Simulation and Statistical Analysis of Discrete-Event Simulations

(Chapters 3 and 4 in Textbook [excluding Sect. 4.5])

- Definition
- Queuing models
- Static models
- Reliability example
- Confidence intervals
- Batch means for variance estimation
- Regenerative processes for dynamic systems

Some Advantages of Simulation (from week 1 handout)

- Often the **only type of model possible** for complex systems
  - Analytical models frequently infeasible
- Process of building simulation can **clarify understanding** of real system
  - Sometimes more useful than actual application of final simulation
- Allows for sensitivity analysis and optimization of real system **without need to operate real system**
- Can maintain **better control over experimental conditions** than real system
- **Time compression/expansion:** Can evaluate system on slower or faster time scale than real system

2

Classification of Simulation Models (from week 1 handout)

- **Static vs. dynamic**
  - ***Static:*** E.g., Simulation solution to integral $\int_{\Omega} f(\mathbf{x})d\mathbf{x}$
  - ***Dynamic:*** Systems that evolve over time; simulation of traffic system over morning or evening rush period
- **Deterministic vs. stochastic**
  - ***Deterministic:*** No randomness; solution of complex differential equation in aerodynamics
  - ***Stochastic (Monte Carlo):*** Operations of grocery store with randomly modeled arrivals (customers) and purchases
- **Continuous vs. discrete**
  - ***Continuous:*** Differential equations; "smooth" motion of object, price, etc. (physics, economics, etc.); Black–Scholes model in mathematical finance for option pricing
  - ***Discrete:*** Events occur at discrete times; queuing networks (discrete-event dynamic systems is core subject of books such as Cassandras and Lafortune, 2008, Law, 2007, etc.); buy/sell orders, etc.

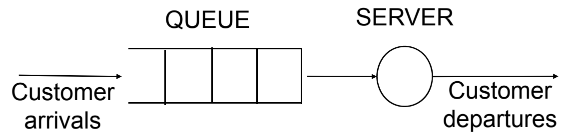
3

Discrete-Event Simulations

- Discrete event systems (DES) apply to systems with events occurring at particular (discrete) times discrete
  - DES versus continuous dynamic systems
  - Examples of DES (e.g., buy or sell order in financial markets)
- Types of DES
  - Static vs. dynamic
  - Dynamic is case of interest here
- Mechanisms for Discrete-Event Simulations
  - Event scheduling scheme
  - Process-oriented scheme (for use with object-oriented programming language)
  - Discrete event simulation languages: GASP, GPSS, SIMAN, SIMSCRIPT, etc.
- Main event characteristics: event type and time of occurrence

4

Queueing Models



- Model specification:
  - *Stochastic models*: distributions of arrival and service processes.
  - *Structural parameters*: capacity of queue; number of servers
  - *Operating policies*: admission control policy, prioritization, etc.
- *A*/*B*/*m*/*K* notation

capacity of the queue
number of servers
service time distribution
interarrival time distribution

- G: General distributions
- GI: General and i.i.d.
- D: Deterministic
- M: Exponential (Markovian)

- Example: *M*/*M*/1 queue is classical single-server queue with exponentially distributed service and interarrival times

Chapter 4 of Rubinstein and Kroese (2017)
(Textbook)

STATIC SIMULATION MODELS
Sections 4.1–4.3

6

Static Simulation Models

- Consider system of interest with state vector $\boldsymbol{x}$
- Have Monte Carlo simulation of system that produces sample values of $\boldsymbol{x}$, say $\boldsymbol{X}_i$
- "Static" simulation models may involve performance over time, but need to have specified termination point (time or event) that delivers an output
- Multiple outputs from multiple simulation runs provide i.i.d. sample for statistical analysis
- Consider scalar-valued performance function $H(\boldsymbol{x})$
  - Cost, wait time, reliability, return on investment, profit, etc.
- Analysts frequently need to know average value of $H(\boldsymbol{X})$
  - Sensitivity studies
  - Optimization
  - Bounding (mimimum/maximum values)
  - Etc.

Overall goal is to use simulation to estimate expected system performance

7

Performance Measure for Static Simulation Models

- Recall performance function $H(\boldsymbol{x})$
- Performance function is chosen to have meaning in the real (physical) world
  - While H is from Monte Carlo simulation, it is intended to "look" like real world
- The random H depends on many choices and decisions in real problem
  - E.g., if H represents firm's cumulative profit over the next year, then H depends on large number of managerial decisions
- Frequently need to study ***average*** behavior of H for given set of choices and decisions
  - Average allows us to "integrate out" the various random possibilities that could occur (before they occur)
  - Helps in planning, optimization, risk analysis, etc. to consider "all" random effects

8

Estimation for Static Simulation Models

- Recall performance function $H(\boldsymbol{x})$
- Average value ℓ is

$$\ell \equiv E[H(\boldsymbol{X})] = \int H(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}$$

- Above integral generally requires knowledge of distribution of $\boldsymbol{X}$ (i.e., to know $f(\boldsymbol{x})$)
- Unbiased estimate of ℓ from N independent simulation runs is

$$\hat{\ell}_N \equiv \frac{1}{N} \sum_{i=1}^{N} H(\boldsymbol{X}_i)$$

- Standard statistical results and methods apply to above estimate; e.g.,

$\hat{\ell}_N \to \ell$ as $N \to \infty$ via WLLN or SLLN, and

$$\sqrt{N}(\hat{\ell}_N - \ell) \xrightarrow{\text{dist}} N\big(0, \text{var}(H(\boldsymbol{X}))\big)$$

9

Confidence Intervals

- Confidence intervals are useful mechanism for characterizing uncertainty in estimate
- Confidence intervals for can be based on t-distribution or, for large N, normal distribution
- Confidence interval is random interval that covers "truth" with specified probability
- Given sample variance S^2, a level α interval for estimate of ℓ based on normal distribution approximation is

$$\mathbb{P}\left(\widehat{\ell} - z_{1-\alpha/2}\frac{S}{\sqrt{N}} \leqslant \ell \leqslant \widehat{\ell} + z_{1-\alpha/2}\frac{S}{\sqrt{N}}\right) \approx 1-\alpha$$

($z_{1-\alpha/2}$ is $1 - \alpha/2$ percentile point for normal distribution)

- Typical values for α are 0.10, 0.05, or 0.01, yielding 90%, 95%, or 99% confidence intervals:

$$\left(\widehat{\ell} \pm z_{1-\alpha/2}\frac{S}{\sqrt{N}}\right)$$

10

Hypothetical Application of Simulation and Central Limit Theorem (as in Chap. 1 slides)

- Suppose have large Monte Carlo simulation of traffic in downtown Baltimore; simulation output has unknown pdf
- City planners and traffic engineers want to consider (costly) modification in traffic signal timings and street configuration
- Hypothetical scenario: Monetary value of people's time waiting in traffic and current daily cost of road maintenance is $1 million/day (made up number!!)
- Run simulation 100 times for modified system:

  Gives sample mean cost of $0.9 million/day (including costs for changing system) and S^2 = 0.0625 (=1/16)
- From CLT, know that

$$\sqrt{N}(\hat{\ell}_N - \mu) \xrightarrow{\text{dist}} N(0, \sigma^2)$$

- Can approximate σ^2 by S^2
- Above values give (approximate) z-statistic of –4.0, providing statistical evidence of improvement (low P-value)
  - May suggest Baltimore should move forward with modification

11

Example: Estimating Reliability of Complex Systems by Monte Carlo

- Reliability determination is essential for assessment of many systems
  - Power grid
  - Aircraft
  - Computer networks
  - Software
  - Biological (human body) (used, e.g., by insurance companies)
  - Etc.
- Difficult or impossible to directly evaluate reliability of "big" systems due to complexity of relationship between subsystems and full system
- Simulation may be good tool for estimating reliability subject to having good model

12

Example: Estimating Reliability of Complex Systems by Monte Carlo (cont'd)

- Assume system is composed of *n* subsystems
- Subsystems may be probabilistically *dependent* or *independent* when operating as part of the full system
- Let X_i represent Bernoulli 0 or 1 random variable representing failure on non-failure in subsystem *i*
  - Let $\boldsymbol{X} \equiv [X_1, X_2, \ldots, X_n]$ and $p_i = P(X_i = 1)$
- Operational state of system (0 or 1) expressed as

$$Y = H(\boldsymbol{X}),\ \{0, 1\}^n \rightarrow \{0, 1\}$$

- For case of independent subsystems, system reliability given by:

$$\ell = P(Y = 1) = E(Y) = \sum_{\boldsymbol{x} \in \{0,1\}^n} H(\boldsymbol{x}) P(\boldsymbol{X} = \boldsymbol{x})$$

$$= \sum_{\boldsymbol{x} \in \{0,1\}^n} H(\boldsymbol{x}) \prod_{i=1}^{n} p_i^{x_i} (1 - p_i)^{(1 - x_i)}$$

13

Example: Estimating Reliability of Complex Systems (cont'd)

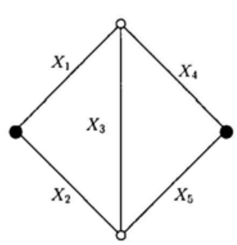
- Famous special cases of series and parallel systems:

Series system fails if ***any*** subsystem fails: $Y = \prod_{i=1}^{n} X_i$

Parallel system fails if ***all*** subsystems fail: $Y = 1 - \prod_{i=1}^{n} (1 - X_i)$

- Most practical systems more complicated than pure series or parallel above
- Example of "bridge network" from textbook (p. 111) below; $H(\boldsymbol{X})$ = 1 if connection between two black nodes shown

$$H(\mathbf{x}) = 1 - (1 - x_1 x_4)(1 - x_2 x_5)(1 - x_1 x_3 x_5)(1 - x_2 x_3 x_4)$$



(Source: Fig. 4.1 in Rubinstein and Kroese, 2017)

14

Example: Estimating Reliability of Complex Systems (cont'd)

- Difficult to carry out expectation defining ℓ if n is large
  - Structure function H not known in many real-world systems
  - Summation over 2^n options in $\{0, 1\}^n$ not feasible for large n
- Monte Carlo simulation may be useful through generating "many" (N) sample Y values to form $\hat{\ell}_N$
  - Need credible simulation (of course!)
  - Need huge number of samples for "useful" estimate with high reliability system (i.e., reliabilities 0.99 and 0.999 are very different when considering critical systems)
- Need individual subsystem reliabilities (the p_i) in executing simulation
  - The p_i can be estimated based on physical principles (e.g., Newton's laws, $F = ma$) and/or system testing
  - Method of maximum likelihood can also be used (see current project of TA Long Wang; discussed later in 553.633)

15

Chapter 4 of Rubinstein and Kroese (2017)
(Textbook)

DYNAMIC SIMULATION MODELS
Section 4.4

16

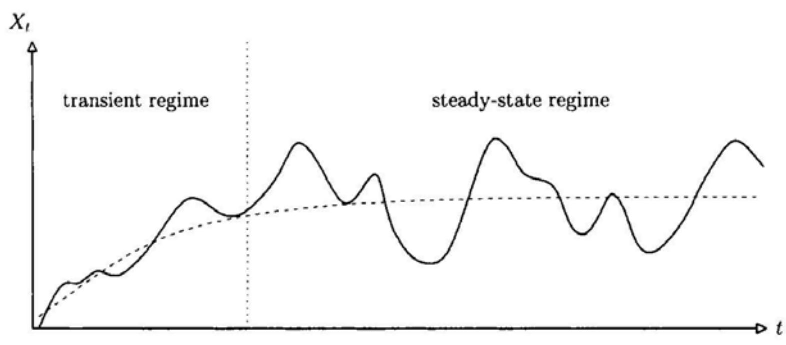
Introduction to Dynamic Simulation Models

- ***Dynamic simulation models*** pertain to systems that evolve over time and do not have natural termination time
- Overall goal (as in static case): Use simulation to estimate expected system performance
- Dynamics imply need to be concerned with finite-horizon or steady-state behavior
  - Finite horizon considers performance over time interval [0, T]
  - Steady-state considers limiting behavior, if such behavior exists; steady state requires that the distribution function of dynamic state $\boldsymbol{X}_t$ is not a function of t (or initial state)
- Transient phase of dynamic process occurs between initial state and limiting steady state
- Note that in both transient phase and steady state, the $\{\boldsymbol{X}_t\}$ are generally ***not*** independent and ***not*** constant

17

Introduction to Dynamic Simulation Models, cont'd

- Plot below is schematic of transient and steady state
- Most practical analytical analysis is for steady state (exception is Kalman filter and related particle filter; later in class)
- Dashed line in plot is mean $E(X_t)$

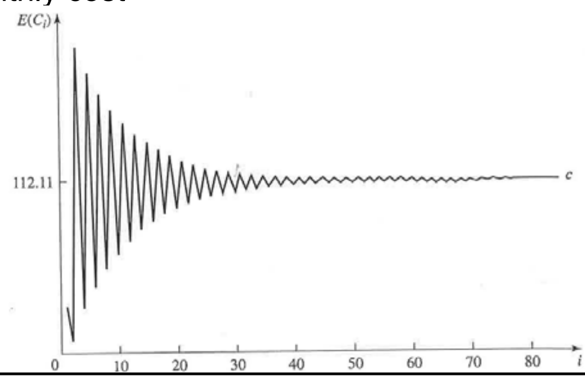


(Source: Fig. 4.3 in Rubinstein and Kroese, 2017)

18

Example of Dynamic Simulation Models: Transient and Steady State in Inventory System

- (s, S) inventory systems widely used in industry: When product inventory drops below quantity s, place order to bring inventory to up to quantity S
- Monthly costs to company C_i are cost for product, holding cost (warehouse, etc.), ordering costs, and backlog costs
- Plot below shows non-monotonic convergence to steady state mean monthly cost



(Source: Law 2014, Chap. 9)

19

Statistical Analysis of Dynamic Simulation Models

- Handle finite-horizon (transient) and steady-state phases differently
- For finite horizon analysis, run system multiple (N) independent times, and compute relevant measures of performance
  - E.g., If Y_i represents average cost over transient interval for run i, then form sample mean and variance of the Y_i across multiple simulation runs in usual manner
- For steady-state analysis, can use ***one*** simulation run (contrast with finite-horizon above)
  - Can estimate mean and variance across the one run
  - Sample mean taken across X_t in one run
  - Must account for dependence in $\{X_t\}$ in computing sample variance; see textbook, pp. 115–116
- Confidence bounds computed in standard general way:

$$\left(\hat{\ell} \pm z_{1-\alpha/2}\frac{\tilde{S}}{\sqrt{T}}\right)$$ ($\tilde{S}^2$ is sample variance estimate)

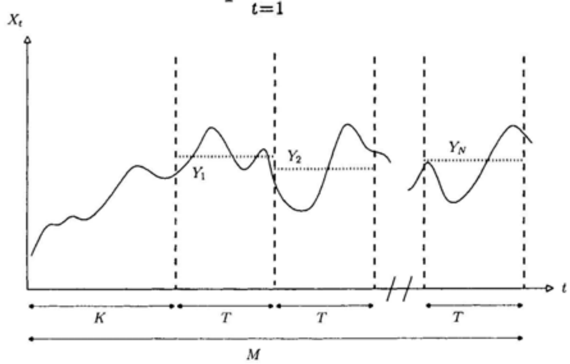
Note: This result does not always hold due to non-independence

20

Batch Means for Dynamic Systems

- Plot below shows partitioning for use in creating near-independent samples Y_i from dependent data

$$Y_i = \frac{1}{T}\sum_{t=1}^{T} X_{ti}, \quad i = 1, \dots, N$$



(Source: Fig. 4.4 in Rubinstein and Kroese, 2017)

- Analysis proceeds in usual way (t-tests, etc.) based on "near-i.i.d." data Y_i instead of original data X_t
  - But smaller sample size (fewer Y_i than X_t) due to dependence

21

Regenerative Systems

- Common issue in analysis of simulation output is need for creating independent or near-independent samples
- Regeneration is useful for addressing issue
- Regenerative systems have property of returning periodically to some particular probabilistic state; system effectively starts anew with each period
- Markov chains and queuing systems are common examples
  - Day-to-day traffic flow; inventory control; communications networks; etc.
- Advantage is that regeneration periods may be considered i.i.d. random processes
- Typical measure of performance ℓ has form:

$$\ell \equiv \frac{E(\text{accrued cost or reward over period})}{E(\text{length of period})}$$

22

Queuing System with Regeneration; Periods Begin with Arrivals 1, 3, 4, 7, 11, 16 (Example 14.2 in Spall, 2003)

Rationale for Performance Measure ℓ

- Let R = "reward" and τ be length of time for one regenerative cycle (both random variables in general)
- Then,

$$\ell = E[H(\boldsymbol{X})] = \frac{E(R)}{E(\tau)}$$

- ℓ represents mean long-run reward per unit time; specifically, can be shown:

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{E(R)}{E(\tau)} \quad \text{w/ prob. 1}$$

where $R(t)$ is cumulative reward to time t
- That is,

For "almost all" sample paths (i.e., w/ probability 1), long-run average reward per unit time is equal to the ratio of expected reward in one cycle over expected length of one cycle (i.e., equal to ℓ)

Care Needed in Estimators of ℓ for Regenerative Systems

- As before, seek estimate of ℓ
- By regenerative property samples of R and τ (R_i and τ_i) are each i.i.d. across i; R_i and τ_i may be dependent at given i
- Straightforward ratio estimator of ℓ is

$$\hat{\ell} = \hat{\ell}_N \equiv \frac{\text{Sample mean of cost/reward } R_i \text{ over } N \text{ periods}}{\text{Sample mean of length } \tau_i \text{ of } N \text{ periods}}$$

- Ratio estimator is ***biased*** in general (i.e., $E(\hat{\ell}_N) \neq \ell$)
  - Biasedness follows from relationship $E(1/\tau) \neq 1/E(\tau)$ for positive random variable τ, although analysis even more complicated due to dependence of R_i and τ_i at given i
- Some applications require unbiased estimator (e.g., stochastic gradient methods in Chap. 7 of textbook)
- Despite bias, $\hat{\ell}$ is convergent to ℓ (next slide): "good enough" some cases

25

Some Properties of Estimator of ℓ

- $\hat{\ell}$ may still be useful even though it is biased estimator of ℓ
- One way is from SLLN:

$$\bar{R} \to R \text{ w/ prob.1 } (\bar{R} \text{ is sample mean of } R_i)$$

$$\bar{\tau} \to \tau \text{ w/ prob.1 } (\bar{\tau} \text{ is sample mean of } \tau_i)$$

- Therefore:

$$\hat{\ell} = \hat{\ell}_N = \frac{\bar{R}}{\bar{\tau}} \to \ell \text{ w/ prob.1}$$

- Note that $\hat{\ell}$ may be near-unbiased or unbiased when R or τ are deterministic
  - When R is deterministic, can use Kantorovich inequality of probability theory to sometimes show near-unbiasedness
  - When τ is deterministic, then $E(\hat{\ell}_N) = E\left(\frac{\bar{R}}{\tau}\right) = \frac{1}{\tau} E(\bar{R}) = \ell$

26

References

- Law, A. M. (2014), *Simulation Modeling and Analysis* (5th ed.), McGraw-Hill, New York.
- Rubinstein, R. Y. and Kroese, D. P. (2017), *Simulation and the Monte Carlo Method* (3rd ed.), Wiley, New York. **[Course text for 553.633]**
- Spall, J. C. (2003), *Introduction to Stochastic Search and Optimization*, Wiley, Hoboken, NJ.