STATISTICAL SIMULATION

Daniel McFadden, Director
Econometrics Laboratory
University of California
Berkeley CA 94720-3880 USA
mcfadden@econ.berkeley.edu
http://econ.berkeley.edu/~mcfadden

Applied science uses simulation models in the laboratory to describe, visualize, and predict phenomena prior to their realization in the field. Contemporary examples are virtual reality trainers for pilots, computer models for weather or for traffic flows through a transportation network, economic forecasting models for product choices in supermarkets, and agent-based evolutionary models for financial markets. These examples all involve computer simulation, which will also be the focus of this paper. However, simulation is both older and more general than computer modeling. Leonardo Di Vinci's working models of gears and irrigation systems are simulation models, rats are used to simulate human biology, and an architect's elevations are used to simulate the appearance of a new building. It may be fruitful to think of these extended forms of simulation in the same manner as the computer simulation models discussed here.

In statistical terminology, simulation is an artificial data generation process, driven by model design and parameter settings, whose output is a synthetic sample. A simulation model is useful if it can be designed and calibrated so that in terms of relevant criteria the synthetic samples it produces approximate well the output of the real data generation process. In this language, the design and calibration process is the statistical problem of model specification and estimation, with success judged in terms of measures of congruence of real and simulated samples from models trained on historical data. For example, in transportation networks, effective capacities and travel times on links are adjusted to match historical traffic data. Stated in these terms, simulation modeling is a proper subject for application of statistical tools and methods. Further, the emphasis on judging simulation models in terms of output performance rather than in terms of structural or parameter accuracy coincides with the modern emphasis in statistics on nonparametric and empirical process methods. These observations will be transparent to any statistician, but are not universally recognized in applied science. Often, calibration is treated as a computational problem in which statistical properties are neglected. Consequently, simulation exercises may fail to take advantage of the opportunities statistics can provide for controlling simulation error and producing statistically reliable results. This paper examines the basic statistical features of simulation models, and describes how calibration procedures can be made consistent with sound statistical inference. The statistical tools I will use are relatively elementary, but it has proven quite fruitful to draw out their implications in the language of simulation modeling. I will outline results on several topics, and use illustrations drawn mostly from economics, finance, and transportation to illustrate their usefulness in applications. The message for statisticians is that there is a large domain of applied science, often viewed as outside the pale of serious statistics, that in fact offers productive opportunities for application of modern statistical methods.

1. Through the elementary example of computation of goodness of fit statistics and minimum chi-square estimation, I introduce the basic issue of accumulation of numerical approximation errors in statistical data processing, and illustrate how ancillary randomization can be used to control error accumulation. In contrast to conventional applications of numerical analysis methods where one wants to approximate a single intractable expression as closely as possible, statistical numerical analysis often seeks to approximate averages of intractable expressions. In the latter case, a "less is more" rule may apply, with inaccurate but unbiased approximations of individual terms leading via a law of large numbers to accurate approximations of averages.

2. A commonly used calibration method for simulation models is to draw a sample from a modeled data generation process (DGP), and then adjust trial parameter values for this process until the simulated sample resembles a real sample. Closeness may be judged in terms of moments or other sample features, such as parameter estimates in a simple, but not necessarily correctly specified, model. The statistical properties of calibration can be developed by recognizing that simulated samples drawn from the modeled DGP form an empirical process indexed by the model parameters. When the calibration process can be interpreted as generalized method of moments estimation of the model parameters, I outline an elementary argument for the large-sample properties of calibration estimators, and use the framework of this argument to make two points. First, if calibration is to have satisfactory statistical properties, it is essential that the simulated sample empirical process be stochastically equicontinuous. This can be achieved through some simple rules for sampling from the modeled DGP that avoid chatter as trial parameters vary. Second, estimation error arising from the simulation process is similar to that arising from real data noise, and is subject to the same statistical treatment, such as control of noise in simulation sample averages through laws of large numbers. The approach is illustrated by the use of mixed multinomial logit models to describe economic choice behavior.

3. Goodness of Fit tests and Minimum Chi-Square estimators are traditional tools of applied statistics, applicable when responses are discrete or can be partitioned into categories. By refining the partition with sample size and using simulation methods to approximate cell probabilities, these tools provide a practical route to statistical inference that is asymptotically efficient for regular parametric problems, well-behaved for many non-regular problems, and adaptable via the method of sieves to non-parametric estimation.

4. There are many common features and tools in simulation modeling and the use of resampling methods to infer population features. A non-parametric test for second-degree stochastic dominance among stock portfolios illustrates the use of simulation methods to form a test statistic and to approximate its distribution under the null or local alternatives.