Simulation
Lecture 17 - Output Analysis for Terminating Simulations

Burak Büke

School of Mathematics
University of Edinburgh

Output Analysis for Terminating Simulations

- Terminating Simulation: Definite start and end time/conditions
- Customer arrivals to a bank
  - System starts at 9am with no customers
  - System ends around 6pm
- Project Management
  - Starts with the first activity
  - Ends with the final activity

Output Analysis for Terminating Simulations

- Terminating Simulation: Definite start and end time/conditions
- Customer arrivals to a bank
    - System starts at 9am with no customers
    - System ends around 6pm
- Project Management
    - Starts with the first activity
    - Ends with the final activity

Output Analysis for Terminating Simulations

- Terminating Simulation: Definite start and end time/conditions
- Customer arrivals to a bank
  - System starts at 9am with no customers
  - System ends around 6pm
- Project Management
  - Starts with the first activity
  - Ends with the final activity

Output Analysis for Terminating Simulations

- Consider estimating average cycle time (time spent in the system)
- Run the system once $\Rightarrow$ Estimate average cycle time
  - Random quantity
  - How can we assess the quality of the estimate?

Output Analysis for Terminating Simulations

- Consider estimating average cycle time (time spent in the system)
- Run the system once$\Rightarrow$ Estimate average cycle time
  - Random quantity
  - How can we assess the quality of the estimate?

Output Analysis for Terminating Simulations

- Consider estimating average cycle time (time spent in the system)
- Run the system once $\Rightarrow$ Estimate average cycle time
    - Random quantity
    - How can we assess the quality of the estimate?

Central Limit Theorem

Theorem 1

Central Limit Theorem. *Let $X_1, X_2, \cdots$ be independent and identically distributed (i.i.d.) random variables with expected value μ and variance σ^2. Then*

$$\frac{\frac{\sum_{i=1}^{n} X_i}{n} - \mu}{\sigma / \sqrt{n}}$$

approaches to a standard normal random variable.

Output Analysis for Terminating Simulation

- Suppose we run a single replication (single day)

$$X_1, X_2, X_3, \ldots, X_n \Rightarrow \bar{X}, S^2$$

- Can we build a confidence interval based on these?
- NOT INDEPENDENT!!!

Output Analysis for Terminating Simulation

- Suppose we run a single replication (single day)

$$X_1, X_2, X_3, \ldots, X_n \Rightarrow \bar{X}, S^2$$

- Can we build a confidence interval based on these?
- NOT INDEPENDENT!!!

Output Analysis for Terminating Simulation

- Suppose we run a single replication (single day)

$$X_1, X_2, X_3, \ldots, X_n \Rightarrow \bar{X}, S^2$$

- Can we build a confidence interval based on these?
- NOT INDEPENDENT!!!

Output Analysis for Terminating Simulations

- Run m (independent) replications

$$\begin{array}{c} X_1^1, X_2^1, \ldots, X_{n_1}^1 \Rightarrow \bar{X}_1 \\ X_1^2, X_2^2, \ldots, X_{n_2}^2 \Rightarrow \bar{X}_2 \\ \vdots \\ X_1^m, X_2^m, \ldots, X_{n_m}^m \Rightarrow \bar{X}_m \end{array}$$

- $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_m$ are indepedent
- Find the sample mean and variance of these

$$\bar{X} = \frac{\sum_{i=1}^m \bar{X}_i}{m}, S^2 = \frac{\sum_{i=1}^m (\bar{X}_i - \bar{X})^2}{m-1}.$$

- Build $(1-\alpha) \times 100$ confidence interval as

$$[\bar{X} - t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}, \bar{X} + t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}]$$

Output Analysis for Terminating Simulations

- Run m (independent) replications

$$\begin{array}{c} X_1^1, X_2^1, \ldots, X_{n_1}^1 \Rightarrow \bar{X}_1 \\ X_1^2, X_2^2, \ldots, X_{n_2}^2 \Rightarrow \bar{X}_2 \\ \vdots \\ X_1^m, X_2^m, \ldots, X_{n_m}^m \Rightarrow \bar{X}_m \end{array}$$

- $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_m$ are indepedent
- Find the sample mean and variance of these

$$\bar{X} = \frac{\sum_{i=1}^m \bar{X}_i}{m}, S^2 = \frac{\sum_{i=1}^m (\bar{X}_i - \bar{X})^2}{m-1}.$$

- Build $(1-\alpha) \times 100$ confidence interval as

$$[\bar{X} - t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}, \bar{X} + t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}]$$

Output Analysis for Terminating Simulations

- Run m (independent) replications

$$\begin{gathered} X_1^1, X_2^1, \ldots, X_{n_1}^1 \Rightarrow \bar{X}_1 \\ X_1^2, X_2^2, \ldots, X_{n_2}^2 \Rightarrow \bar{X}_2 \\ \vdots \\ X_1^m, X_2^m, \ldots, X_{n_m}^m \Rightarrow \bar{X}_m \end{gathered}$$

- $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_m$ are indepedent
- Find the sample mean and variance of these

$$\bar{X} = \frac{\sum_{i=1}^m \bar{X}_i}{m}, S^2 = \frac{\sum_{i=1}^m (\bar{X}_i - \bar{X})^2}{m-1}.$$

- Build $(1-\alpha) \times 100$ confidence interval as

$$[\bar{X} - t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}, \bar{X} + t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}]$$

Output Analysis for Terminating Simulations

- Run m (independent) replications

$$\begin{array}{c} X_1^1, X_2^1, \ldots, X_{n_1}^1 \Rightarrow \bar{X}_1 \\ X_1^2, X_2^2, \ldots, X_{n_2}^2 \Rightarrow \bar{X}_2 \\ \vdots \\ X_1^m, X_2^m, \ldots, X_{n_m}^m \Rightarrow \bar{X}_m \end{array}$$

- $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_m$ are indepedent
- Find the sample mean and variance of these

$$\bar{X} = \frac{\sum_{i=1}^m \bar{X}_i}{m}, S^2 = \frac{\sum_{i=1}^m (\bar{X}_i - \bar{X})^2}{m-1}.$$

- Build $(1-\alpha) \times 100$ confidence interval as

$$[\bar{X} - t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}, \bar{X} + t_{m-1,1-\alpha/2}\frac{S}{\sqrt{m}}]$$

Simulation Example – Project Management

- We are doing a project which is composed of 6 activities.
- All the predecessors of n activity should be completed before starting the activity

Activity	Predecessor(s)	Duration
A	–	Expon(0.2/day)
B	A	Expon(0.3/day)
C	A	Uniform(1,4)days
D	B, C	Uniform(3,5) days
E	C	Uniform(1, 10)
F	D, E	Expon(1/day)

Output Analysis for Terminating Simulations

- How many replications should we run? $m = ?$
- First determine the precision you wish for the confidence interval
  - We wish to have $[\bar{X} - \epsilon, \bar{X} + \epsilon]$
- If we know σ, we need to set

$$\epsilon = z_{1-\alpha/2} \frac{\sigma}{\sqrt{m}}$$

$$m = \left\lceil \left(\frac{z_{1-\alpha/2}\sigma}{\epsilon} \right)^2 \right\rceil$$

- What if we don't know σ?

Output Analysis for Terminating Simulations

- If we were to have data, we could estimate σ and use t distributions
    - We have not done the replications yet!
- Use a two-stage procedure

Output Analysis for Terminating Simulations

- If we were to have data, we could estimate σ and use t distributions
    - We have not done the replications yet!
- Use a two-stage procedure

Output Analysis for Terminating Simulations

- If we were to have data, we could estimate σ and use t distributions
  - We have not done the replications yet!
- Use a two-stage procedure

Two-Stage Method

1. Run m_0 replications to obtain $Y^1, Y_2, \cdots, Y^{m_0}$.
2. Estimate

$$\bar{Y} = \frac{\sum_{i=1}^{m_0} Y^i}{n_0}, S_{m_0}^2 = \frac{\sum_{i=1}^{m_0} (Y^i - \bar{Y})^2}{m_0 - 1}$$

3. Let $m = \min\{n : n \geq \frac{t_{n-1,1-\alpha}^2 S_{m_0}^2}{\epsilon^2}\}$.
4. Run m "new" replications to obtain $X^1, X^2, \ldots, X^m$.
5. Form

$$\bar{X} = \frac{\sum_{i=1}^{m} X^i}{m}, \text{ and } s^2 = \frac{\sum_{i=1}^{m} (X^i - \bar{X})^2}{m-1}.$$

6. Find the confidence interval as

$$(\bar{X} - \frac{t_{m-1,1-\alpha} s}{\sqrt{m}}, \bar{X} + \frac{t_{m-1,1-\alpha} s}{\sqrt{m}}).$$

Sequential Procedure

- Two stage method does not "guarantee" the desired width!(Why?)
- The following sequential procedure guarantees the desired width:
  1. Perform m_0 replications and get $X^1, X^2, \ldots, X^{m_0}$ and set $m = m_0$
  2. Form
  $$\bar{X} = \frac{\sum_{i=1}^{m} X^i}{m}, \text{ and } s^2 = \frac{\sum_{i=1}^{m}(X^i - \bar{X})^2}{m-1}.$$
  3. If $\frac{t_{m-1,1-\alpha} s}{\sqrt{m}} < \epsilon$ stop and output the confidence interval $[\bar{X} - \epsilon, \bar{X} + \epsilon]$.
  4. Otherwise generate X^{m+1} by performing another replication, set $m = m + 1$ and go to Step 2.