

Introduction to Computer Simulation

Simulation – The process of designing a model of a real system, implementing the model as a computer program, and conducting experiments with the model for the purpose of understanding the behavior of the system, or evaluating strategies for the operation of the system.

Monte Carlo---refers to the generation of random numbers

Similar to a roulette wheel

Adds statistical variation to simulation models

Uses of Simulation

- Allows companies to ask many “what if” questions about changes in their systems without actually changing the systems themselves
- Used in situations where no tractable mathematical model exists
 - Supermarket queues where customers can move freely between lanes
 - Which of several possible investments to choose (risk analysis) in the face of possible future governmental regulation scenarios
 - Inventory decisions in the face of unknown demand
 - Competitive bidding situations

Uses, cont.

- **Understand the current system and explain why it is behaving as it is**
- **Explore changes to the current system, typically to try and improve it**
 - **For example, increase factory output by**
 - **adding more machines...?**
 - **speeding up existing machines...?**
 - **reduce idle time by better maintenance...?**
- **Design a new system from scratch**
 - **Typically to try and design a system that fulfills certain (sometimes statistical) requirements at minimum cost**

Limitations of Simulation

- **It is not an optimization tool---to seek improvements you must...**
 - **make a change**
 - **run the simulation to see if an improvement has been achieved or not**
 - **repeat**
- **Large amounts of computer time can be consumed by this process!**
 - **Many runs are necessary for statistical significance**

Advantages and Disadvantages of Simulation

- **Advantages**

- Easy to build the model
- Can handle otherwise intractable problems
- Not subject to so many assumptions
- Easy for decision maker to understand
- Cheap and safe
- Good for comparisons

- **Disadvantages**

- Hard to debug (randomness)
- No guarantee of optimality
- Hard to establish validity
- Costly to build and maintain complex models
- Can't produce absolute numbers

**BOTTOM LINE: USE SIMULATION ONLY WHEN
ADVANTAGES OUTWEIGH DISADVANTAGES**

Other Issues

- **Computer languages:**

- **General purpose languages can be used**

- Fortran
- Pascal
- C

- **Simulation languages**

- GPSS
- Simscript
- SLAM

- **Randomness**

- Random number generation
- Utilities on computer
- Built in to simulation languages

Other Issues (continued)

- **How do you start? – Initialization bias**
- **How do you stop? – Stopping rules**
 - **Number of a type of event**
 - **Time**
- **Statistics**
 - **How many runs?**
 - **Experimental design**
 - **Analysis of output**
- **Validation and Verification (Is the model any good?)**

Types of Simulation Models

- **Static Simulation (we will do these in this class)**
 - **No element of time included**
 - **Useful for developing statistical distribution information about a system (that can't be done in closed form)**
- **Discrete Event Simulation**
 - **Dynamic - looks at system over time**
 - **May be “time-stepped”**
 - **May be “event-stepped”**
- **Continuous time simulation**
 - **Based on differential equations that describe system**
 - **No “stepping”**

Random Number Generation

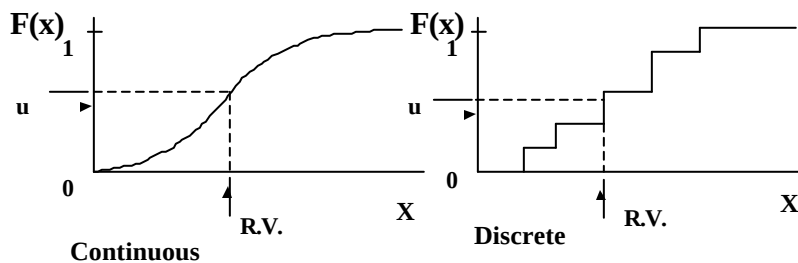
- Random numbers from RNG's are not really random
- Desirable characteristics of RNG's
 - (1) Know distribution (usually $U(0,1)$)
 - (2) Independent
 - (3) Reproducible but changeable
 - (4) Fast computation
 - (5) Small memory requirements
 - (6) Portable
 - (7) Easy to implement
 - (8) Numbers should be dense

Non-Uniform Random Numbers

- These algorithms used to generate $U(0,1)$ variates
- Need to generate other PRN with other distributions

Inverse Method:

- Want $X \sim F(\bullet)$ CDF
- Draw $u \sim U(0,1)$
- $X = F^{-1}(u)$ where $F(\bullet)$ is invertible



Non-Uniform Random Numbers

(2) Convolution – Add them up

ex. Add several Bernoulli RV's to get a Binomial

(3) Composition – Use when CDF can be written as convex combination of others

– Draw $U(0,1)$ to find the piece, then draw from that piece

(4) Acceptance - Rejection

– Useful when no CDF available

– Range of X must be finite and $f(x)$ must have upper and lower bound

★ *Fortunately - many applications have generators for different distributions readily available as functions*

Input Distributions

Discrete Distributions

- Bernoulli - model situation with 2 choices
 - Flipping a (maybe biased) coin
 - e.g., success/failure, right/left, 0/1
- Binomial - n independent Bernoulli trials
 - Number of successes in n trials
- Poisson - used to model number of occurrences of an event in some given time period
 - Number of arrivals to a bank in an hour
 - Number of accidents at an intersection in a year
- Geometric - number of trials until the first success
 - Number times light is turned on before it fails

Input Distributions

Continuous Distributions

- **Uniform** – models process when outcome is equally likely to be between a and b
 - Compute as $x = a + (b - a)u$, with $u \sim U(0,1)$
- **Exponential** – models time until some occurrence of some event
 - has range of 0 - ∞
 - one parameter distribution $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$
 $E[X] = 1/\lambda$, Arrival rate = λ
 - Memoryless: $P[X > x + x_0 | X > x_0] = P[X > x]$
 - Nice mathematical properties
 - Basis of queuing theory

13

Input Distributions

- **Normal Distribution** - Model error
 - Model anything symmetrically distributed around a central point
 - Hard to generate Normal Random Numbers
(No closed form for CDF)
 - Remember: $-\infty < x < \infty$
- **Triangular Distribution** - useful and easy to use
 - Range of X is [a,b] Mode is c
 - $E[X] = (a + b + c)/3$
 - $Var[X] = (a^2 + b^2 + c^2 - ab - ac - bc)/18$

14

Input Distributions

- **Triangular Distribution (continued)**
 - Can be used with limited information
 - Estimate (a) minimum (2) maximum (c) most likely
 - Set $a = \min$, $b = \max$,
 - Method:
 - (1) Draw $u \sim U(0,1)$
 - (2) If $u \leq (c - a)/(b - a)$ then $x = a + \sqrt{(b - a)(c - a)u}$
Else $u > (c - a)/(b - a)$ then
$$x = b - \sqrt{(b - a)(b - c)(1 - u)}$$

15

Empirical CDF

Have sample of data (e.g., completion times for jobs on machine)

$$x_1, x_2, x_3, \dots, x_n$$

– Get order statistics:

$$x_{[1]}, x_{[2]}, x_{[3]}, \dots, x_{[n]} \text{ such that } x_{[i]} = x_{[i+1]}$$

- approximate: $F(x_{[i]}) = (i-1)/(n-1)$
- define $F(x)$ for $x_{[i-1]} < x < x_{[i]}$ by interpolation
- Generate RN's using inversion
- Packages handle these
- Careful: Ignores the tails of the distribution

16

In Excel...

- **Random number generation**---generate sample values for each uncertain input variable
 - A sample, by definition, does not represent the entire population---but as sample size increases, the sample becomes more representative of the underlying population
- A random number formula must be placed in each cell that represents a random, or uncertain, independent variable in the simulation

Random Numbers (Without Add-ins)

- All spreadsheets can generate random sequences of numbers between 0 and 1
- **In Excel: =RAND()**
 - Uniformly distributed and probabilistically independent
- Other distributions available, but can use Uniform to generate others
 - $\text{INT}(n*\text{RAND()})+1$ returns numbers between 1 and n
 - Example: with $n = 6$, this simulates rolling a die
- If you recalculate (F9) all of the numbers will change
 - To freeze a certain set of random numbers:
Select range---Copy---Paste Special (Values)

Some Common Distributions

- Discrete uniform distribution (a to a+n-1)
 - $=\text{INT}(n*\text{RAND}())+a$
- Continuous uniform distribution (a to b)
 - $=a+(b-a)*\text{RAND}()$
- Symmetric triangular distribution (a to b)
 - $=a+(b-a)*(\text{RAND}()+\text{RAND}())/2$
- Normal distribution
 - $=\text{NORMINV}(\text{RAND}(),\mu,\sigma)$
- Exponential distribution
 - $=(-1/\lambda)*\text{LN}(\text{RAND}())$

19

Steps for Using a Discrete Distribution (Without Add-ins)

Establish probabilities for potential outcomes (e.g., .2, .6, .2)

Determine cumulative probabilities (e.g., .2, .8, 1.0)

Define a correspondence between the outcomes and the random number values

As a random number is “drawn” the corresponding random event “occurs” in the simulation model (with the same relative frequency as in the probability distribution)

LOOKUP functions (V or H) pick appropriate value from the table

20

VLOOKUP Function

“Lookup table” has cumulative probabilities and possible demand values

Cumulative probabilities must contain beginning of range

VLOOKUP selects the row that contains the largest value that is *less than* the random number

Format: =VLOOKUP(x,y,z) where

x = *cell of (random) number to look up*

y = *table cell range (cum prob #s in first column)*

z = *column of table to get desired value from*

Example: in Cell C13, =VLOOKUP(B13, E5:F9,2)

Random Number Generation Using Crystal Ball

- RNG function in each cell representing a random event...
 - CB.Uniform(*min,max*)
 - CB.Normal(*m,s,min,max*)
 - CB.Triangular(*min,most likely,max*)
 - CB.Custom(*range*)
 - etc.
- Keep in mind continuous vs. discrete
- Deselect “Set to Distribution Mean” in Cell Preferences

Replicating the Model (Without Add-ins)

Many iterations are necessary for statistical significance

A one-way Data Table can be used to replicate model

Fill in a series from 1 to the number of iterations you want (at least 100)

Next to your series, in the first row, set the cell = to the simulation results cell

Select entire table, go to Data menu, select Table, select **any blank cell** as Column input cell

To save these results, Copy and Paste Special (Values)

Using Data Tables

- As used here, Data Table command “tricks” Excel into calculating simulation many times and storing the solutions
 - Model recalculates for each value in series, even though this value (1--100, for example) have nothing to do with the problem
- Two-way data tables allow several values to be substituted in the problem
 - Formula must be in upper left corner of table

Replicating the Model with Crystal Ball

- Select the output cells to track using **Define Forecast**
 - Repeat for additional output cells
 - Select **When Stopped** for fastest run time
- Select the number of iterations using **Run Preferences**
 - At least 100 replications recommended
- Start simulation

Data Analysis (Without Add-ins)

- Use **Data Analysis** tool to summarize and analyze the data in order to draw conclusions
 - **Descriptive Statistics** dialog box provides mean, median, mode, standard deviation, min, max, confidence level, and so on
 - **Histogram** helps determine the distribution of outcomes
 - Cumulative frequencies also useful in answering questions

Data Analysis using Crystal Ball

- View, Statistics will show summary stats
- View, Frequency Chart will show distribution (histogram) of results
- Run, Forecast Windows will display histograms of other output cells selected
- Run, Create Report allows you to save graphs or stats from CB in your spreadsheet
- Run, Extract Data allows you to save replication data for further analysis

Confidence Interval for Pop. Mean

Use descriptive statistics to calculate confidence interval, if desired

$\text{mean} \pm 1.96 * (\text{std dev}) / \sqrt{(\text{sample size } n)}$ gives 95% interval

Alternately, mean $\pm$ Confidence Level (95%) from table gives approximately same information

Same can be done for population proportion

Larger sample sizes will make intervals tighter

Other Features of Crystal Ball

- **Decision Table** allows you to evaluate alternate values of decision variables in problem
 - Equivalent to using two-way data table in Excel
- **OptQuest** allows you to determine optimal value of decision variables in simulation models
 - Involves many simulation replications

Special Purpose Simulation Tools

- **Computer simulation** is one of the most versatile of all management science tools
- **RAND** function in Excel allows simulation of a wide variety of problems
 - This is only a fraction of what simulation can accomplish
- **Add-in packages** to Excel (e.g., @RISK, Crystal Ball) facilitate simulation, particularly additional distributions, statistics and presentation
- Many special-purpose packages exist for more complex modeling (e.g., SIMAN, SLAM, GPSS)