Using Simulation to Explore the Power of Hypothesis Tests

Bessie Kirkwood
Sweet Briar College
Sweet Briar, VA 24595
bkirk@sbc.edu

Partial support for this work was provided by the National Science Foundation's Course, Curriculum, and Laboratory Improvement program under grant 0410586.

The exercise outlined below is one I used in a new Biostatistics course at Sweet Briar College this past fall. Its purpose is to use simulations to demonstrate how the probability of a Type II error depends on sample size, the population mean, and the population standard deviation. Students work in a classroom equipped with desktop computers. In the class meeting after this activity, I developed some sample size formulas. Students commented that they were glad they had done the lab exploration first.

The Biostatistics course is designed to be a "second" statistics course. For more information about the course, and for the full lab activity, please visit the website:

http://www.faculty.sbc.edu/bkirk/Biostatistics/

Although students can compute the power of a test of hypotheses about the mean without resorting to simulations, they can get so involved in the calculations that they lose sight of the main ideas. Simulations reinforce the idea that hypothesis tests result in errors randomly, but with known probability.

Outline of the activity:

The class is divided into teams of two or three students. Each team carries out 5 simulations using MINITAB commands provided in the handout. They sample from a normal distribution with a fixed standard deviation, using a different specified mean for each simulation, and using a sample size assigned to the team. The simulations count the number of times that a one-sided hypothesis test results in a Type II error, thus obtaining approximate probabilities of a Type II error. Teams report their results to the class and everybody creates plots of the results. An example plot is shown below.

In a second set of simulations, mean and sample size are fixed while the standard deviation is varied.

In a third set of simulations, students explore the relationship between the significance level and the probability of Type II error.