Systems Engineering 520

Monte Carlo Simulation

Huseyin Topaloglu
School of Operations Research
and Industrial Engineering

Lecture 3

Cornell University

Overview

- Setting up a Monte Carlo simulation model
- Estimating expected values via Monte Carlo simulation
- Building confidence intervals for the estimates
- Estimating probabilities via Monte Carlo simulation

A Simple Model

- Interested in collecting baseball cards coming out of cereal boxes
- There are 5 types of cards
- Buy 4 boxes, come home, open all boxes and check all cards
- What is the expected number of different cards?
- E.g. if we get the cards 1, 3, 3, 5 out of 4 boxes, then the number of different cards is 3

A Simple Model

- Build a spreadsheet model
- The probability of getting each card out of a box is equal
- Generate the card coming out of a box by using the formula =FLOOR(RAND()*5,1)+1
- This formula generates an integer value between 1 and 5 with equal probabilities
- Count the number of different cards
- Replicate the experiment many times

A Simple Model

	A	B	C	D	E	F	G	H	I	J	K
1	Trial No	Box1	Box2	Box3	Box4	Card1	Card2	Card3	Card4	Card5	Differents
2	1	5	3	5	5	0	0	1	0	3	2
3	2	4	2	1	1	2	1	0	1	0	3
4	3	5	2	3	1	1	1	1	0	1	4
5	4	4	5	2	1	1	1	0	1	1	4
6	5	5	2	2	1	1	2	0	0	1	3
7	6	1	4	3	4	1	0	1	2	0	3
8	7	1	4	5	5	1	0	0	1	2	3
9	8	4	5	2	2	0	2	0	1	1	3
10	9	3	2	5	2	0	2	1	0	1	3
11	10	5	1	2	5	1	1	0	0	2	3

- The average of column K should give an idea about the expected number of different cards

Building a Confidence Interval

- Every time we recompute the cells, our estimate of the expected number of different cards changes
- How much confidence can we put on our estimate?
- Let X be the random variable that represents the number of different cards that we get
- We are interested in estimating E{X}
- Each row of the spreadsheet generates an independent sample of X

Building a Confidence Interval

- Assume that the spreadsheet has N rows
- That is, we generate N samples of X
- Call these samples $X_1, X_2, \ldots, X_N$
- Use standard confidence interval methodology

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i, \quad s_N^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X}_N)^2$$

$$\bar{X}_N \mp z_{\alpha/2} \frac{s_N}{N^{1/2}}$$

Building a Confidence Interval

- Build a confidence interval by using N = 40

$$\bar{X}_N = 2.92, \;\; s_N = 0.69,$$
$$\mathrm{CI}(95\%) = [2.70, 3.14] = 2.92 \mp 0.22$$

- Build a confidence interval by using N = 1000

$$\bar{X}_N = 2.94, \;\; s_N = 0.67,$$
$$\mathrm{CI}(95\%) = [2.90, 2.98] = 2.94 \mp 0.04$$

Building a Confidence Interval

- If we make 40 trials, then the expected number of different cards lie in the interval [2.70, 3.14] with probability 0.95
- That is, if we make 40 trials, then we can state with 95% confidence that the expected number of different cards lie in the interval [2.70, 3.14]
- The expected number of different cards is roughly 2.92
- Therefore, if we make 40 trials, then our precision is roughly +/-0.22 or +/-8% (0.22 / 2.92 ~ 0.08)

Building a Confidence Interval

- If we make 1000 trials, then the expected number of different cards lie in the interval [2.90, 2.98] with probability 0.95
- That is, if we make 1000 trials, then we can state with 95% confidence that the expected number of different cards lie in the interval [2.90, 2.98]
- The expected number of different cards is roughly 2.94
- Therefore, if we make 1000 trials, then our precision is roughly +/-0.04 or +/-1% (0.04 / 2.94 ~ 0.01)

Number of Trials Needed for Certain Precision

- How many trials are needed to estimate the expected number of different cards with a certain precision?
- We want to estimate the expected number of different cards with +/- q precision and 95% confidence
- How many trials are needed?

$$\bar{X}_N \mp z_{\alpha/2} \frac{s_N}{N^{1/2}}$$

q

Number of Trials Needed for Certain Precision

- Choose N such that

$$z_{\alpha/2} \, \frac{s_N}{N^{1/2}} = q$$

$$N = \left[\frac{z_{\alpha/2} \, s_N}{q} \right]^2$$

- We cannot make this computation yet because we do not know s_N without making the trials

Number of Trials Needed for Certain Precision

- Note that s_{40} = 0.69 and s_{1000} = 0.68
- The value of s_N is insensitive to N
- To deal with the difficulty that we do not know s_N without making the trials, we first make a small number of pilot runs (say K) and use s_K to estimate s_N
- We compute the number of trials needed by

$$N \approx \left[z_{\alpha/2} \, s_K / q \right]^2$$

Number of Trials Needed for Certain Precision

- E.g. assume that we carry out a pilot run of 40 trials and $s_{40} = 0.69$
- If we want to estimate the expected number of different cards with +/-0.01 precision and 95% confidence, then the number of trials needed is

$$N = [1.96 \times 0.69/0.01]^2 = 18289$$

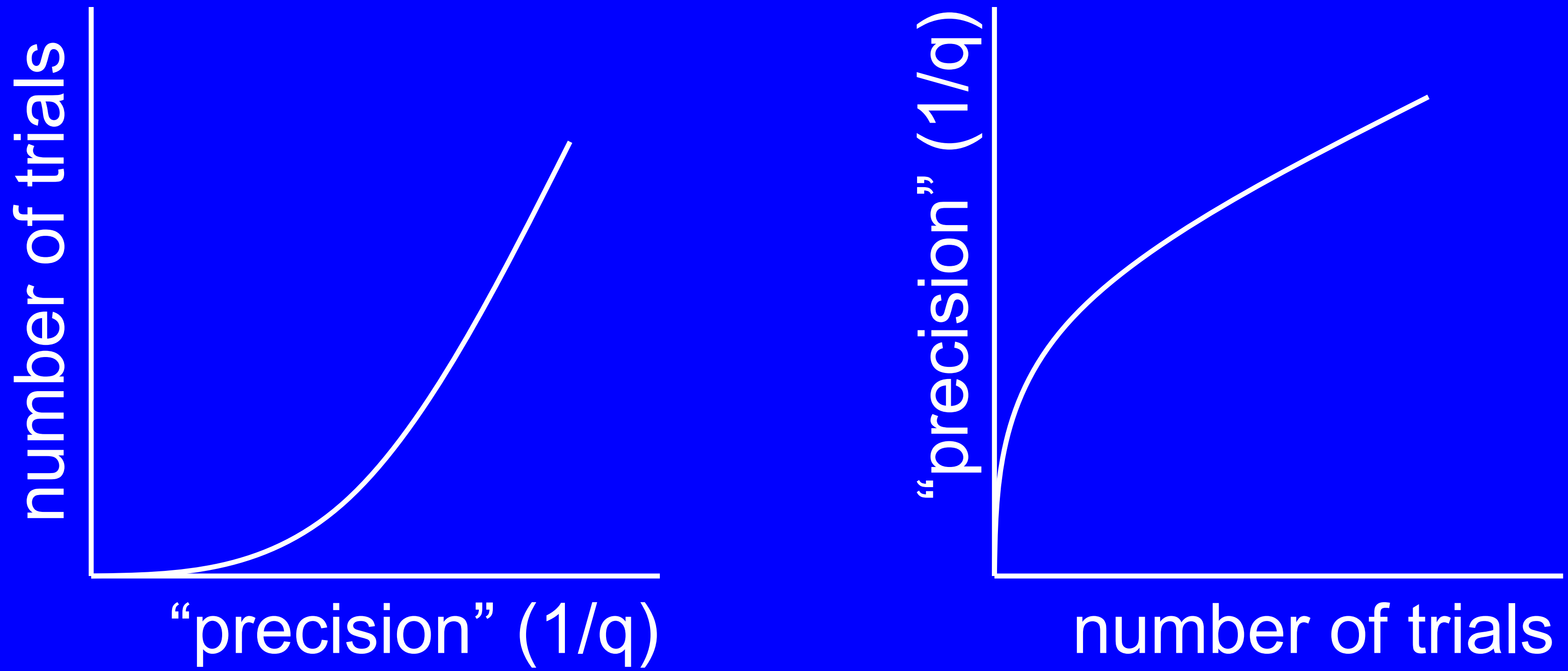
Number of Trials Needed for Certain Precision

- If we want to estimate the expected number of different cards with +/-0.005 precision and 95% confidence, then the number of trials needed is

$$N = [1.96 \times 0.69/0.005]^2 = 73159$$

- If we want to increase the precision by a factor of 2, then we need to increase the number of trials by 4
- For this reason, Monte Carlo simulation is not too useful to obtain extremely precise estimates

Number of Trials Needed for Certain Precision



Estimating Probability of an Event

- How to modify the simulation model to estimate the probability that we get 2 or more different cards?
- …
- We simply compute the proportion of trials that we get 2 or more different cards
- How precise is this probability estimate?
- We can use the same confidence interval methodology

Estimating Probability of an Event

- Why does the same confidence interval methodology work?
- Define the random variable Y

$$Y = \begin{cases} 1 & \text{if we get 2 or more different cards} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \mathbb{E}\{Y\} &= 1\,\mathbb{P}\{Y=1\} + 0\,\mathbb{P}\{Y=0\} = \mathbb{P}\{Y=1\} \\ &= \mathbb{P}\{\text{we get 2 or more different cards}\} \end{aligned}$$

- Estimating the probability that we get 2 or more different cards is equivalent to estimating E{Y}

Estimating Probability of an Event

- Compute the number of trials needed to estimate the probability that we get 2 or more cards with +/- 0.02 precision and 0.95 confidence

Embedding More Complicated Randomness

- Consider the phone company problem in the first homework assignment
- The probability density function for the length of a phone call is $f(x) = (1/4)\, e^{-x/4}$
- How do we generate a sample of the length of a phone call?
- The formula =-4*LN(RAND()) can be used to generate a sample of the length of a phone call
- What about other distributions?