ACE Experimental Design: An Introduction

Alan G. Isaac

American University

2015-07-03

Simulations as Experiments

"A simulation study is an experiment that needs to be designed."

—[kelton-1999-dse]

Purposes of Simulations

Find a **useful** simple model of a complex process.

- understanding: why results emerge; which factors are important; how specific factors influence the results; hypothesis testing
- prediction: anticipate real-world responses to controllable factors
- control/optimization:
  - search for “optimal” configuration of controllable factors
  - search for “robust” configurations of controllable factors

In this course, we emphasize the role of simulation modeling in producing new understanding.

Experimental Design in CSS

Need for planning: avoid time-consuming inefficiencies by adequate design

- stopping criteria (terminal or steady state)
- choice of metrics (moments, quantiles, ...)
- role of randomness (seeds for PRNGs; use of common random numbers (CRNs))

Design of Experiments (DoE):

- rules for the effective collection of responses from experimental designs
- originates from real world experimentation; we apply DoE techniques to experiments with artificial computer worlds

Experimental Design in CSS

Design of Experiments (DoE) is not used as widely or effectively in the practice of simulation as it should be.

Possible explanations:

- ignorance: DoE literature is unfamiliar to CSS researchers (different specialization).
- myopia: CSS researchers focus on building not analyzing their models.
- empahsis: DoE literature primarily addresses a different audience (real-world experimentation) and not the needs of simulation research.

Terminology

factor an input or parameter (quantitative or qualitative) that can take on more than one value

- decision factors (controllable)
- noise factors (exogenous)

factor levels values of factors (usually numerically coded)

scenario a parameter configuration: specifies a factor level for each factor

replicate a scenario + a random number sequence

simulation maps a replicate into outputs (e.g., produces an observation or an observation time series)

meta-model simplified representation of the scenario$\rightarrow$output relationship (e.g., a regression model) responses

- focal outcome variations in response to treatment variations

experimental design

1. an experimental framework
2. predictions of responses based on one or more hypotheses (what are the questions?)
3. a detailed description of how the hypotheses will be experimentally tested.
    - the treatments
    - the data collection
    - the approach to data analysis

Experimental Design in CSS (cont.)

experimental framework a simulation model that permits systematic variation in one or more treatment factors.

hypothesis a testable statement about the relationship between treatment factors and responses (simulation results)

treatment a configuration of treatment factors

treatment factor a model parameter whose values can be manipulated by the experimenter.

experiment multiple simulation runs with specified variation in treatment factors

Hypothesis Testing

1. predict the effects of changes in the treatment factors
2. run the simulation to confirm or disconfirm those predictions.

requires a specification of the:

- treatment factor variation in the experiment,
- simulation results that will be used to test the hypothesis,
- criteria determining whether or not these results reject the hypothesis.

Basic Components of Experimental Design

- decide what responses to examine (i.e., what will count as results)
- choose one or more treatment factors (i.e., model parameters to vary systematically)
- specify what values of the treatment factor you will consider
- specify the fixed values of all other model parameters.
- specify your hypotheses about the effect of your treatment factor(s) on your responses
- specify the number of replicates for each scenario
- specify the random number generator and seed(s) for each run
- specify the stopping criterion for the simulation (e.g., the number of iterations for each replicate)

Basic Components of Results Reporting

For each value of the treatment parameter, make a table showing the following:

- the treatment value,
- the random seed,
- and the final values of your focus variables at iteration (i.e., the values when the simulation stops)

For each value of the treatment parameter, report descriptive statistics for your focus variables. (At the very least, report the mean, standard deviation, and histogram.)

Design Matrix

design matrix: specifies all considered combinations of factor levels.

design point: a row of the design matrix; a scenario

E.g., if we have just two factors (A and B) with levels A1, A2, B1, B2, B3 we end up with six scenarios.

A	B
A1	B1
A1	B2
A1	B3
A2	B1
A2	B2
A2	B3

Problem of Dimensionality

Rule: if k is a big number, then 2^k is a very big number.
If we have k factors each with just 2 possible levels, that gives us 2^k possible combinations.

full-factorial experiment considers all 2^k scenarios

fractional-factorial experiment: restrict the scenarios considered by imposing constraints (e.g., C=A*B) `http://www.itl.nist.gov/div898/handbook/pri/section3/pri334.htm`

Response Surface

images/kelton-response01.png

Response Surface: Contours

images/kelton-response02.png

Estimated Response Surface

images/kelton-response03.png

Estimated Response Surface: Contours

images/kelton-response04.png

Estimated Response Surface

images/kelton-response05.png

Estimated Response Surface: Contours

images/kelton-response06.png

Response of Response

How sensitive is the response to the factors?

Direct approach (finite differences):

- make a run with the factor of interest set to a value
- perturb the factor value and make another run
- compute slope (difference quotient)

Indirect approach:

- fit a meta-model
- consider the partial derivatives

Single-run methods:

- perturbation analysis—move the factor during the run, track new trajectory as well as trajectory if the perturbation had not been made

Meta-Model

response = f(factors)
E.g., summarize with a regression.
Comment: we may have many different responses in a simulation, with a different meta-model for each.

Optimum Response

Indirect:

- fit a meta-model
- optimize it using calculus

Staged

- fit a meta-model of low dimensionality
- optimize it using calculus
- optimum determines a search region for further simulations

Black box

- simulation is a black box function
- optimization run by any external tool (e.g., NetLogo's Mathematica link)

Optimization using External Services

images/kelton-optimum.png

Example: Schelling Segregation Model

The simulation model is specified by our code but should be given a supporting description.
Model parameters include:

- the size of the world (default: 8 by 8),
- the number of agent types (default: 2)
- the number of agents of each type (or, the total number of agents and their ratio)
- the “happiness cutoff” parameter
- the relative size of the different agent classes;
- the initial location of the agents.

Example: Schelling Segregation Model

We can pick any parameter as a treatment factor.

Illustrative hypothesis: increases in the happiness cutoff parameter increase the likelihood of a segregated final outcome.

Turning the hypothesis into a prediction: given 25 agents each of two agent types and an initial random distribution of agents in the world, a rise in the happiness cutoff increases the final (T=1000) segregation measure.

Seeding the PRNG

Note that this experimental design is probabilistic.
Therefore, for each variation in the model's treatment factors, you must run multiple simulations.
Typically, each run will be based on a different programmer-specified setting of a random seed for the pseudo-random number generator (PRNG) used in the simulation.

Seeding the PRNG

Specification of the PRNG and seed renders the results deterministic (i.e., capable of exact replication).

Therefore, you must record as data the random seed used for each run (along with the values of your treatment-factors) and all user-specified or default settings for maintained parameter values and simulation control options.

Recording the Data

- For each run, record the degree of segregation displayed by the resulting agent location pattern.
- Report “descriptive statistics” that summarize these experimental segregation findings.
- Based on these descriptive statistics, draw conclusions regarding whether or not your hypothesis appears to be supported by your observations.

These descriptive statistics would typically include (at a minimum) the sample mean value, the sample standard deviation, and possibly also a histogram, for the degree of segregation observed across runs.

Descriptive Statistics (Very Basic Review)

Suppose you have conducted N runs of an experiment using N different seed values for your pseudo-random number generator.

Suppose you have recorded N observations X1, X2,..., XN regarding some quantifiable experimental outcome of interest X (e.g., the degree of segregation).

The sample mean is the arithmetic average value of your N observations.

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

Descriptive Statistics (basic review continued)

Variance (σ^2): the average squared deviation from the mean.

standard deviation: the square root of the variance.

- often good measures of how dispersed the data are: a low variance means that the data are more tightly clustered around the mean.
- we usually use the standard deviation, because it has the same units as the mean, which makes it easier to interpret.

Descriptive Statistics (review continued)

For example, given observations X1,...,XN:

$$\sigma_X^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2$$
$$= \frac{1}{N}\sum_{i=1}^{N}(X_i^2 - 2X_i\bar{X} + \bar{X}^2)$$
$$= \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \bar{X}^2$$

$$\sigma_X = \sqrt{\sigma_X^2}$$

Note: there are two common measures of variance (and thus of standard deviation). A common alternative to the measure we adopt divides by (N-1) instead of N. See `http://en.wikipedia.org/wiki/Standard_deviation` for details.

Graphics

histogram

- a standard graphic for simulation outcomes.
- puts observations into bins based on their values
- plots the number (or relative frequency) of values in each bin.

Histogram

You might construct a frequency histogram for N observations as follows.

1. Bin your observations into bins B1 ... BN
2. For each possible bin, Bi, let #Bi denote the number of observations falling in that bin.
3. Produce the bin relative frequencies bi = #Bi/N
4. plot the points (i, bi) This is your relative frequency histogram.

Histogram Example

Suppose N = 1000 and you have two bins, for outcomes above (bin 1) or below (bin 2) some segregation measure cutoff.
Suppose 300 observation fall into bin 1.
Then your frequency histogram plots two points: (1,.30) and (2,.70).

Histogram

A histogram can indicate a failure of the data to have a central tendency. (Pareto Distribution)

This is important for our interpretation of our descriptive statistics. E.g., sample mean and sample standard deviation are most useful observations cluster around the mean value.

Schelling Model: Response Surface

images/schelling-response01.png

Schelling Model: Response Estimate

After looking at the response surface, seek a meta-model:

$$y = beta_0 + beta_1 x_1 + \beta_2 x_2 + beta_3 x_1^2 + \beta_4 x_2^2$$

- y pct_same_type
- x1 population_density
- x2 pct_same_type_desired

Schelling Model: Responses (blue) v. Fitted Values (yellow)

images/schelling-response02.png

Schelling Model: Response Estimate

```
import scikits.statsmodels as sm
y = pct_same_type
x1 = population_density
x2 = pct_same_type_desired
dummy = x2 >= 80
X = np.column_stack((x1, x2, x1**2, x2**2, dummy))
X = sm.add_constant(X)
results = sm.OLS(pct_same_type, X).fit()
print results.summary()
yf = results.fittedvalues
```

Schelling Model: Response Estimate

```
=========================================
|            coefficient           prob.|
-----------------------------------------
| x1           -73.2431       2.138E-08|
| x2             1.1802       1.080E-13|
| x3            40.3979        0.001051|
| x4            -0.0074       3.685E-05|
| x5           -28.4147       1.077E-12|
| const         73.8049       4.204E-45|
=========================================
```

Schelling Model: Responses (blue) v. Fitted Values (yellow)

images/schelling-response03.png

References

[kelton-1999-dse] Kelton, W David. (1999) "Designing simulation experiments". In (Eds.) *Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future - Volume 1*, Phoenix, Arizona, United States: ACM.

[kleijnen-2005-ijoc] Kleijnen, Jack P C, Susan M Sanchez, and Thomas W Lucas and Thomas M Cioppa. 2005. A User's Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing* 17, 263--289.