УДК 519.22

MONTE CARLO OUTPUT ANALYSIS

Babina O.I.
Scientific supervisor - Associate professor Suray V.N.
Siberian Federal University

Introduction

An impossibility to experiment on real object in economics often leads to the use of imitating modeling. One of the methods of imitating modeling is Monte-Carlo simulation. Monte Carlo simulation is a type of simulation that relies on repeated random sampling and statistical analysis to compute the results. This method of simulation is very closely related to random experiments, experiments for which the specific result is not known in advance. In this context, Monte Carlo simulation can be considered as a methodical way of doing so-called what-if analysis.

We use mathematical models in natural sciences, social sciences, and engineering disciplines to describe the interactions in a system using mathematical expressions. These models typically depend on a number of input parameters, which when processed through the mathematical formulas in the model, results in one or more outputs. A schematic diagram of the process is shown in Figure 1.

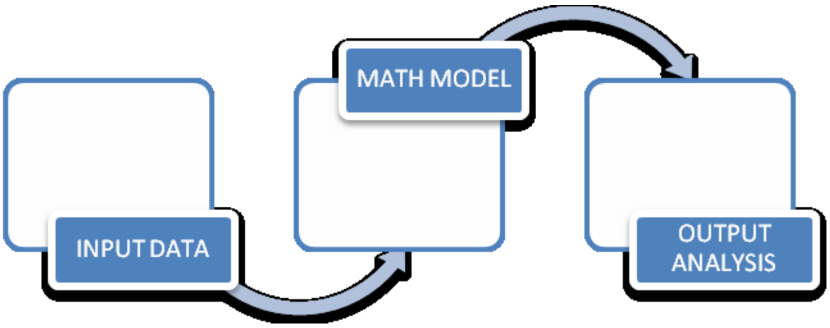

Figure 1: Mathematical models

The input parameters for the models depend on various external factors. Because of these factors, realistic models are subject to risk from the systematic variation of the input parameters. A deterministic model, which does not consider these variations, is often termed as a base case, since the values of these input parameters are their most likely values. An effective model should take into consideration the risks associated with various input parameters. In most circumstances, experimenters develop several versions of a model, which can include the base case, the best possible scenario, and the worst possible scenario for the values of the input variables (Figure 2).

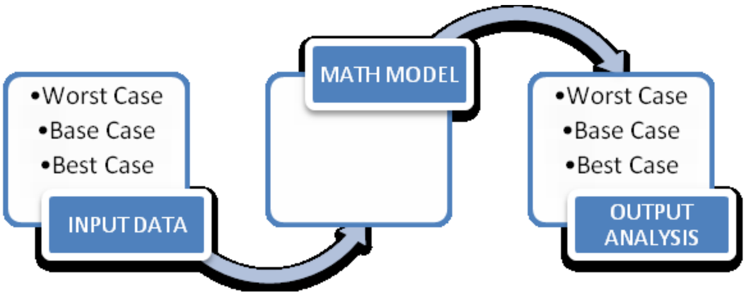

Figure 2: Case-based modeling.

This approach has various disadvantages. First, it might be difficult to evaluate the best and worst case scenarios for each input variable. Second, all the input variables may not be at their best or worst levels at the same time. Decision making tends to be difficult as well, since now we are considering more than one scenario. Also, as an experimenter increases the number of cases to consider, model versioning and storing becomes difficult. An experimenter might be tempted to run various ad-hoc values of the input parameters, often called what-if analysis, but it is not practical to go through all possible values of each input parameter. Monte Carlo simulation can help an experimenter to methodically investigate the complete range of risk associated with each risky input variable.

In Monte Carlo simulation, we identify a statistical distribution which we can use as the source for each of the input parameters. Then, we draw random samples from each distribution, which then represent the values of the input variables. For each set of input parameters, we get a set of output parameters. The value of each output parameter is one particular outcome scenario in the simulation run.

We collect such output values from a number of simulation runs. Finally, we perform statistical analysis on the values of the output parameters, to make decisions about the course of action (whatever it may be). We can use the sampling statistics of the output parameters to characterize the output variation.

Monte carlo simulation output analysis

The result of the Monte Carlo simulation of a model is typically subjected to statistical analysis. For each set of random numbers (or trials) generated for each of the random variable, the model formula is used to arrive at a trial value for the output variable(s). When the trials are complete, the stored values are analyzed. Averaging trial output values result in an expected value of each of the output variables. Aggregating the output values into groups by size and displaying the values as a frequency histogram provides the approximate shape of the probability density function of an output variable.

The output values can themselves be used as an empirical distribution, thereby calculating the percentiles and other statistics. Alternatively, the output values can be fitted to a probability distribution, and the theoretical statistics of the distribution can be calculated. These statistics can then be used for developing confidence bands. The precision of the expected value of the variable and the distribution shape approximations improve as the number of simulation trials increases.

Formulas for basic statistical analysis

In this section, we show the formulas for the basic statistical analysis for a set of output values. Let us assume that we have N values for each of the output parameters, each value represented as x_i; $i = 1...N$. Note that, these are the estimates of the complete population from the simulated sample, so we use sample statistics. The formulas for the basic statistical analysis for a set of output values are:

1. **Mean** $(\bar{x}) = \frac{1}{n}\sum_i x_i$
2. **Median 50th percentile**
3. **Standard Deviation** $s = \sqrt{\frac{1}{N-1}\sum_i (x_i - \bar{x})^2}$
4. **Variance** $s^2 = \frac{1}{N-1}\sum_i (x_i - \bar{x})^2$
5. **Skewness** $Skewness = \frac{\sum_i (x_i - \bar{x})^3}{(N-1)s^3}$
6. **Kurtosis** $Kurtosis = \frac{\sum_i (x_i - \bar{x})^4}{(N-1)s^4} - 3$

7. **Coeff. of Variability** $Coeff.of\ Variability = \frac{s}{\bar{x}}$

8. **Minimum** ($x_{\min}$) $x_{\min} = \min_i x_i$

9. **Maximum** ($x_{\max}$) $x_{\max} = \max_i x_i$

10. **Range Width** $Range\ Width = x_{\max} - x_{\min}$

$$Mean\ Std.Error = \frac{s}{\sqrt{n}}$$

11. Mean Std. Error

Other than calculating the basic statistics, one can also calculate the capability statistics in case of a six-sigma-based simulation, or perform sensitivity analysis to find out the input variables which cause the predominance of variation in the values of the output parameter of interest.

Example of simulation model output of estimation of risks of the investment project at the industrial enterprise

Table 1 shows an example of the calculated basic statistics which result after a Monte Carlo simulation has been performed. The table shows the output analysis from 501 trials. For each method of simulation (simulation model of estimation of risks of the investment project at the industrial enterprise), the table shows the basic statistics involved with the values of the output parameter, like mean, median, mode, variance, standard deviation, skewness, kurtosis, coefficient of variability, and so on.

Table 1: Basic statistics of simulation model of estimation of risks of the investment project at the industrial enterprise

Parameters	Values
Number of Trials	501
Mean	49851,71528
Median	45915,73055
Mode	--
Standard Deviation	28154,79167
Skewness	-0,811789632
Kurtosis	0,265040814
Range Width	119209,6681
Minimum	304,7804035
Maximum	119514,4485
Sum	24975709,36
Coeff. Of Variability	0,580

Conclusion

Monte Carlo simulation is a very useful mathematical technique for analyzing uncertain scenarios and providing probabilistic analysis of different situations. In this article, we have discussed the methodology, theoretical basis, and Monte Carlo output analysis**.**