Using R as a Wrapper in Simulation Studies

Tim F. Liao
University of Illinois

R is the computing software of choice today for many statisticians across a variety of disciplines, and is becoming the lingua franca of statistical computing because of its flexibility and availability. However, there are many statistical software packages that have been in use for the last two to three decades, and for many data analysts it will be useful if features of an existing program can be employed while the flexibility of R is taken advantage of. The inflexibility of many existing statistical packages makes a simulation study rather difficult, notably when a program, command, or subroutine must be called repeatedly a large number of times with a set of changing input values and with the output collected in an easily executable file.

In this paper I present a simulation application where the software LEM, a popular software for categorical data in the social and behavioral sciences, is called repeatedly from an R wrapper. The application is a simulation within a simulation. That is, R handles the external simulation of a large number times, allowing for varying input parameters. The specialty software handles the internal simulation of estimating a statistical model on randomly generated data. The wrapper allows the data analyst to change the input, store the output, and produce graphic interpretation of the simulation output. Although the example involves LEM in DOS mode, the same idea (not the code) can be generalized to other statistical programs in other platforms (with necessary modifications).