Advanced Practical 2022/2023
Operations Research Case

Lecture: Ingredients of a Simulation Project

Guanlian Xiao

Department Operations Analytics
Vrije Universiteit Amsterdam

June 2023

Steps in a Simulation Project

1. Define the problem of interest.
2. Gather relevant data.
3. Formulate a mathematical model that represents the problem.
4. Validate the model [redefine when needed].
5. Construct a simulation model.
6. Write computer code.
7. Verify the computer program.
8. Run experiments.
9. Output statistics and analysis.

Input Modelling

Given data $x_1, \ldots, x_n$ of an input variable X, and suppose to generate samples of X.

- Fit a theoretical distribution:
  1. Hypothize a parameterized family of distributions, e.g. Gamma(α, λ).
  2. Estimate parameters, e.g. by maximum likelihood.
  3. Test goodness-of-fit, visually by histogram and Q-Q plot, and empirically by chi-square or Kolmogorov-Smirnov test.
- When no fit, an alternative is to simulate directly from the data.
- Another alternative: construct the empirical distribution function of the data, en simulate.

Validation

- Is the process of determining whether a simulation model (not the program!) is an accurate representation of the system.
- In other words, did we build the right model?
- Techniques include
  - *(i).* Validation of model assumptions: concerns the type of model (queueing, inventory, etc), variables and parameters.
  - *(ii).* Validating input-output relations.
  - *(iii).* Validating input-output data: establishing whether the output data from the simulation model resemble the output data from the actual system. It makes heavily use of statistical procedures and testing.

Statistical Validation

- We apply statistical testing for:
  - *(i).* validating the choice of input variables and processes of the simulation model;
  - *(ii).* validating the simulation model by comparing statistically its output results with the output data of the real system.
- For 1. we apply maximum likelihood estimation and goodness-of-fit tests to ascertain whether an assumed distribution is consistent with a given set of data.
- For 2. we test the hypothsis that two seperate samples of data come from the same underlying population. (One sample is generated by simulation, the other are real data.)

Random Numbers

- Random number generator for sampling from uniform $(0, 1)$.
- Sampling from distributions by general methods (inverse transform or accept-reject), or by special designed algorithms.
- The Python module `scipy.stats` facilitates sampling from many theoretical distributions.
- See supporting document in Canvas.

Verification

- Has the conceptual simulation model been correctly translated into a computer program?
- In other words, does the program what it is supposed to do?
- Techniques include
  - *(i).* Write and debug the program in small subroutines.
  - *(ii).* Make flow diagrams.
  - *(iii).* Run the simulation under a variety of input parameters (specifically extreme parameter values).
  - *(iv).* Utilise a trace (print state, event list, counter variables after each event).
  - *(v).* Run the the model under simplifying assumptions for which the performance measures are known or can be computed or satisfy known relations (e.g. Little's formula).

Output Analysis

Inspect the simulation output for satisfying theoretical properties.

- Convergence for increasing sample size (strong law).
- Is the estimator(approximately) normal distributed (central limit theorem)?
- Are the observations of different runs independent?
- Is the sample size large enough (but too large)?