ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

Unit 14: Ranking, Selection, and Optimization

Sid Banerjee

School of ORIE, Cornell University

comparison of alternate systems

till now we have seen how to:

– simulate complex discrete-event systems

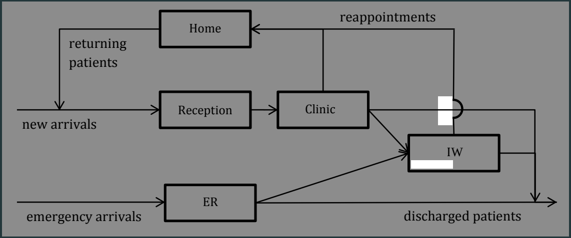
– compute statistics about these models.

we now want to use these to compare different system configs

comparing systems: main ideas

- use of simultaneous confidence intervals
- practical significance and indifference zones
- use of common random numbers

example: staffing the Fingerlakes hospital



hospital employs 15 doctors

Q: how should we allocate doctors to optimize service?

questions and models

- how do we divide the doctors between ER and clinic?
- what is the added benefit of hiring another doctor?
- is it useful to have ‘floaters’ who can go to the clinic/ER as needed?

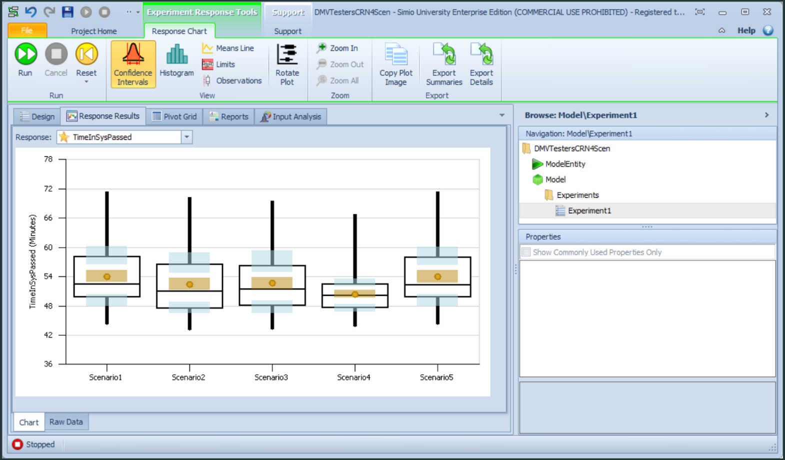
example: combating lack of diversity in companies

example: combating lack of diversity in companies

'Rooney rule': for every position, interview top male and female candidate

simultaneous confidence intervals

how do we use CIs to compare different scenarios?



does this mean that with prob 0.75, scenario 4 is the best?

simultaneous confidence intervals: the union bound

the union bound

let $A_1, A_2, \ldots, A_k$ be events. then

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_k) &= 1 - \mathbb{P}(A_1^c \cup A_2^c \cup \cdots \cup A_k^c) \\ &\geq 1 - (\mathbb{P}(A_1^c) + \mathbb{P}(A_2^c) + \cdots + \mathbb{P}(A_k^c))\end{aligned}$$

let A_i = event that the ith ci contains its true mean...

practical and statistical significance

a practically meaningful difference depends on the problem at hand:

- \$10,000 on a portfolios return
- 5 minutes in waiting time for COVID test
- 20 people being unable to connect to a Zoom meeting

statistical significance depends on sampling variability in estimates:

- a 95% confidence interval for the difference in expected time between scenarios is 4 ± 7 minutes. what can we conclude?
- what if it was 4 ± 1 minute?

controlling significance

- we use **statistical procedures** to tell us whether we can believe the difference we see in the results from two or more scenarios.
- we use the **number of replications** to control the size of the difference that is detectable; that is, to control the error in our estimates.
- **you** have to decide what difference is practically significant.

ranking and selection

- given a set of systems, simulates each for a random amount of time and returns a single system i that is estimated to be the best
- to keep this from running for ages in the event of ties or near ties, we specify an indifference zone δ – smallest difference worth detecting (practical significance)
- *"with probability ≥ 0.95, system i is the best, or is within δ of the best"*
- be careful! run time increases as $\delta \to 0$.

is comparing different simulations fair?

common random numbers

- let X and Y be rvs giving output from two different scenarios.
- want $\mu_X - \mu_Y = \mathbb{E}[X] - \mathbb{E}[Y] = \mathbb{E}[X - Y]$
- if X, Y independent (different streams) then

  $Var(D) = Var(X - Y) =$

- in general (whether independent or not)

  $Var(D) = Var[X - Y] =$

- use CRN to try to make $Cov(X, Y) > 0$.

clicker question: comparing queueing disciplines

consider an M/M/1 queue with arrival rate λ and service rate $\mu > \lambda$

suppose you build two simulation models

- in the first, you serve jobs in a First-In First-Out (FIFO) order
- in the second in a Last-In First-Out (LIFO) order

(a) the average queue length in FIFO is smaller than that in LIFO

(b) the average queue length in LIFO is smaller than that in FIFO

(c) the averages are same, but LIFO has higher variance in queue lengths

(d) the queue length distributions are identical in the two

clicker question: comparing queueing disciplines

consider an M/M/1 queue with arrival rate λ and service rate $\mu > \lambda$

suppose you build two simulation models

- in the first, you serve jobs in a First-In First-Out (FIFO) order
- in the second in a Last-In First-Out (LIFO) order

clicker question: comparing queueing disciplines part 2

in the previous setting (M/M/1 queue with arrival rate λ, service rate $\mu > \lambda$ under FIFO and LIFO service), what can we say about the time in system?

(a) the average time in system in FIFO is smaller than in LIFO
(b) the average time in system in LIFO is smaller than in FIFO
(c) the averages are same, but LIFO has higher variance in time in system
(d) the time in system distributions are identical in the two

RNG streams

- original model used a single stream (stream 0) for everything
  - scrambles the sequence
- fix using streams

streams in python

- create different rng objects for each stream
- eg. in numpy:

```
arrival_stream = np.random.randomstate(seed=0)
service_stream = np.random.randomstate(seed=1)
t = arrival_stream.exponential(1.0/arrival_rate)
```

using CRN in Markovian simulations

using CRN in Markovian simulations

using CRN in Markovian simulations

using CRN in Markovian simulations

simulation optimization

ranking and selection:

– comparing small number of systems

– need to simulate each system at least a bit

what can we do for bigger problems?

simulation optimization

- simulation optimization: **search** over different systems
- Markov decision processes: **optimize** over decisions (controls)

simulation optimization is hard

- local vs global optima
- many decision variables means huge decision space. e.g., shifts start on the hour, up to 11 agents can start each hour, 11^{24} possible solutions (systems)
- **estimation error** means we can never be certain that one solution x is better than another y.
- **simulation noise** (estimation error) can swamp the signal.

optimization bias

the estimated objective value for a minimization problem is always lower than it should be

tools and techniques

sample average approximation use a fixed run-length and common random numbers. minimize estimated function with deterministic optimization software

metamodeling fit a simpler function, e.g., polynomial, to simulation output, minimize the simpler function

stochastic approximation somehow estimate the slope (gradient) of the objective function at current point, and take a step in the opposite direction; repeat

random search given a current point or set of points, randomly choose new ones and simulate

- easily adapted to broad classes of problems
- no guarantees
- not much good with lots of variables