IE630: Simulation Modelling & Analysis
Output Analysis – Steady State Systems

Saurabh Jain

Assistant Professor

Department of Industrial Engineering and Operations Research

IIT Bombay

Quick Recap

Topics of Discussion

- Obtain a CI for a steady-state mean, $\nu = E(Y)$, of the output process $Y_1, Y_2, \ldots \ldots \ldots \ldots$

- On the other hand, $E(Y_2) \neq \nu$ for "small" i because of the initial condition

Topics to be covered:

- Determining and eliminating the initial warm-up bias
- Obtaining sample observations in this case
- Determining run length
  - Theoretically Infinite, but practical number?

Introduction

Things to check ….

- Starting and stopping is essence of model
- Interested in the long-run characteristics of the system

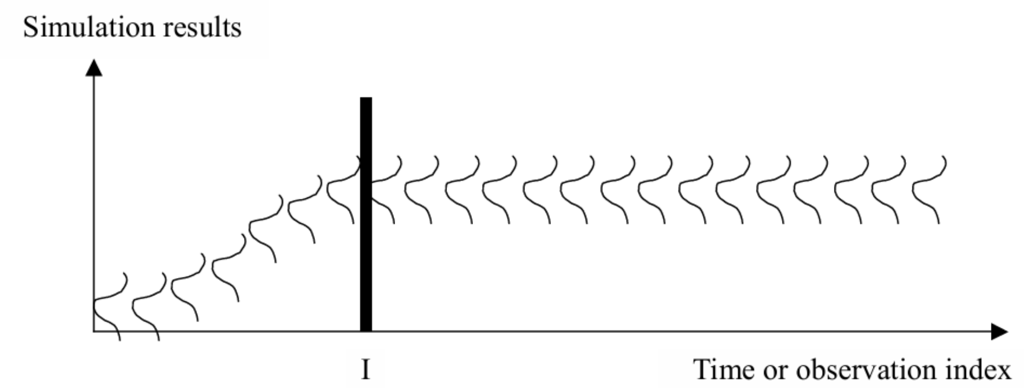
Let $Y_1, Y_2, \dots \dots \dots \dots$ be the output → single run of the non-terminating simulation

Issues…..

- The problem of initial start-up due to arbitrary initial conditions
- Eliminate warm-up bias
- Obtaining sample observations to build CIs - Need to choose, IID data points, which are "representative" of the "steady-state behavior"
- Determining run-length - Theoretically infinite, but practically how long?

How to compute Warm-up Period?

- System starts from empty and idle



i. Run simulation model with a ‘warm-up’ period and remove data collected during warm-up, and use only remaining data
ii. Set initial conditions such that model starts in realistic / steady state conditions
iii. Set partial initial condition, then warm-up the model, and remove the warm-up data
iv. Run model for long time, making the bias effect negligible

How to compute Warm-up Period?

- **Methods:**
  - Visually inspect the performance and choose warm-up period based on your judgement
  - Time series inspection
  - Ensemble methods
  - Welch’s moving average method
  - Autocorrelation estimators
  - Marginal Error rule
  - Neural Networks
  - …..
  - Statistical Tests
  - Goodness-of-Fit Test
  - Kalman Filters
  - …..

Welch moving average method

Welch moving average method

- Conduct n replications ($n \geq 5$) of simulation, each of length m
  - $Y_{ji} \rightarrow i^{th}$ observation from j^{th} replication (j=1…n; i=1…m).

- Compute the averaged process

- Compute ensemble moving averages, $\bar{Y}_i\ (w)$
  - For i-th datapoint, construct moving average centered at i with w values on either side.

- Choose T_0 to be value of i beyond which ensemble averages seems to have converged

Welch moving average method… TIPS

- Initially, make n = 5 replications;
  - m should as large as possible (larger than anticipated T0 , allows infrequent events to occur)

- Plot $\bar{Y}_i\ (w)$ for several different values of w.
  - Choose the smallest value of w for which the corresponding plot is “reasonably smooth”. Use this plot to determine the length of the warm up period, T_0.

- If no value of w is satisfactory, make 5 more replications of length m. Repeat step 2 using all available data.

Examples

MSER Heuristic

- Collect 1 replication of the finite stochastic output data
- Compute the Truncated Mean

$$\bar{Y}_{n,d} = \frac{1}{n-d} \sum_{j=d+1}^{n} Y_j \,, \qquad \forall d = 1..n$$

- Compute optimal truncation point (Test Statistic)

$$d^* = \underset{n \gg d \geq 0}{\arg min} \left[\frac{1}{(n-d)^2} \sum_{j=d+1}^{n} \left(Y_j - \bar{Y}_{n,d}\right)^2 \right],$$

- Advantages
  - Does not require multiple replication data. Can obtain truncation point/ warm-up period for each output data series.

MSER Heuristic

- Variants
  - Using batch average of size, say, m ➔ MSER-m
- Takeaways:
  - If the MSER test statistic is non-increasing, perform MSER with larger data set.
  - Reject d* value if (d* > n/2) ➔ Indicates that transient period is long. Perform MSER again with ( )
  - MSER can be performed for each replication!

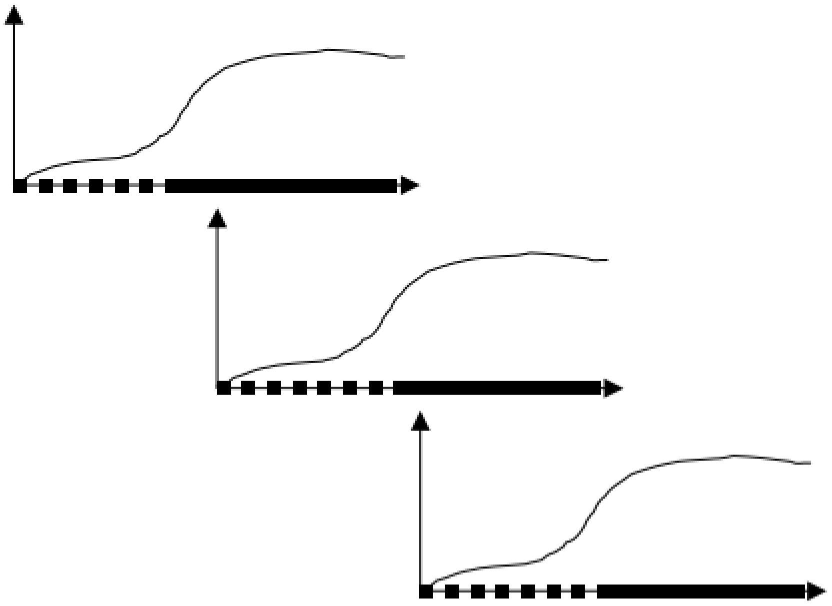
Examples

Issues

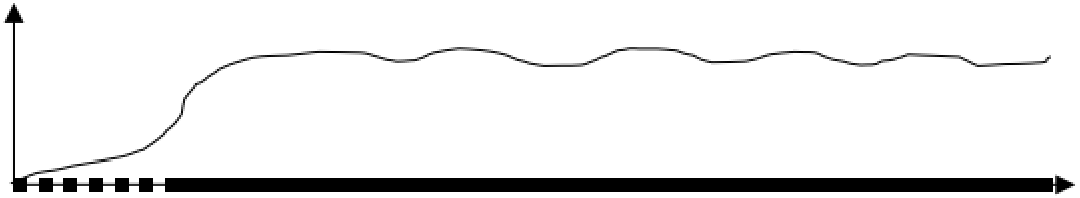
- Determining and eliminating the initial warm-up bias
- Obtaining sample observations in this case
  - Need to choose, (   ), which are (   ) of the (   )
- Determining run length
  - Theoretically Infinite, but practical number?

Obtaining Sample Observation

- There are two methods obtaining independent sample observations for non-terminating simulations:
- **Truncated replications** (multiple replications)

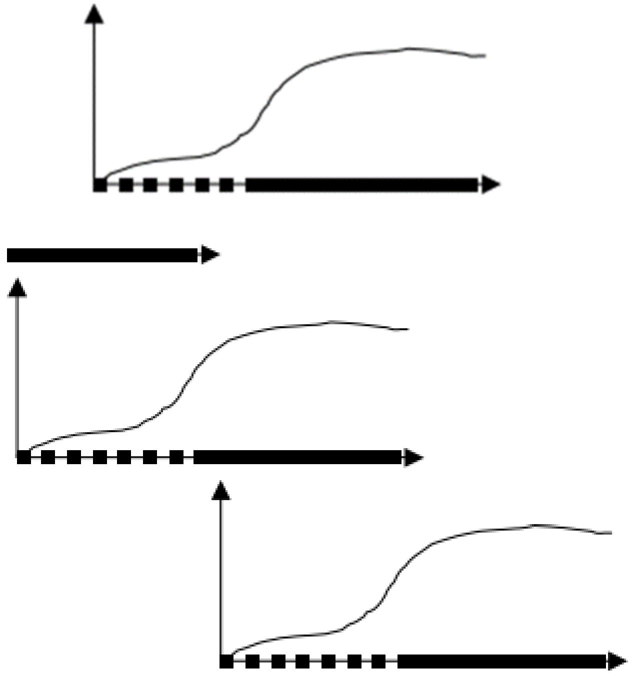


- **Interval batching methods** (batch mean techniques)



Truncated replications (multiple replications)

- Same with the terminating simulation case except:
- I (bias) must be determined and removed (Welch's method)
- An appropriate run length must be determined (will be explained soon)
- Advantage
  - It ensures that samples are independent
- Disadvantage
  - Running through the warm up phase for every replication
  - takes a lot of time
  - Throw away n*I amount of data, where I is warm up period



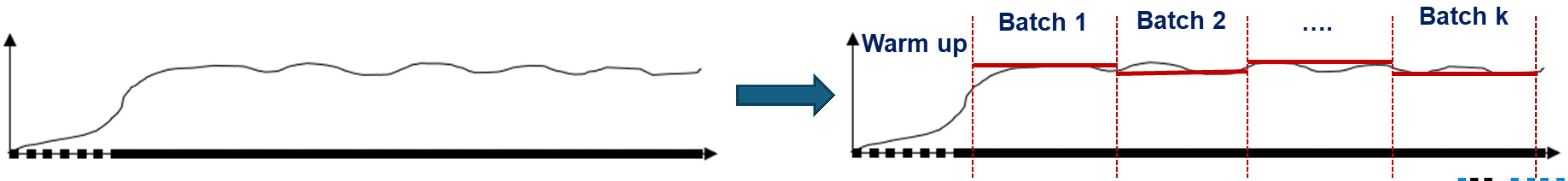
Truncated replications (multiple replications)

- Warm-up Period = 30 hours
- Execute simulation for 10* Warm-up = 10*30 hrs = 300 hours
- Number of replications =
- Average throughput for each replication =

60.8	59.42	60.93	58.51	59.48
59.8	59.16	58.71	58.64	60.96

Batch Means Technique

- Still need to run about 5 to figure out the warm up period
- Then, a super long run (only 1 replication) is made, and statistics are collected during different periods (batches) of time => Beyond the warm up period, the run is divided into non-overlapping intervals
- Intervals
  - Equal amount of time (time-based): work in process, number of parts in the queue
  - Equal number of observations (observation-based): time spent in the queue, throughput times



Batch Means Technique

- Obtain output data from a single long replication
  - Remove warmup period data
  - Collect at least 10 times the data as deleted.
- For batches from output data
  - Batch means will be correlated if batch size is small (recall we need uncorrelated data for CI)
  - Need to check auto-correlation of data at different lags
- Determine the run length,
  - That is, the number of batches!
- Construct CI
  - Observations → batch means
  - n → number of batches

Batch Run Length Determination

- Given a batch run length, it is suggested to have at least 10 batches => so that the percentiles of the t-distribution will be relatively close to those of normal
- The size of batch to be at least 100 under most circumstances

- If a trade-off must be made between the number of batch intervals and the batch run length, it's better to have a few independent observations than to have several autocorrelated observations
- The batch means method is robust with respect to errors in determining the length of the warm up period
  - Because it is likely that only the first batch mean has significant bias (assuming the batch sizes are large)

Batch Run Length Determination

- Lag
- Distances apart from each other

- Correlogram (or autocorrelation plot) Plot of the correlation between points at various lags

Autocorrelation diagram

- Mathematical Representation of Correlogram

$$\rho = \sum_{i=1}^{n-j} \frac{(x_i - \bar{x})(x_{i+j} - \bar{x})}{\sigma^2 (n-j)}$$

where j is the lag and σ is the standard deviation of the sample (approximated by “s”).

Summary

- Terminating system
  - n*m data => use only n data (last average column) to calculate a CI
- Steady state system
  - Determination of a warm-up period
    - Informal visualization (multiple replications)
    - Formal method: Welch’s moving average method
  - Run length determination
    - Informal method:
      - length which includes most of events (even rare ones)
      - Formal method: correlogram
    - Given a warm up period and run length
      - Truncated method: same with terminating except removal of warm up for each replication
      - Batch means method: one lengthy replication, decomposed to n independent samples