment effects then be attributed to chance alone? Here is a numerical example. Take 16 controls with mean 59.2 and 9 treated animals with mean 70.7. Let the pooled standard deviation equal 12. Under the assumptions of the paper, the 95% confidence intervals for the mean effects of the treated and control groups overlap. But a 95% confidence interval for the difference between the two population means does not contain zero!

Herman Callaert
Center for Statistics
Limburgs Universitair Centrum
Universitaire Campus, Building D
B - 3590 Diepenbeek, Belgium

herman.callaert@luc.ac.be

Response

Professor Callaert suggests that my description of an Excel spreadsheet for demonstrating experimental power requires clarification. I am happy to attempt this.

He states that power can be illustrated through simulation in a different way. No doubt this is so; there must be many ways to do it. I have found that my spreadsheet does help to explain a key point in understanding experimental power: that p-values are subject to experimental variability, and that they depend on the truth and the sample size. The calculation of the power is then done otherwise, by tables or PC packages such as Minitab-12, and this is illustrated in class. The spreadsheet is not intended to be used for such calculations, not even approximately.

The picture used, with separate confidence intervals for Treated and Control, mimics what is normally done in the scientific literature that my audience will read and publish in. (This may be because there are often more than

two groups to compare.) It is true that these intervals may overlap, and yet the difference be significant at 5%. It is something I sometimes draw their attention to; many are not aware of this possibility. I agree that the level of confidence (which is 95%) should be stated, and I have amended the spreadsheet accordingly.

Graham W. Horgan
Biomathematics and Statistics Scotland
Rowett Research Institute, Aberdeen
AB21 9SB, Scotland, UK

G.Horgan@bioss.sari.ac.uk

The revised spreadsheet is available at http://www.amstat.org/publications/jse/secure/v7n1/esim2.xls

---

JSE Homepage | Subscription Information | Current Issue | JSE Archive (1993-1998) | Data Archive | Index | Search JSE | JSE Information Service | Editorial Board | Information for Authors | Contact JSE | ASA Publications