Output Analysis for a Single Model

Output Analysis

Radu Trîmbițaș

Faculty of Math and CS - UBB

1st Semester 2010-2011

Purpose

- Objective: Estimate system performance via simulation
- If θ is the system performance, the precision of the estimator $\widehat{\theta}$ can be measured by:
  - The standard error of $\widehat{\theta}$
  - The width of a confidence interval (CI) for θ.
- Purpose of statistical analysis:
  - To estimate the standard error or confidence interval.
  - To figure out the **number of observations** required to achieve a desired error or confidence interval.
- Potential issues to overcome:
  - Autocorrelation, e.g. inventory cost for subsequent weeks lack statistical independence.
  - Initial conditions, e.g. inventory on hand and number of backorders at time 0 would most likely influence the performance of week 1.

Types of Simulation

- Distinguish the two types of simulation:
  - transient vs.
  - steady state - a simulation whose objective is to study long-run or steady-state behavior of nonterminating system
- Illustrate the inherent variability in a stochastic discrete-event simulation.
- Cover the statistical estimation of performance measures.
- Discusses the analysis of transient simulations.
- Discusses the analysis of steady-state simulations.

Types of Simulation I

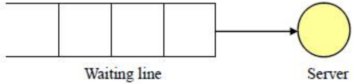
- A distinction is made between *terminating* (or transient) versus *non-terminating simulations*
- **Terminating simulation:**
  - Runs for some duration of time T_E, where E is a specified event or set of events that stops the simulation.
  - Starts at time 0 under well-specified initial conditions.
  - 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)
    - 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.
  - T_E may be known from the beginning or it may not

Types of Simulation II

- **Non-terminating simulation:**
  - Runs continuously, or at least over a very long period of time.
  - Examples: assembly lines that shut down infrequently, hospital emergency rooms, telephone systems, network of routers, Internet.
  - 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 nonterminating 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 because the model is an input-output transformation and the input variables are random variables.
- M/G/1 queueing example:
  - Poisson arrival rate = 0.1 per minute and 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 5000 minutes
    - Divide the time interval $[0, 5000)$ into 5 equal subintervals of 1000 minutes.
    - Average number of customers in queue from time $(j-1) \cdot 1000$ to $j \cdot (1000)$ is Y_j.



$$L_Q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{\rho^2}{1 - \rho}$$

Stochastic Nature of Output Data II

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

Batching Interval (minutes)	Batch, j	Replication 1, Y_{1j}	Replication 2, Y_{2j}	Replication 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, $\overline{Y}_1$, $\overline{Y}_2$, $\overline{Y}_3$, can be regarded as independent observations, but averages within a replication, $Y_{11}, \ldots, Y_{15}$, 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: $\{Y(t), 0 \leq t \leq T_E\}$ with time-weighted mean: ϕ
- **Point estimation for discrete time data**.
  - The point estimator:

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

  - unbiased if $E(\widehat{\theta}) = \theta$ (Desired)
  - biased if $E(\widehat{\theta}) \neq \theta$; $E(\widehat{\theta}) - \theta$ is called bias of $\widehat{\theta}$

Point estimator I

- **Point estimation for continuous-time data**. The point estimator:

$$\widehat{\phi} = \frac{1}{T_E} \int_0^{T_E} Y(t)\mathrm{d}t$$

  - Is biased in general where: $E(\widehat{\phi}) \neq \phi$.
  - An unbiased or a low-bias estimator is desired.
- Usually, system performance measures can be put into the common framework of θ or ϕ:
- Example: The proportion of days on which sales are lost through an outof-stock situation, let:

$$Y(i) = \begin{cases} 1, & \text{if out of stock on day } i \\ 0, & \text{otherwise} \end{cases}$$

Point estimator II

- Performance measure that does not fit – quantile or percentile:

$$P(Y \geq \theta) = p$$

  - Estimating quantiles: the inverse of the problem of estimating a proportion or probability.
  - Consider a histogram of the observed values Y:
    - Find such that $100p\%$ of the histogram is to the left of (smaller than) $\widehat{\theta}$
  - A widely used performance measure is the median, which is the 0.5 quantile or 50-th percentile.

Confidence-Interval Estimation I

- Suppose $X_1, X_2, \ldots, X_n$ are independent sample from a normally distributed population with mean μ and variance σ^2.
- If the sample mean and sample variance are

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2,$$

  then

$$T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$$

  has Student's t-distribution with $n-1$ degrees of freedom
- If c is the p-th quantile of this distribution, then $P(-c < T < c) = p$

Confidence-Interval Estimation II

- Consequently

$$P\left(\overline{X} - c\frac{S}{\sqrt{n}} < \mu < \overline{X} + c\frac{S}{\sqrt{n}}\right) = p$$

- Confidence Interval (CI):
- A measure of error, where Y_i are normally distributed

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

- We cannot know for certain how far $\overline{Y}$ 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 $\overline{Y}$ and θ
- The more replications we make, the less error there is in $\overline{Y}$ (converging to 0 as R goes to infinity).

Confidence-Interval Estimation III

- Prediction Interval (PI):
  - A measure of risk.
  - A good guess for the average cycle time on a particular day is our estimator but it is unlikely to be exactly right.
  - PI is designed to be wide enough to contain the actual average cycle time on any particular day with high probability.
  - Normal-theory prediction interval:

$$\overline{Y} \pm t_{\alpha/2,R-1} S \sqrt{1 + \frac{1}{R}}$$

  - The length of PI will not go to 0 as R increases because we can never simulate away risk.
  - Prediction Intervals limit is $\theta \pm z_{\alpha/2}\sigma$

Output Analysis for Terminating Simulations

- A terminating simulation: runs over a simulated time interval $[0, T_E]$.
- A common goal is to estimate:

$$\theta = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right), \qquad \text{for discrete output}$$

$$\phi = E\left(\frac{1}{T_E}\int_0^{T_E} Y(t)dt\right), \text{ for continuous output } Y(t)$$

- In general, independent replications are used, each run using a *different random number stream* and independently chosen initial conditions.

Statistical Background I

- Important to distinguish within-replication data from across-replication data.
- For example, simulation of a manufacturing system
  - Two performance measures of that system: cycle time for parts and work in process (WIP).
  - Let Y_{ij} be the cycle time for the j-th part produced in the i-th replication.
  - Across-replication data are formed by summarizing within-replication data .

Within-Replication Data				Across-Rep. Data
Y_{11}	Y_{12}	$\cdots$	Y_{1n_1}	$\bar{Y}_{1\cdot}, S_1^2, H_1$
Y_{21}	Y_{22}	$\cdots$	Y_{2n_2}	$\bar{Y}_{2\cdot}, S_2^2, H_2$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
Y_{R1}	Y_{R2}	$\cdots$	Y_{Rn_R}	$\bar{Y}_{R\cdot}, S_R^2, H_R$
				$\bar{Y}_{\cdot\cdot}, S^2, H$

Statistical Background II

- Across Replication:
  - For example: the daily cycle time averages (discrete time data)

$$\overline{Y} = \frac{1}{R}\sum_{i=1}^{R} Y_i, \qquad \text{the average}$$

$$S^2 = \frac{1}{R-1}\sum_{i=1}^{R}\left(Y_{i.} - \overline{Y}_{..}\right), \qquad \text{the sample variance}$$

$$H = t_{\alpha/2,R-1}\frac{S}{\sqrt{R}}, \qquad \text{the CI halfwidth}$$

- Within replication:

Statistical Background III

- For example: the WIP (a continuous time data)

$$\overline{Y}_{i.} = \frac{1}{T_{E_i}} \int_0^{T_{E_i}} Y_i(t)dt, \qquad \text{the average}$$

$$S_i^2 \frac{1}{T_{E_i}} \int_0^{T_{E_i}} \left(Y_i(t) - \overline{Y}_i\right)^2 dt, \qquad \text{the sample variance}$$

- Overall sample average, $\overline{Y}_{..}$, and the interval replication sample averages, $\overline{Y}_{i.}$, are always unbiased estimators of the expected daily average cycle time or daily average WIP.
- Across-replication data are independent (different random numbers) and identically distributed (same model), but within-replication data do not have these properties.

Confidence Intervals with Specified Precision I

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

$$H = t_{\alpha/2,R-1}\frac{S}{\sqrt{S}} \quad (1)$$

  R number of replications, S^2 sample variance

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

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

- Assume that an initial sample of size R0 (independent) replications has been observed.

Confidence Intervals with Specified Precision II

- Obtain an initial estimate S_0^2 of the population variance σ^2

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

- Then, choose sample size R such that $R \geq R_0$
- Solving for R

$$R \geq \left(\frac{t_{\alpha/2,R-1} S_0}{\varepsilon} \right)^2$$

- Call Center Example: estimate the agent'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 $S_0^2 = (0.072)^2 = 0.00518$.

Confidence Intervals with Specified Precision III

- The error criterion is $\varepsilon = 0.04$ and confidence coefficient is $1 - \alpha = 0.95$, hence, the final sample size must be at least:

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

- For the final sample size:

R	13	14	15
$t_{0.025,R-1}$	2.18	2.16	2.14
$(t_{\alpha/2,R-1}S_0)^2$	15.39	15.10	14.83

- $R = 15$ is the smallest integer satisfying the error criterion, so $R - R_0 = 11$ additional replications are needed.
- After obtaining additional outputs, half-width checked.

Quantiles I

- Here, a proportion or probability is treated as a special case of a mean.
- When the number of independent replications $Y_1, \ldots, Y_R$ is large enough that $t_{\alpha/2,n-1} \approx z_{\alpha/2}$, the confidence interval for a probability p is often written as:

$$\widehat{p} \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{R-1}}$$

- Quantile determination is the inverse of the probability estimation problem: find θ such that $P(Y \leq \theta) = p$, p dat
- The best way is to sort the outputs and use the $(R^* p)$-th smallest value, i.e., find θ such that $100p\%$ of the data in a histogram of Y is to the left of θ.

Quantiles II

- **Example**: If we have $R = 10$ replications and we want the $p = 0.8$ quantile, first sort, then estimate θ by the $(10)(0.8) = 8$-th smallest value (round if necessary).

5.6	←sorted data
7.1	
8.8	
8.9	
9.5	
9.7	
10.1	
12.2	←this is our point estimates
12.5	
12.9	

- Confidence Interval of Quantiles: An approximate $(1-\alpha)100\%$ confidence interval for θ can be obtained by finding two values θ_l and θ_u.

Quantiles III

- θ_l cuts off $100p_l\%$ of the histogram (the $R \cdot p_l$ smallest value of the sorted data).
- θ_u cuts off $100p_u\%$ of the histogram (the $R \cdot p_u$ smallest value of the sorted data).

$$p_l = p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{R-1}}$$

$$p_u = p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{R-1}}$$

- **Example**: Suppose $R = 1000$ replications, to estimate the $p = 0.8$ quantile with a 95% confidence interval.
- First, sort the data from smallest to largest.
- Then estimate of θ by the $(1000)(0.8) = 800$-th smallest value, and the point estimate is 212.03.

Quantiles IV

- And find the confidence interval:

$$p_l = 0.8 - \eta = 0.78$$
$$p_u = 0.8 + \eta = 0.82$$
$$\eta = 1.96\sqrt{\frac{0.8(1-0.8)}{1000-1}}$$

The CI is the 780th and the 820th smallest value

- The point estimate is 212.03
- The 95% CI is [188.96, 256.79]

A portion of the 1000 sorted values:

Output	Rank
180.92	779
188.96	**780**
190.55	781
208.58	799
212.03	**800**
216.99	801
250.32	819
256.79	**820**
256.99	821

Estimating Probabilities and Quantiles from Summary Data I

- To compute standard deviation from H

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

- Estimating a probability or quantile from summary data is more difficult: given sample mean and CI halfwidth, using normal theory

$$P\left(\overline{Y}_{i.} \leq c\right) \approx P\left(Z \leq \frac{c - \overline{Y}_{..}}{S}\right)$$

$$\widehat{\theta} \approx \overline{Y}_{..} + z_p S$$

Estimating Probabilities and Quantiles from Summary Data II

- **Example:** For 25 replications and a 90% CI for the daily average WIP (work in process) 218 ± 32, we wish $P(\overline{Y}_{i.} \leq 350)$ and the 85-th percentile

$$S = \frac{H\sqrt{R}}{t_{0.05,24}} = 93$$

$$P(\overline{Y}_{i.} \leq 350) = P\left(Z \leq \frac{350 - 218}{93}\right) = 0.92$$

$$\widehat{\theta} \approx \overline{Y}_{..} + z_{0.85}S = 218 + 1.04 \cdot 93 = 315$$

- the quality depends on the validity of normality assumption

Output Analysis for Steady-State Simulations I

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

$$\theta = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i, \qquad \text{for discrete measure}$$

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

independent of initial conditions, both with probability 1

- The sample size is a design choice, with several considerations in mind:

Output Analysis for Steady-State Simulations II

  - 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.

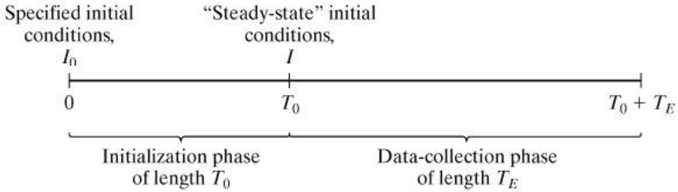
- Notation: the estimation of θ from a discrete-time output process.
  - One replication (or run), the output data: $Y_1, Y_2, Y_3, \ldots$
  - With several replications, the output data for replication r: $Y_{r_1}, Y_{r_2}, Y_{r_3}, \ldots$

Initialization Bias I

- 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.
- Divide each simulation into two phases:
  - An initialization phase, from time 0 to time T_0.

Initialization Bias II

- 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).



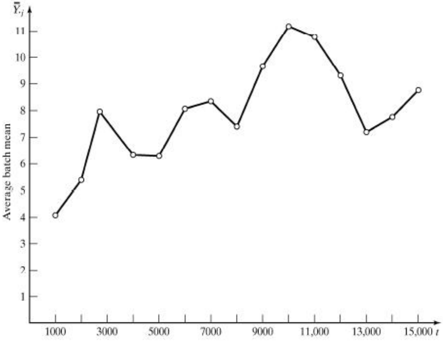
Initialization Bias III

- 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 $T_0 + T_E = 15000$ minutes.
  - Response variable: queue length, $L_Q(t, r)$ (at time t of the r-th replication).
  - Batching intervals of 1000 minutes, batch means
- Ensemble averages:
  - To identify trend in the data due to initialization bias
  - The average corresponding batch means across replications (R replications):

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

- A plot of the ensemble averages, $\overline{Y}_{.j}(n, d)$, versus $1000j$, for $m = 1, 2, \ldots, 15$

Initialization Bias IV

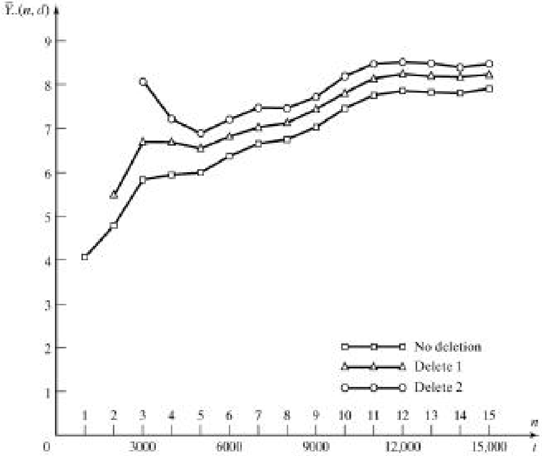


- Cumulative average sample mean (after deleting d observations):

$$\overline{Y}_{..}(n,d) = \frac{1}{n-d}\sum_{j=d+1}^{n} \overline{Y}_{.j}$$

Initialization Bias V

- Not recommended to determine the initialization phase.



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

Initialization Bias VI

- 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 number 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 to approach steady state $\overline{Y}_{.j}$
  - Different performance measures could approach steady state at different rates.

Error Estimation I

- If $\{Y_1, \ldots, Y_n\}$ are not statistically independent, then S^2/n is a biased estimator of the true variance.
- Almost always the case when $\{Y_1, \ldots, Y_n\}$ is a sequence of output observations from within a single replication (autocorrelated sequence, time-series).
- Suppose the point estimator $\widehat{\theta}$ is the sample mean

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Variance of $\overline{Y}$ is very hard to estimate.
- For systems with steady state, produce an output process that is approximately covariance stationary (after passing the transient phase).

Error Estimation II

- The covariance between two random variables in the time series depends only on the lag, i.e. the number of observations between them.

- For a covariance stationary time series, $\{Y_1, \ldots, Y_n\}$:
  - Lag-k autocovariance is: $\gamma_k = cov(Y_1, Y_{1+k}) = cov(Y_i, Y_{i+k})$
  - Lag-k autocorrelation is: $\rho_k = \frac{\gamma_k}{\sigma^2}$, $-1 \le \rho_k \le 1$

- If a time series is covariance stationary, then the variance of $\overline{Y}$ is:

$$V(\overline{Y}) = \frac{\sigma^2}{n} \left[\underbrace{1 + 2\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\rho_k}_{c} \right]$$

Error Estimation III

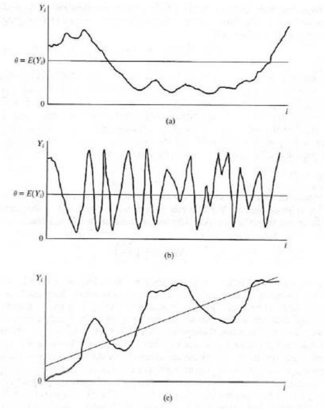
- The expected value of the variance estimator is:

$$E\left(\frac{S^2}{n}\right) = B \cdot V(\overline{Y}), \text{ where } B = \frac{\frac{n}{c} - 1}{n - 1}$$

(a) $\rho_k > 0$ for most k. Stationary time series (Y_i) exhibiting positive autocorrelation. Serie slowly drifts above and then below the mean.

(b) $\rho_k < 0$ for most k. Stationary time series (Y_i) exhibiting positive autocorrelation.

(c) Nonstationary time series with an upward trend



Error Estimation IV

- The expected value of the variance estimator is:

$$E\left(\frac{S^2}{n}\right) = B \cdot V(\overline{Y}),$$

where $B = \frac{n/c-1}{n-1}$, and $V(\overline{Y})$ is the variance of $\overline{Y}$.

  - If (Y_i) are independent, then S^2/n is an unbiased estimator of $V(\overline{Y})$
  - If the autocorrelation ρ_k are primarily positive, then S^2/n is biased low as an estimator of $V(\overline{Y})$.
  - If the autocorrelation ρ_k are primarily negative, then S^2/n is biased high as an estimator of $V(\overline{Y})$.

Replication Method I

- Use to estimate point-estimator variability and to construct a confidence interval.
- Approach: make R replications, initializing and deleting from each one the same way.
- Important to do a thorough job of investigating the initial-condition bias:
  - Bias is not affected by the number of replications, instead, it is affected only by deleting more data (i.e., increasing T_0) or extending the length of each run (i.e. increasing T_E).
- Basic raw output data $\{Y_{rj}, r = 1, \ldots, R,\ j = 1, \ldots, n\}$ is derived by:
  - Individual observation from within replication r.
  - Batch mean from within replication r of some number of discrete-time observations.

Replication Method II

- Batch mean of a continuous-time process over time interval j.

- Each replication is regarded as a single sample for estimating θ. For replication r:

$$\overline{Y}_{r\cdot}(n,d) = \frac{1}{n-d}\sum_{j=d+1}^{n} Y_{rj}$$

- The overall point estimator:

$$\overline{Y}_{\cdot\cdot}(n,d) = \frac{1}{R}\sum_{r=1}^{r} \overline{Y}_{r\cdot}(n,d) \text{ and } E\left[\overline{Y}_{\cdot\cdot}(n,d)\right] = \theta_{nd}$$

- If d and n are chosen sufficiently large:
  - $\theta_{nd} \approx \theta$
  - $\overline{Y}_{\cdot\cdot}(n,d)$ is an approximately unbiased estimator of θ.

Replication Method III

- To estimate the standard error of $\overline{Y}_{..}$, compute the sample variance and standard error:

$$S^2 = \frac{1}{R-1}\sum_{r=1}^{R}(\overline{Y}_{r.} - \overline{Y}_{..})^2 = \frac{1}{R-1}\left(\sum_{r=1}^{R}\overline{Y}_{r.}^{\,2} - R\overline{Y}_{..}^{\,2}\right) \quad \text{and} \quad s.e.(\overline{Y}_{..}) = \frac{S}{\sqrt{R}}$$

Mean of the undeleted observations from the r-th replication.

Mean of $\overline{Y}_{1.}(n,d),\ldots,\overline{Y}_{R.}(n,d)$

Standard error

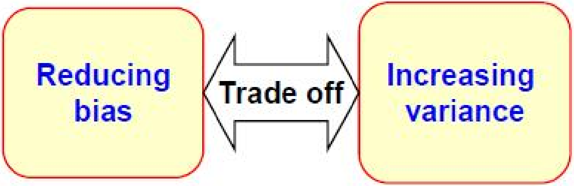
- Length of each replication (n) beyond deletion point (d):

$$(n-d) > 10d \text{ or } T_E > 10T_0$$

- Number of replications (R) should be as many as time permits, up to about 25 replications.

Replication Method IV

- For a fixed total sample size (n), as fewer data are deleted ($\downarrow d$):
  - Confidence interval shifts: greater bias.
  - Standard error of $\overline{Y}_{..}(n, d)$ decreases: decrease variance.



M/G/1 queueing example I

- Suppose $R = 10$, each of length $T_E = 15000$ minutes, starting at time 0 in the empty and idle state, initialized for $T_0 = 2000$ minutes before data collection begins.
- Each batch means is the average number of customers in queue for a 1000-minute interval.
- The 1-st two batch means are deleted ($d = 2$).
- estimator and standard error are:

$$\overline{Y}_{..}(15,2) = 8.43, \qquad s.e.\left(\overline{Y}_{..}(15,2)\right) = 1.59$$

- The 95% CI for long-run mean queue length is:

$$\overline{Y}_{..} - t_{\alpha/2,R-1}\frac{S}{\sqrt{R}} \le \theta \le \overline{Y}_{..} + t_{\alpha/2,R-1}\frac{S}{\sqrt{R}}$$
$$8.43 - 2.26(1.59) \le L_Q \le 8.43 + 2.26(1.59)$$

M/G/1 queueing example II

- A high degree of confidence that the long-run mean queue length is between 4.84 and 12.02 (if d and n are "large" enough).

Replication r	Sample Mean for Replication r		
	No deletion $\overline{Y}_{r.}(15,0)$	Delete 1 $\overline{Y}_{r.}(15,1)$	Delete 2 $\overline{Y}_{r.}(15,2)$
1	3.27	3.24	3.25
2	16.25	17.20	17.83
3	15.19	15.72	15.43
4	7.24	7.28	7.71
5	2.93	2.98	3.11
6	4.56	4.82	4.91
7	8.44	8.96	9.45
8	5.06	5.32	5.27
9	6.33	6.14	6.24
10	10.10	10.48	11.07
$\overline{Y}_{d.}(15,d)$	7.94	8.21	8.43
$\sum_{i=1}^{R}\overline{Y}_{r.}^{2}$	826.20	894.68	938.34
S^2	21.75	24.52	25.30
S	4.66	4.95	5.03
$S/\sqrt{10} = s.e.(\overline{Y})$	1.47	1.57	1.59

Sample Size I

- To estimate a long-run performance measure, θ, within $\pm\varepsilon$ with confidence $100(1-\alpha)\%$.
- M/G/1 queueing example (cont.):
  - We know: $R_0 = 10$, $d = 2$ deleted and $S_0^2 = 25.30$.
  - To estimate the long-run mean queue length, L_Q, within $\varepsilon = 2$ customers with 90% confidence ($\alpha = 10\%$).
  - Initial estimate:

$$R \geq \left(\frac{z_{0.05} S_0}{\varepsilon}\right)^2 = \frac{1.645^2(25.30)}{2^2} = 17.1$$

  - Hence, at least 18 replications are needed, next try $R = 18, 19, \ldots$ using $R \geq \left(t_{0.05,R-1} S_0/\varepsilon\right)^2$. We found that

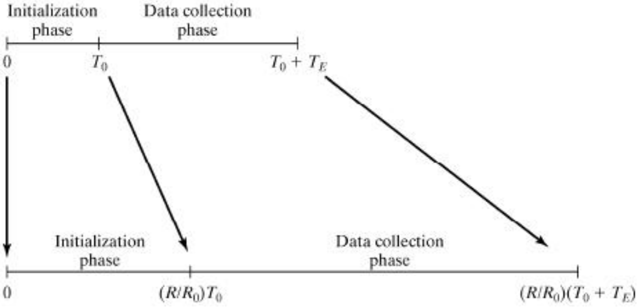
$$R = 19 \geq \left(t_{0.05,R-1} S_0/\varepsilon\right)^2 = \left(1.73^2 \cdot 25.3/4\right) = 18.93$$

  - Additional replications needed is $R - R_0 = 19 - 10 = 9$.

Sample Size II

- An alternative to increasing R is to increase total run length $T_0 + T_E$ within each replication.
  - Approach:
    - Increase run length from $T_0 + T_E$ to $(R/R_0)(T_0 + T_E)$, and
    - Delete additional amount of data, from time 0 to time $(R/R_0)T_0$.
  - Advantage: any residual bias in the point estimator should be further reduced.
  - However, it is necessary to have saved the state of the model at time $T_0 + T_E$ and to be able to restart the model.

Sample Size III



Batch Means for Interval Estimation

- 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:

$$\overline{Y}_j = \frac{1}{m}\int_{(j-1)m}^{jm} Y(t+T_0)dt,\ j = 1, 2, \ldots, k$$

  - A discrete-time process, $\{Y_i, i = d+1, d+2, \ldots, n\}$:
    - k batches of size $m = (n-d)/k$, batch means:

$$\overline{Y}_j = \frac{1}{m}\sum_{i=(j-1)m+1}^{jm} Y_{i+d}, \qquad j = 1, 2, \ldots, k$$

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.
- Batch mean for a long run replication

References

- Averill M. Law, *Simulation Modeling and Analysis*, McGraw-Hill, 2007
- J. Banks, J. S. Carson II, B. L. Nelson, D. M. Nicol, *Discrete-Event System Simulation*, Prentice Hall, 2005
- J. Banks (ed), *Handbook of Simulation*, Wiley, 1998, Chapter 11