Estimation of Absolute Performance

Outline:

1. Type of Simulation w.r.t. Output Analysis
2. Stochastic Nature of Output Data
3. Absolute Measures
4. Output Analysis for Terminating Simulation
5. Output Analysis for Steady-State Simulation

1. Type of Simulation w.r.t. Output Analysis

Output analysis is the examination of the data generated by a simulation

Its purpose is either to predict the performance of a system or to compare the performance of two or more alternate system designs

The need for statistical output analysis is based on the observation that the output data from a simulation exhibits random variability

- due to use of random numbers to produce input variables
- Two different streams or sequences of random variables will produce two sets of outputs which will differ

1. Type of Simulation w.r.t. Output Analysis

Objective: Estimate system performance via simulation

If the system performance is measured by θ, the result of a set of simulation experiments will be an estimator $\hat{\theta}$ of θ

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. inventory cost for subsequent weeks lack statistical independence.
- Initial conditions, e.g. inventory on hand and # of backorders at time 0 would most likely influence the performance of week 1.

1. Type of Simulation w.r.t. Output Analysis

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.

Discusses the analysis of transient simulations.

Discusses the analysis of steady-state simulations.

1. Type of Simulation w.r.t. Output Analysis

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.

1. Type of Simulation w.r.t. Output Analysis

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.

2. 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 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 .

2. 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.

3. Absolute Measures

Consider the estimation of a performance parameter, θ (or ϕ), of a simulated system.

It is desired to have a "point estimate" and an "interval estimate" of θ (or ϕ)

- In many cases, there is an obvious or natural choice candidate for a point estimator. Sample mean is such an example
- Interval estimates expand on "point estimates" by incorporating the uncertainty of point estimates
  - Different samples from different intervals may have different means
  - An interval estimate quantifies this uncertainty by computing lower and upper values with a given level of confidence (i.e., probability)

3. Absolute Measures

Simulation output data are of the form $\{Y_1, Y_2, \ldots, Y_n\}$ for estimating θ is referred to as discrete-time data, because the index n is discrete valued

The simulation data of the form $\{Y(t),\ 0 \le t \le T_E\}$ is referred to as continuous-time data with time-weighted mean ϕ because the index t is continuous valued.

Point estimation for discrete time data.

- The point estimator:

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

  - Is unbiased if its expected value is θ, that is if:

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

Desired

  - Is biased if:

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

3. Absolute Measures: Point Estimator

Point estimation for continuous-time data.

- The point estimator:

$$\hat{\phi} = \frac{1}{T_E} \int_0^{T_E} Y(t)dt$$

  - Is biased in general where: $E(\hat{\phi}) \neq \phi$.
  - An unbiased or low-bias estimator is desired.

Usually, system performance measures can be put into the common framework of θ or ϕ:

- e.g., the proportion of days on which sales are lost through an out-of-stock situation, let:

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

3. Absolute Measures : Point Estimator

Performance measure that does not fit this common framework is a "quantile" or "percentile"

$$\Pr\{Y \leq \theta\} = p$$

- e.g., $p=0.85$; 85% of the customers will experience a delay of θ minutes are less. Or a customer has only a 0.15 probability of experiencing a delay longer than θ minutes.
- Estimating quantiles: the inverse of the problem of estimating a proportion or probability. In estimating probability, a proportion θ is given and p is to be estimated; but in estimating a quantile, p is given and θ is to be estimated.
- Consider a histogram of the observed values Y:
  - Find $\hat{\theta}$ such that $100p$% of the histogram is to the left of (smaller than) $\hat{\theta}$.
  - e.g., if we observe $n=250$ customer delays, then an estimate $\hat{\theta}$ of the 85^{th} percentile of delay is a value such that (0.85)(250)=212.5 ≈213 of the observed values are less than or equal to θ.

3. Absolute Measures: Confidence-Interval Estimation

To understand confidence intervals fully, it is important to distinguish between *measures of error*, and *measures of risk*

- contrast the confidence interval with a prediction interval (another useful output-analysis tool).
- Both confidence and prediction intervals are based on premise that the data being produced by the simulation is well represented by a probability model

3. Absolute Measures: Confidence-Interval Estimation

Consider a manufacturing system producing parts and the performance measure is cycle time for parts (time from release into the factory until completion). Y_{ij} is the cycle time for j^{th} part produced in i replication.

Within Replication Data	Across Replication Data
Y_{11} Y_{12} …… Y_{1n1}	$\bar{Y}_{1.}, S_1^2, H_1$
Y_{21} Y_{22} …… Y_{2n2}	$\bar{Y}_{2.}, S_2^2, H_2$
………	
Y_{R1} Y_{R2} …… Y_{RnR}	$\bar{Y}_{R.}, S_R^2, H_R$

H is confidence interval half-width

$\bar{Y}_{..}, S^2, H$

3. Absolute Measures: Confidence-Interval Estimation

Suppose the model is the normal distribution with mean θ, variance σ^2 (both unknown).

- Let $\bar{Y}_{i.}$ be the average cycle time for parts produced on the i^{th} replication (representing a day of production) of the simulation.
  - Therefore, its mathematical expectation is θ and let σ be the day-to-day variation of the average cycle-time
- Suppose our goal is to estimate θ
- Average cycle time will vary from day to day, but over the long-run the *average of the averages* will be close to θ.
- The natural estimator for θ is the overall sample mean of R independent replications, $\bar{Y}_{..} = \sum_{i=1}^{R} Y_{i.}/R$, but it is not θ, is only estimate
- A confidence interval (CI) is a measure of that error $S^2 = \frac{1}{R-1}\sum_{i=1}^{R}(Y_{i.} - Y_{..})^2$
- Let Sample variance across R replications:

3. Absolute Measures: Confidence-Interval Estimation

Confidence Interval (CI):

- A measure of error.
- Assumes $Y_{i\cdot}$ are normally distributed.

$$\bar{Y}_{\cdot\cdot} \pm t_{\alpha/2,R-1} \frac{S}{\sqrt{R}}, \quad \text{where } t_{\alpha/2,R-1} \text{ is the quantile of t-distribution}$$

- 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).
- Unfortunately, the confidence interval itself may be wrong!!

3. Absolute Measures: Confidence-Interval Estimation

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 as the daily average varies.
- 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:

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

- The length of PI will not go to *0* as *R* increases because we can never simulate away risk.
- PI's limit is: $\theta \pm z_{\alpha/2}\sigma$ indicating no matter how much we simulate, the daily average still varies.

3. Absolute Measures: Confidence-Interval Estimation

Example:

- Suppose that the overall average of the average cycle time on 120 replications of a manufacturing simulation is 5.80 hours, with a sample standard deviation of 1.60 hours
- Since $t_{0.025,119}$=1.98, a 95% confidence interval for the long-run expected daily average cycle time is $5.80 \pm 1.98(1.60/\sqrt{120})$ or 5.80 ± 0.29 hours.
  - Our best guess for average cycle time is 5.80 hours, but there could be as much as ±0.29 hours error in that estimate
- On any particular day, we are 95% confident that the average cycle time for all parts produced on that day will be $5.80 \pm 1.98(1.60)\sqrt{(1+1/120)} = 5.80 \pm 3.18$ hours!!

4. Output Analysis for Terminating Simulations

A terminating simulation: runs over a simulated time interval $[0, T_E]$ and results in observations $Y_1, \ldots, Y_n$

The sample size n may be a fixed number or a random variable.

A common goal is to estimate:

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

$$\phi = E\left(\frac{1}{T_E}\int_0^{T_E} Y(t)\,dt\right), \quad \text{for continuous output } Y(t), 0 \le t \le T_E$$

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

4. Output Analysis for Terminating Simulations: Statistical Background

It is very important to distinguish within-replication data from across-replication data.

The issue is further confused by the fact that simulation languages only provide summary of the measures and not the raw data.

For example, consider simulation of a manufacturing system

- Two performance measures of that system: cycle time $\bar{Y}_{i.}$ 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

4. Output Analysis for Terminating Simulations: Statistical Background

Across Replication:

- For example: the daily cycle time averages (discrete time data)
  - The average: $\bar{Y}_{..} = \frac{1}{R}\sum_{i=1}^{R} Y_{i.}$
  - The sample variance: $S^2 = \frac{1}{R-1}\sum_{i=1}^{R}(Y_{i.} - \bar{Y}_{..})^2$
  - The confidence-interval half-width: $H = t_{\alpha/2,R-1}\frac{S}{\sqrt{R}}$

Within replication:

- For example: the WIP (a continuous time data)
  - The average: $\bar{Y}_{i.} = \frac{1}{T_{Ei}}\int_0^{T_{Ei}} Y_i(t)\,dt$
  - The sample variance: $S_i^2 = \frac{1}{T_{Ei}}\int_0^{T_{Ei}} \left(Y_i(t) - \bar{Y}_{i.}\right)^2 dt$

4. Output Analysis for Terminating Simulations: Statistical Background

Overall sample average, $\bar{Y}$, and the interval replication sample averages, $\bar{Y}_{i.}$, 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.

4. Output Analysis for Terminating Simulations: C.I. with Specified Precision

Sometimes we would like to estimate CI with a specified precision

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

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

R is the # of replications

S^2 is the sample variance

Suppose that an error criterion ε is specified with probability $1 - \alpha$, a sufficiently large sample size should satisfy: $P\left(\left|\bar{Y}_{..} - \theta\right| < \varepsilon\right) \geq 1 - \alpha$

(in other words, it is desired to estimate θ by $\bar{Y}_{..}$)

4. Output Analysis for Terminating Simulations: C.I. with Specified Precision

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: $\bar{Y}_{..} \pm t_{\alpha/2,R-1} \dfrac{S}{\sqrt{R}}$

4. Output Analysis for Terminating Simulations: C.I. with Specified Precision

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 $S_0^2 = (0.072)^2 = 0.00518$.
- 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 must be greater than this #

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

4. Output Analysis for Terminating Simulations: Quantiles

To present the interval estimator for quantiles,

- it is helpful to look at the interval estimator for a mean in the special case when mean represents a proportion or probability, p

In this book, a proportion or probability is treated as a special case of a mean.

4. Output Analysis for Terminating Simulations: Quantiles

When the number of independent replications $Y_1, \ldots, Y_R$ is large enough that $t_{\alpha/2,n-1} = z_{\alpha/2}$, the confidence interval for a probability p is often written as:

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

The sample proportion

A quantile is the inverse of the probability to the probability estimation problem:

p is given

Find θ such that $Pr(Y \leq \theta) = p$

4. Output Analysis for Terminating Simulations: Quantiles

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.

- 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 estimate**
12.5
12.9

4. Output Analysis for Terminating Simulations: Quantiles

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

$$\text{where} \quad p_\ell = p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{R-1}}$$

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

4. Output Analysis for Terminating Simulations: Quantiles

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:

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

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

  - Independent of the initial conditions.

5. 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: $Y_{r1}, Y_{r2}, Y_{r3}, \ldots$

5. Output Analysis for Steady-State Simulation: Initialization Bias

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.

5. Output Analysis for Steady-State Simulation: Initialization Bias

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).

5. Output Analysis for Steady-State Simulation: Initialization Bias

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 rth 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.

5. Output Analysis for Steady-State Simulation: Error Estimation

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 θ is the sample mean

$$\bar{Y} = \sum_{i=1}^{n} Y_i / n$$

- Variance of $\bar{Y}$ is almost impossible to estimate.
- For system with steady state, produce an output process that is approximately covariance stationary (after passing the transient phase).
  - The covariance between two random variables in the time series depends only on the lag (the # of observations between them).

5. Output Analysis for Steady-State Simulation: Error Estimation

For a covariance stationary time series, $\{Y_1, \ldots, Y_n\}$:

- Lag-k autocovariance is: $\delta_k = \text{cov}(Y_1, Y_{1+k}) = \text{cov}(Y_i, Y_{i+k})$

- Lag-k autocorrelation is: $\rho_k = \dfrac{\gamma_k}{\sigma^2}$

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

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

The expected value of the variance estimator is:

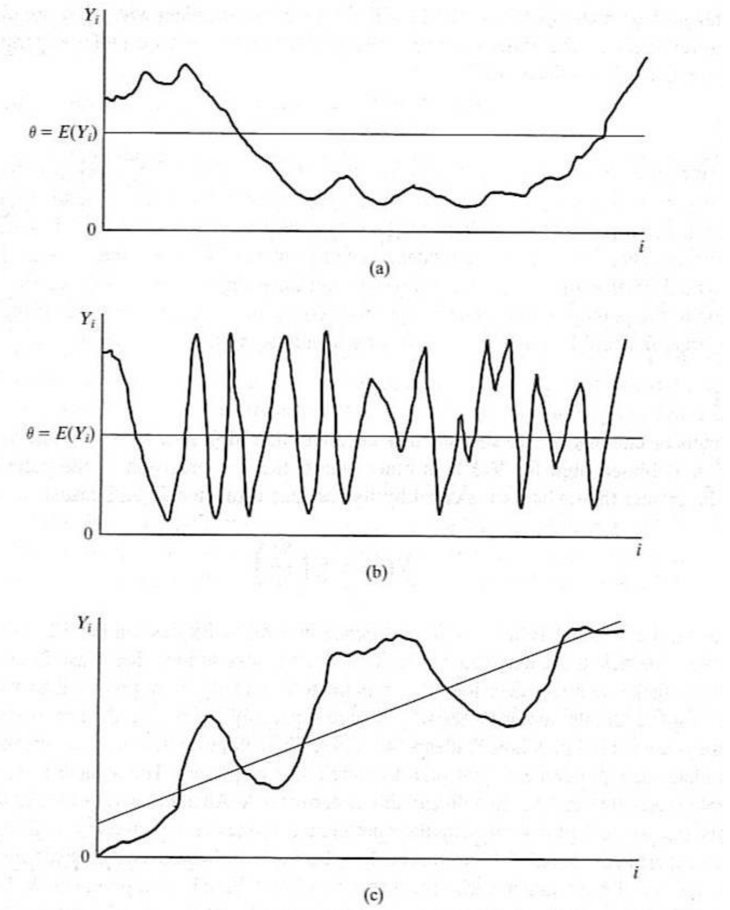
$$E\left(\frac{S^2}{n}\right) = BV(\bar{Y}), \qquad \text{where } B = \frac{n/c - 1}{n - 1}$$

5. Output Analysis for Steady-State Simulation: Error Estimation

Stationary time series Y_i exhibiting positive autocorrelation.

Stationary time series Y_i exhibiting negative autocorrelation.

Nonstationary time series with an upward trend



5. Output Analysis for Steady-State Simulation: Error Estimation

The expected value of the variance estimator is:

$$E\left(\frac{S^2}{n}\right) = BV(\bar{Y}), \qquad \text{where } B = \frac{n/c-1}{n-1} \text{ and } V(\bar{Y}) \text{ is the variance of } \bar{Y}$$

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

5. Output Analysis for Steady-State Simulation: Replication Method

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 $\{Y_{rj}, r = 1, ..., R; j = 1, ..., n\}$ is derived by:

- Individual observation from within replication *r*.
- Batch mean from within replication *r* of some number of discrete-time observations.
- Batch mean of a continuous-time process over time interval *j*.

5. Output Analysis for Steady-State Simulation: 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, $\{Y(t),\ T_0 \le t \le T_0 + T_E\}$:

- 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: $$\bar{Y}_j = \frac{1}{m}\sum_{i=(j-1)m+1}^{jm} Y_{i+d}$$

5. Output Analysis for Steady-State Simulation : Batch Means for Interval Estimation

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

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.