Chapter 11
Output Analysis for a Single Model

Banks, Carson, Nelson & Nicol

Discrete-Event System Simulation

Purpose

- Objective: Estimate system performance via simulation
- If θ is the system performance, the precision of the estimator $\hat{\theta}$ can be measured by:
  - The standard error of $\hat{\theta}$.
  - The width of a confidence interval (CI) for θ.
- Purpose of statistical analysis:
  - To estimate the standard error or CI .
  - To figure out the number of observations required to achieve desired error/CI.
- Potential issues to overcome:
  - Autocorrelation, e.g. time in system for successive jobs lack statistical independence.
  - Initial conditions, e.g. jobs in queue at the start of the simulation influence the time in system for subsequent jobs, for a while

Outline

- Distinguish the two types of simulation: transient vs. steady state.
- Illustrate the inherent variability in a stochastic discrete-event simulation.
- Cover the statistical estimation of performance measures.
- Discuss the analysis of transient simulations.
- Discuss the analysis of steady-state simulations.

Type of Simulations

- Terminating verses non-terminating simulations
- Terminating simulation:
  - Runs for some duration of time T_E, where E is a specified event that stops the simulation.
  - Starts at time *0* under well-specified initial conditions.
  - Ends at the stopping time T_E.
  - Bank example: Opens at 8:30 am (time *0*) with no customers present and *8* of the *11* teller working (initial conditions), and closes at 4:30 pm (Time $T_E = 480$ minutes).
  - The simulation analyst chooses to consider it a terminating system because the object of interest is one day's operation.

Type of Simulations

- Non-terminating simulation:
  - Runs continuously, or at least over a very long period of time.
  - Examples: assembly lines that shut down infrequently, telephone systems, hospital emergency rooms.
  - Initial conditions defined by the analyst.
  - Runs for some analyst-specified period of time T_E.
  - Study the steady-state (long-run) properties of the system, properties that are not influenced by the initial conditions of the model.
- Whether a simulation is considered to be terminating or non-terminating depends on both
  - The objectives of the simulation study and
  - The nature of the system.

Stochastic Nature of Output Data

- Model output consist of one or more random variables (r. v.) because the model is an input-output transformation and the input variables are r.v.'s.
- M/G/1 queueing example:
  - Poisson arrival rate = *0.1* per minute;
    service time ~ $N(\mu = 9.5, \sigma = 1.75)$.
  - System performance: long-run mean queue length, $L_Q(t)$.
  - Suppose we run a single simulation for a total of 5,000 minutes
    - Divide the time interval [*0, 5000*) into *5* equal subintervals of *1000* minutes.
    - Average number of customers in queue from time *(j-1)1000* to *j(1000)* is Y_j.

Stochastic Nature of Output Data

- M/G/1 queueing example (cont.):
  - Batched average queue length for 3 independent replications:

Batching Interval (minutes)	Batch, j	Replication		
		1, Y_{1j}	2, Y_{2j}	3, Y_{3j}
[0, 1000)	1	3.61	2.91	7.67
[1000, 2000)	2	3.21	9.00	19.53
[2000, 3000)	3	2.18	16.15	20.36
[3000, 4000)	4	6.92	24.53	8.11
[4000, 5000)	5	2.82	25.19	12.62
[0, 5000)		3.75	15.56	13.66

  - Inherent variability in stochastic simulation both within a single replication and across different replications.
  - The average across 3 replications, $\bar{Y}_{1.}, \bar{Y}_{2.}, \bar{Y}_{3.}$, can be regarded as independent observations, but averages within a replication, Y_{11}, …, Y_{15}, are not.

Measures of performance

- Consider the estimation of a performance parameter, θ (or ϕ), of a simulated system.
  - Discrete time data: $[Y_1, Y_2, \ldots, Y_n]$, with ordinary mean: θ
  - Continuous-time data: $\{Y(t), 0 \le t \le T_E\}$ with time-weighted mean: ϕ
- Point estimation for discrete time data.
  - The point estimator:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$E(\hat{\theta}) = \theta$$

Desired

    - Is unbiased if its expected value is θ, that is if:
    - Is biased if: $E(\hat{\theta}) \neq \theta$

Confidence-Interval Estimation

[Performance Measures]

- Central Limit Theorem.

If $X_1, X_2, \ldots, X_n$ are i.i.d. samples from a distribution with mean θ and variance ρ^2, then as $n \to \infty$ the random variable $\hat{\theta} = (1/n)(\sum_{i=1}^{n} X_i)$ is Gaussian with mean θ and variance ρ^2/n.

- Key points: sample mean is unbiased estimator
- Variance describes "how close" the sample mean is likely to be to the true unknown μ
- i.i.d. assumption satisfied by using samples *across* replications

Confidence-Interval Estimation

[Performance Measures]

- Suppose the model is the normal distribution with mean θ, variance σ^2 (both unknown) (e.g., sample mean)
  - Let Y_i be the average time in system for i^{th} replication of the simulation (its mathematical expectation is θ).
  - Average time in system will vary from epoch to epoch, but over the long-run the average of the averages will be close to θ.
  - Sample variance across R replications:

$$S^2 = \frac{1}{R-1}\sum_{i=1}^{R}(Y_{i.} - Y_{..})^2$$

Sample mean of sample means

Sample mean from replication

Confidence-Interval Estimation

- Confidence Interval (CI):
  - A measure of error.
  - Where $Y_{i\cdot}$ are normally distributed.

$$\bar{Y}_{\cdot\cdot} \pm t_{\alpha/2,R-1} \frac{S}{\sqrt{R}}$$

  - We cannot know for certain how far $\bar{Y}_{\cdot\cdot}$ is from θ but CI attempts to bound that error.
  - A CI, such as 95%, tells us how much we can trust the interval to actually bound the error between $\bar{Y}_{\cdot\cdot}$ and θ.
  - The more replications we make, the less error there is in $\bar{Y}_{\cdot\cdot}$ (converging to *0* as *R* goes to infinity).

Confidence-Interval Estimation

- Confidence Interval (CI) at X% confidence:
  - Random interval constructed from data that X% of the time will contain the true θ
  - Where $Y_{i\cdot}$ are normally distributed.

$$\bar{Y}_{\cdot\cdot} \pm t_{\alpha/2,R-1} \frac{S}{\sqrt{R}}$$

  - Bottom line—if sample mean is in (1-X)% tail of its distribution, then θ not in interval
  - The more replications we make, the less error there is in (converging to *0* as *R* goes to infinity).

C.I. with Specified Precision

[Terminating Simulations]

- The half-length H of a $100(1 - \alpha)\%$ confidence interval for a mean θ, based on the t distribution, is given by:

$$H = t_{\alpha/2, R-1} \frac{S_0}{\sqrt{R}}$$

R is the # of replications

S_0^2 is the sample variance

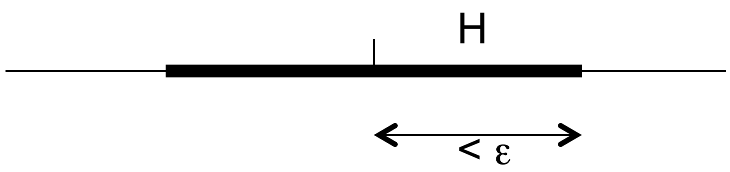
- Suppose that an error criterion ε is specified: with probability $1 - \alpha$, a sufficiently large sample size should satisfy:

$$P\left(\left|\bar{Y}_{\cdot\cdot} - \theta\right| < \varepsilon\right) \geq 1 - \alpha$$

C.I. with Specified Precision

[Terminating Simulations]

This is a statement about half-width of C.I.



$$t_{\alpha/2,R-1} \frac{S_0}{\sqrt{R}} < \epsilon$$

Satisfied when $$\left(\frac{z_{\alpha/2} \cdot S_0}{\epsilon} \right)^{1/2} \leq R$$

Because N(0,1) critical value $z_{\alpha/2} \leq t_{\alpha/2,R-1}$

C.I. with Specified Precision

[Terminating Simulations]

- Assume that an initial sample of size R_0 (independent) replications has been observed.
- Obtain an initial estimate S_0^2 of the population variance σ^2.
- Then, choose sample size R such that $R \geq R_0$:
  - Since $t_{\alpha/2, R-1} \geq z_{\alpha/2}$, an initial estimate of R:

$$R \geq \left(\frac{z_{\alpha/2} S_0}{\varepsilon} \right)^2, \quad z_{\alpha/2} \text{ is the standard normal distribution.}$$

  - R is the smallest integer satisfying $R \geq R_0$ and $R \geq \left(\frac{t_{\alpha/2, R-1} S_0}{\varepsilon} \right)^2$
- Collect $R - R_0$ additional observations.
- The $100(1-\alpha)\%$ C.I. for θ:

$$\bar{Y}_{..} \pm t_{\alpha/2, R-1} \frac{S_0}{\sqrt{R}}$$

C.I. with Specified Precision

[Terminating Simulations]

- Multi-server example: estimate a CPU's utilization ρ over the first 2 hours of the workday.
  - Initial sample of size $R_0 = 4$ is taken and an initial estimate of the population variance is $S_0^2 = (0.072)^2 = 0.00518$.
  - The error criterion is $\varepsilon = 0.04$ and confidence coefficient is $1\text{-}\alpha = 0.95$, hence, the final sample size must be at least:

$$\left(\frac{z_{0.025} S_0}{\varepsilon}\right)^2 = \frac{1.96^2 * 0.00518}{0.04^2} = 12.14$$

  - For the final sample size:

R	13	14	15
$t_{0.025,\, R-1}$	2.18	2.16	2.14
$\left(t_{\alpha/2, R-1} S_0 / \varepsilon\right)^2$	15.39	15.1	14.83

  - $R = 15$ is the smallest integer satisfying constraint (15>14.83) , so $R - R_0 = 11$ additional replications are needed.
  - After obtaining additional outputs, half-width should be checked to ensure constraint met given better estimate of variance

Output Analysis for Steady-State Simulation

- Consider a single run of a simulation model to estimate a steady-state or long-run characteristics of the system.
  - The single run produces observations Y_1, Y_2, ... (generally the samples of an autocorrelated time series).
  - Performance measure:

$$\theta = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i, \quad \text{for discrete measure}$$ (with probability 1)

$$\phi = \lim_{T_E\to\infty} \frac{1}{T_E} \int_0^{T_E} Y(t)\,dt, \quad \text{for continuous measure}$$ (with probability 1)

    - Independent of the initial conditions.

Output Analysis for Steady-State Simulation

- The sample size is a design choice, with several considerations in mind:
  - Any bias in the point estimator that is due to artificial or arbitrary initial conditions (bias can be severe if run length is too short).
  - Desired precision of the point estimator.
  - Budget constraints on computer resources.
- Notation: the estimation of θ from a discrete-time output process.
  - One replication (or run), the output data: $Y_1, Y_2, Y_3, \ldots$
  - With several replications, the output data for replication r: $Y_{r1}, Y_{r2}, Y_{r3}, \ldots$

Initialization Bias [Steady-State Simulations]

- Methods to reduce the point-estimator bias caused by using artificial and unrealistic initial conditions:
  - Intelligent initialization.
  - Divide simulation into an initialization phase and data-collection phase.
- Intelligent initialization
  - Initialize the simulation in a state that is more representative of long-run conditions.
  - If the system exists, collect data on it and use these data to specify more nearly typical initial conditions.
  - If the system can be simplified enough to make it mathematically solvable, e.g. queueing models, solve the simplified model to find long-run expected or most likely conditions, use that to initialize the simulation.

Initialization Bias [Steady-State Simulations]

- Divide each simulation into two phases:
  - An initialization phase, from time *0* to time T_0.
  - A data-collection phase, from T_0 to the stopping time T_0+T_E.
  - The choice of T_0 is important:
    - After T_0, system should be more nearly representative of steady-state behavior.
  - System has reached steady state: the probability distribution of the system state is close to the steady-state probability distribution (bias of response variable is negligible).

Initialization Bias [Steady-State Simulations]

- M/G/1 queueing example: A total of 10 independent replications were made.
  - Each replication beginning in the empty and idle state.
  - Simulation run length on each replication was T_0+T_E = 15,000 minutes.
  - Response variable: queue length, $L_Q(t,r)$ (at time *t* of the *r*th replication).
  - Batching intervals of 1,000 minutes, batch means
- Ensemble averages:
  - To identify trend in the data due to initialization bias
  - The average corresponding batch means *across* replications:

$$\bar{Y}_{.j} = \frac{1}{R}\sum_{r=1}^{R} Y_{rj}$$

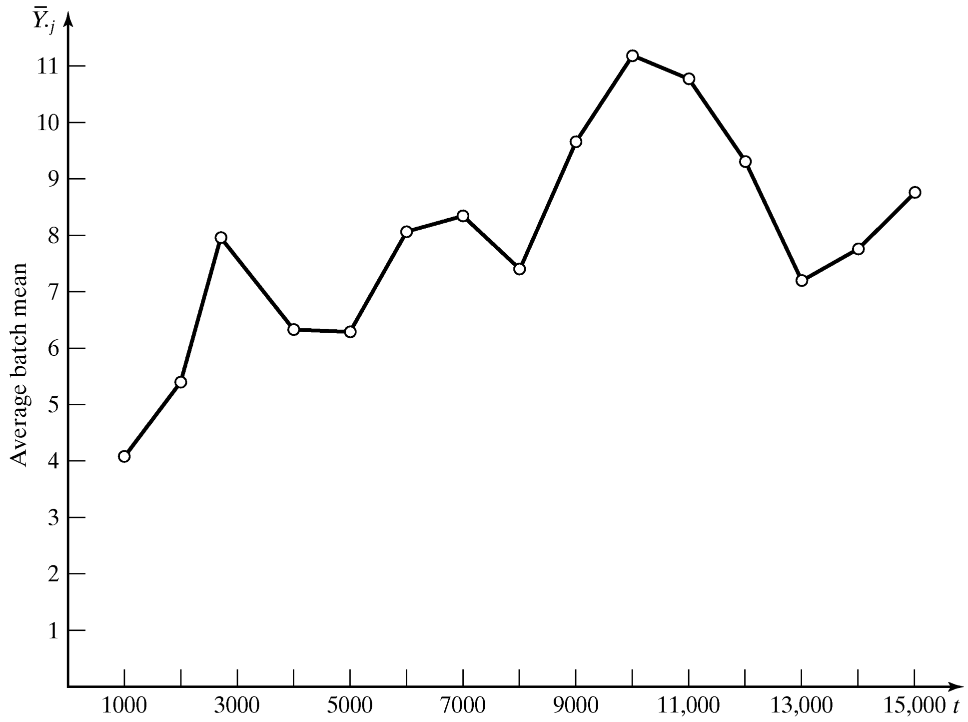
R replications

  - The preferred method to determine deletion point.

Initialization Bias

[Steady-State Simulations]

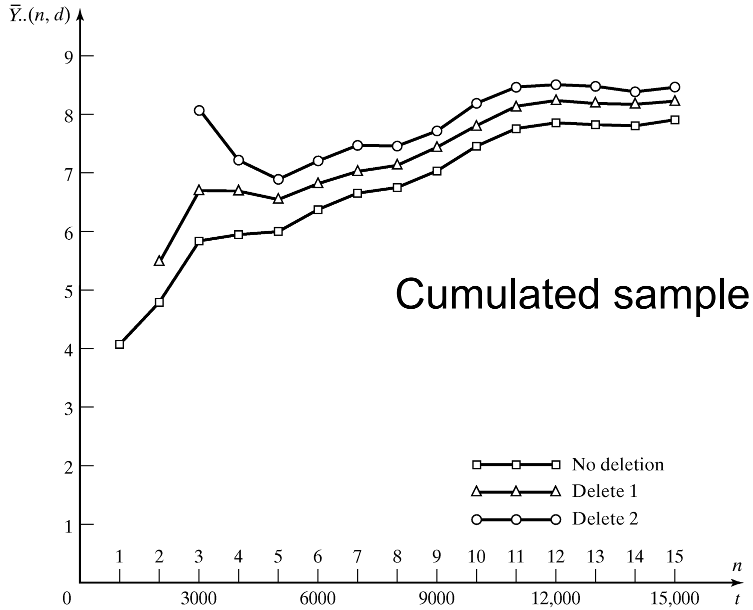
- A plot of the ensemble averages, $\bar{Y}..(n,d)$, versus *1000j*, for *j = 1,2, …,15*.



  - Illustrates the downward bias of the initial observations.

Initialization Bias [Steady-State Simulations]

- *Cumulative* average sample mean (after deleting *d* observations):
$$\bar{Y}_{..}(n,d) = \frac{1}{n-d}\sum_{j=d+1}^{n}\bar{Y}_{.j}$$
  - Not recommended to determine the initialization phase.



Cumulated sample mean by given point in time

  - It is apparent that downward bias is present and this bias can be reduced by deletion of one or more observations.

Initialization Bias [Steady-State Simulations]

- No widely accepted, objective and proven technique to guide how much data to delete to reduce initialization bias to a negligible level.
- Plots can, at times, be misleading but they are still recommended.
  - Ensemble averages reveal a smoother and more precise trend as the # of replications, R, increases.
  - Ensemble averages can be smoothed further by plotting a moving average.
  - Cumulative average becomes less variable as more data are averaged.
  - The more correlation present, the longer it takes for $\bar{Y}_{\cdot j}$ to approach steady state.
  - Different performance measures could approach steady state at different rates.

Batch Means for Interval Estimation

[Steady-State Simulations]

- Using a single, long replication:
  - Problem: data are dependent so the usual estimator is biased.
  - Solution: batch means.
- Batch means: divide the output data from *1* replication (after appropriate deletion) into a few large batches and then treat the means of these batches as if they were independent.
- A continuous-time process, $\{Y(t), T_0 \leq t \leq T_0+T_E\}$:
  - k batches of size $m = T_E/k$, batch means:

$$\bar{Y}_j = \frac{1}{m}\int_{(j-1)m}^{jm} Y(t+T_0)\,dt$$

- A discrete-time process, $\{Y_i,\ i = d+1, d+2, \ldots, n\}$:
  - k batches of size $m = (n-d)/k$, batch means: $\bar{Y}_j = \frac{1}{m}\sum_{i=(j-1)m+1}^{jm} Y_{i+d}$

Batch Means for Interval Estimation

[Steady-State Simulations]

$$\underbrace{Y_1, \ldots, Y_d}_{\text{deleted}}, \underbrace{Y_{d+1}, \ldots, Y_{d+m}}_{\bar{Y}_1}, \underbrace{Y_{d+m+1}, \ldots, Y_{d+2m}}_{\bar{Y}_2}, \;\ldots\; , \underbrace{Y_{d+(k-1)m+1}, \ldots, Y_{d+km}}_{\bar{Y}_k}$$

- Starting either with continuous-time or discrete-time data, the variance of the sample mean is estimated by:

$$\frac{S^2}{k} = \frac{1}{k} \sum_{j=1}^{k} \frac{\left(\bar{Y}_j - \bar{Y}\right)^2}{k-1} = \sum_{j=1}^{k} \frac{\bar{Y}_j^2 - k\bar{Y}^2}{k(k-1)}$$

- If the batch size is sufficiently large, successive batch means will be approximately independent, and the variance estimator will be approximately unbiased.
- No widely accepted and relatively simple method for choosing an acceptable batch size *m* (see text for a suggested approach). Some simulation software does it automatically.

Summary

- Stochastic discrete-event simulation is a statistical experiment.
  - Purpose of statistical experiment: obtain estimates of the performance measures of the system.
  - Purpose of statistical analysis: acquire some assurance that these estimates are sufficiently precise.
- Distinguish: terminating simulations and steady-state simulations.
- Steady-state output data are more difficult to analyze
  - Decisions: initial conditions and run length
  - Possible solutions to bias: deletion of data and increasing run length
- Statistical precision of point estimators are estimated by standard-error or confidence interval
- Method of independent replications was emphasized.