Analysis of Simulation Results

Raj Jain
Washington University
Saint Louis, MO 63130
Jain@cse.wustl.edu

Audio/Video recordings of this lecture are available at:

http://www.cse.wustl.edu/~jain/cse574-08/

Overview

- Analysis of Simulation Results
- Model Verification Techniques
- Model Validation Techniques
- Transient Removal
- Terminating Simulations
- Stopping Criteria: Variance Estimation
- Variance Reduction

Model Verification vs. Validation

- Verification ⇒ Debugging
- Validation ⇒ Model = Real world

- Four Possibilities:
  1. Unverified, Invalid
  2. Unverified, Valid
  3. Verified, Invalid
  4. Verified, Valid

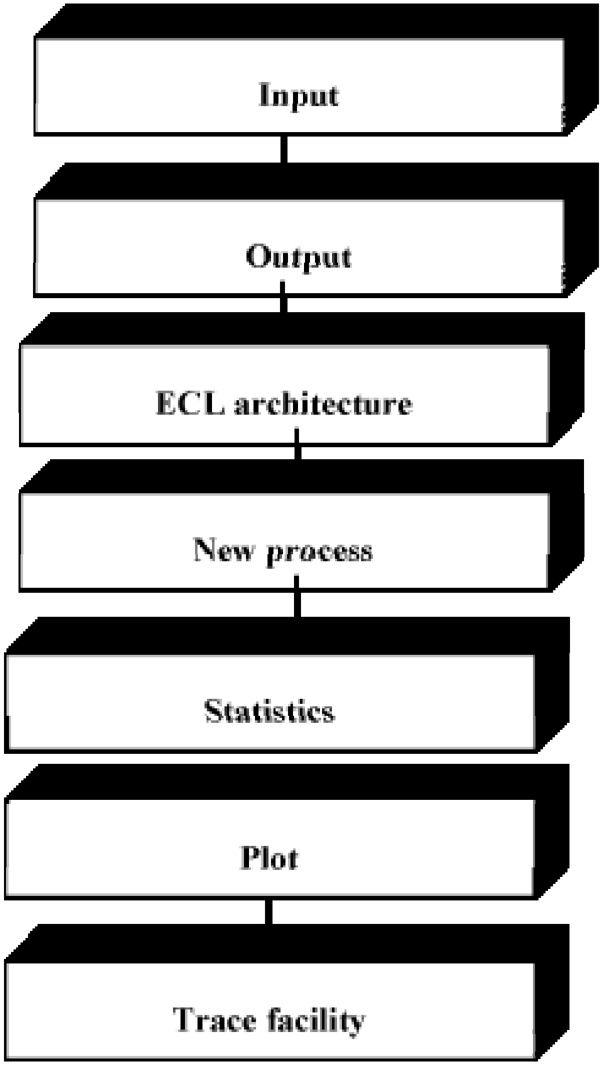
Model Verification Techniques

1. Top Down Modular Design
2. Anti-bugging
3. Structured Walk-Through
4. Deterministic Models
5. Run Simplified Cases
6. Trace
7. On-Line Graphic Displays
8. Continuity Test
9. Degeneracy Tests
10. Consistency Tests
11. Seed Independence

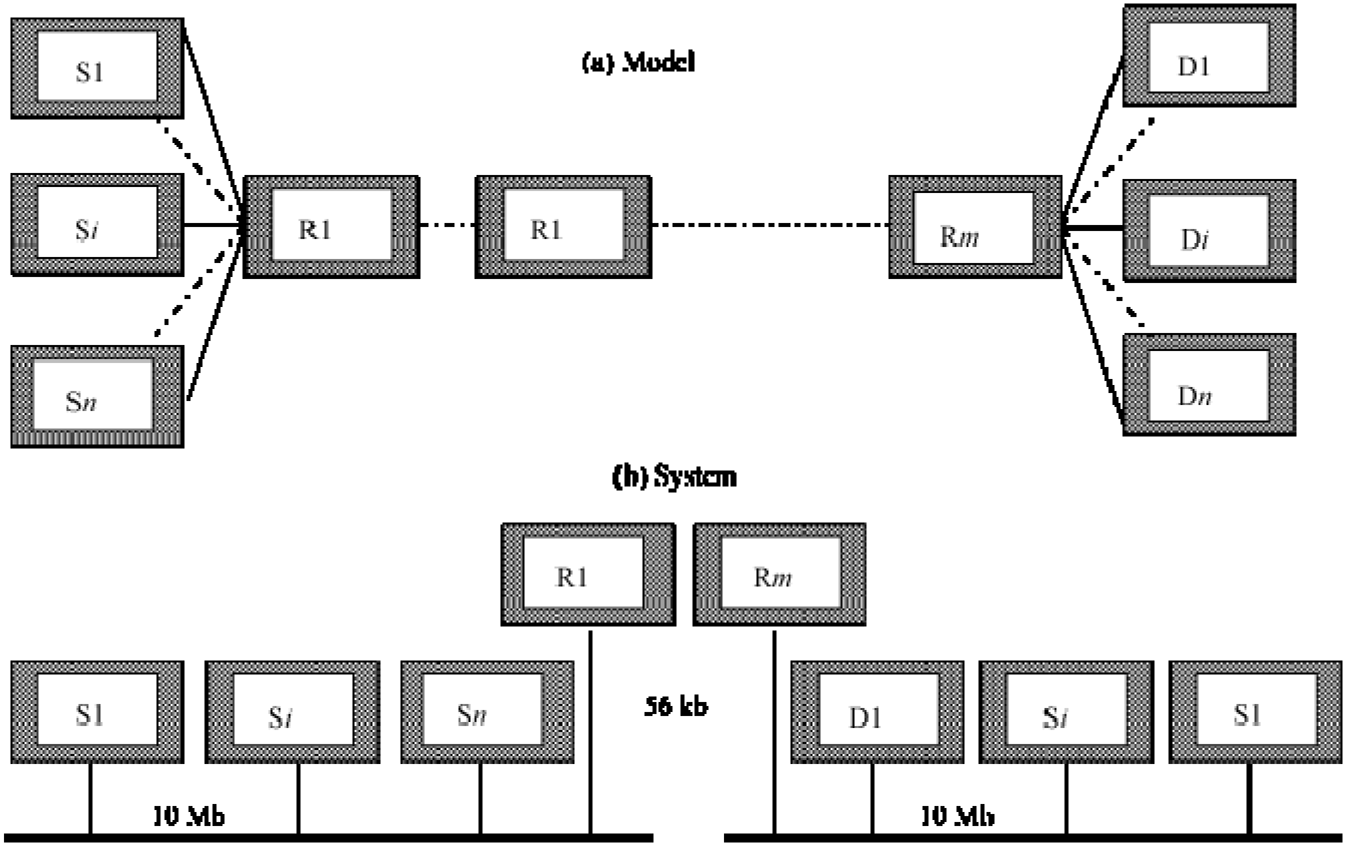
Top Down Modular Design

- Divide and Conquer
- Modules = Subroutines, Subprograms, Procedures
  - Modules have well defined interfaces
  - Can be independently developed, debugged, and maintained
- Top-down design
  ⇒ Hierarchical structure
  ⇒ Modules and sub-modules

Top Down Modular Design (Cont)



Top Down Modular Design (Cont)



Verification Techniques

- **Anti-bugging**: Include self-checks:

  $\sum$ Probabilities = 1

  Jobs left = Generated - Serviced
- **Structured Walk-Through**:
  - Explain the code another person or group
  - Works even if the person is sleeping
- **Deterministic Models**: Use constant values
- **Run Simplified Cases**:
  - Only one packet
  - Only one source
  - Only one intermediate node

Trace

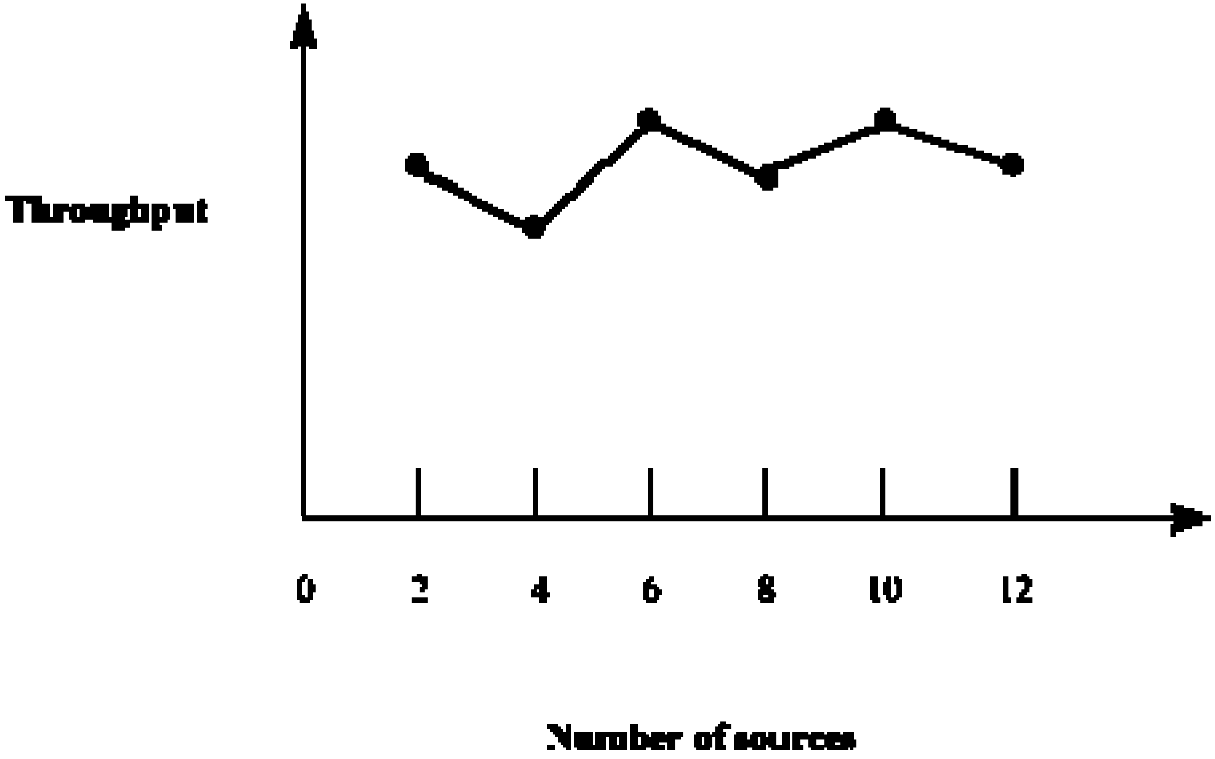
- Trace = Time-ordered list of events and variables
- Several levels of detail:
  - Events trace
  - Procedure trace
  - Variables trace
- User selects the detail
  - Include on and off
- See Fig 25.3 in the Text Book on page 418 for a sample trace

On-Line Graphic Displays

- Make simulation interesting
- Help selling the results
- More comprehensive than trace

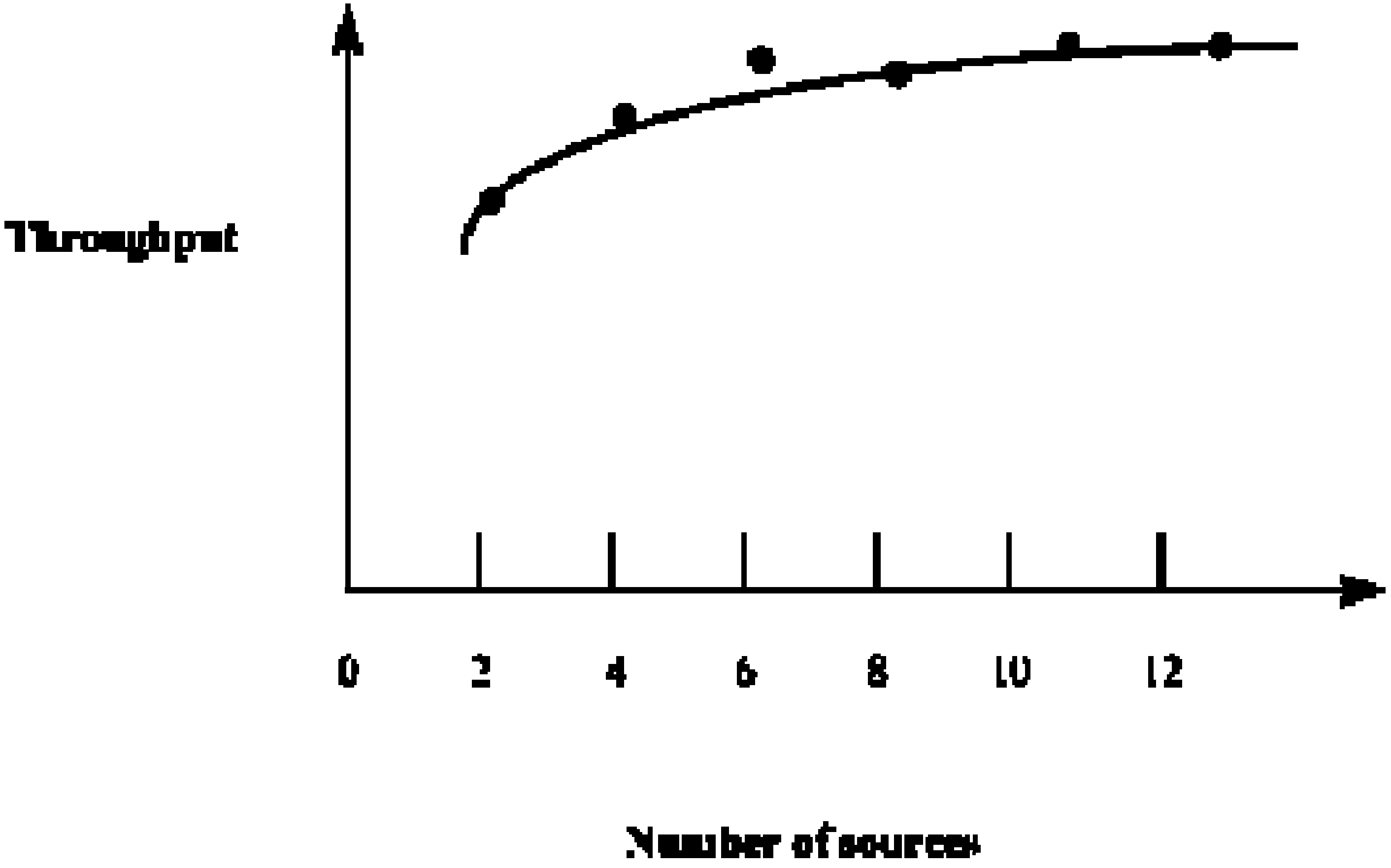
Continuity Test

- Run for different values of input parameters
- Slight change in input ⇒ slight change in output
- Before:



Continuity Test (Cont)

- After:



More Verification Techniques

- **Degeneracy Tests**: Try extreme configuration and workloads
    - One CPU, Zero disk
- **Consistency Tests**:
    - Similar result for inputs that have same effect
        - Four users at 100 Mbps vs. Two at 200 Mbps
    - Build a test library of continuity, degeneracy and consistency tests
- **Seed Independence**: Similar results for different seeds

Model Validation Techniques

- Validation techniques for one problem may not apply to another problem.
- Aspects to Validate:

  1. Assumptions
  2. Input parameter values and distributions
  3. Output values and conclusions

- Techniques:

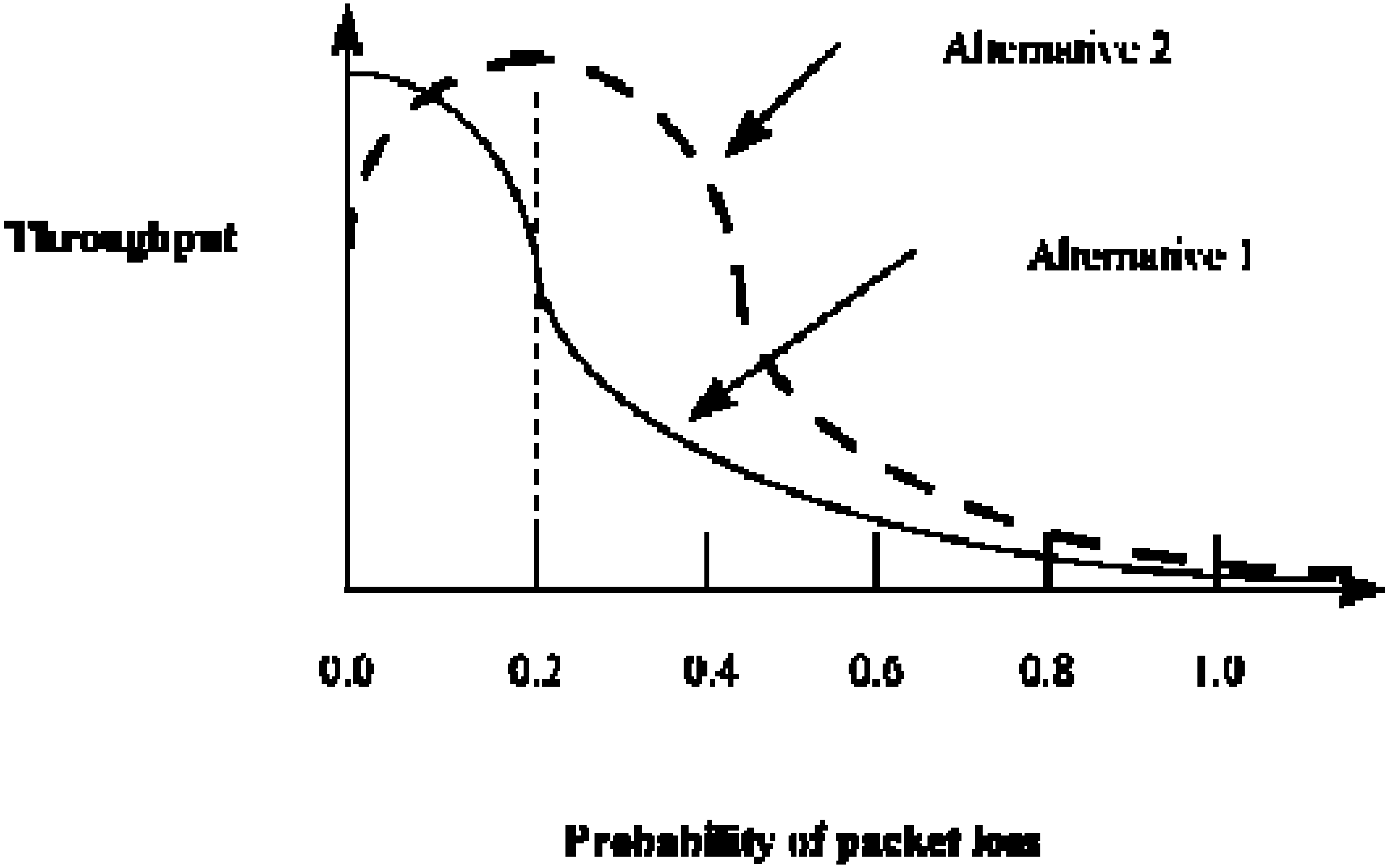
  1. Expert intuition
  2. Real system measurements
  3. Theoretical results

⇒ $3 \times 3 = 9$ validation tests

Expert Intuition

- Most practical and common way
- Experts = Involved in design, architecture, implementation, analysis, marketing, or maintenance of the system
- Selection = fn of Life cycle stage
- Present assumption, input, output
- Better to validate one at a time
- See if the experts can distinguish simulation vs. measurement

Expert Intuition (Cont)



Real System Measurements

- Compare assumptions, input, output with the real world
- Often infeasible or expensive
- Even one or two measurements add to the validity

Theoretical Results

- Analysis = Simulation
- Used to validate analysis also
- Both may be invalid
- Use theory in conjunction with experts' intuition
  - E.g., Use theory for a large configuration
  - Can show that the model is not invalid

Transient Removal

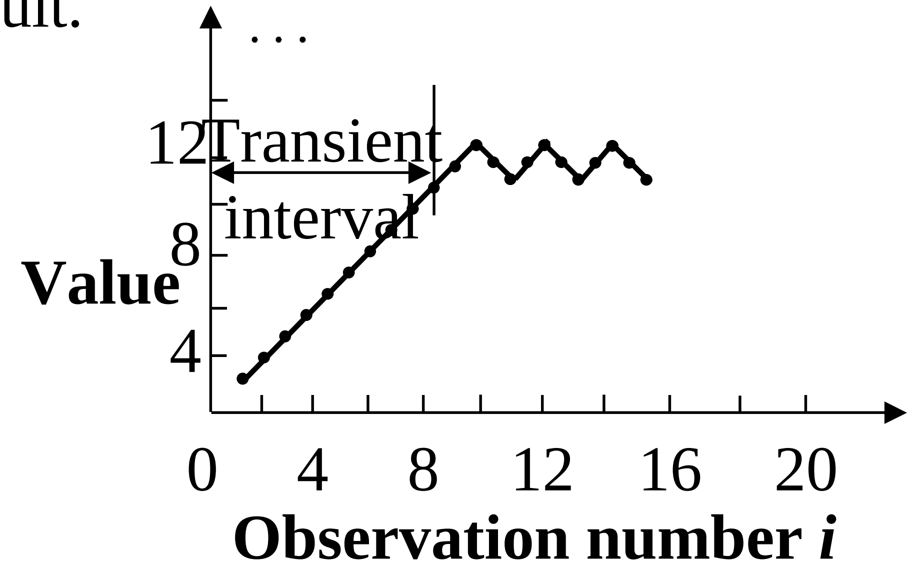
- Generally steady state performance is interesting
- Remove the initial part
- No exact definition ⇒ Heuristics:
  1. Long Runs
  2. Proper Initialization
  3. Truncation
  4. Initial Data Deletion
  5. Moving Average of Independent Replications
  6. Batch Means

Transient Removal Techniques

- **Long Runs**:
  - Wastes resources
  - Difficult to insure that it is long enough
- **Proper Initialization**:
  - Start in a state close to expected steady state
    ⇒ Reduces the length and effect of transient state

Truncation

- Assumes variability is lower during steady state
- Plot max-min of *n-l* observation for *l* =1, 2, ……….
- When *(l+1)*th observation is neither the minimum nor maximum ⇒ transient state ended
- At l = 9, Range = (9, 11), next observation = 10
- Sometimes incorrect result.



Initial Data Deletion

- Delete some initial observation
- Compute average
- No change ⇒ Steady state
- Use several replications to smoothen the average
- m replications of size n each
  x_{ij}= jth observation in the ith replication

Initial Data Deletion (Cont)

Steps:

1. Get a mean trajectory by averaging across replications

$$\bar{x}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij} \quad j = 1, 2, \ldots, n$$

2. Get the overall mean:

$$\bar{\bar{x}} = \frac{1}{n}\sum_{j=1}^{n} \bar{x}_j$$

Set l=1 and proceed to the next step.

Initial Data Deletion (Cont)

3. Delete the first *l* observations and get an overall mean from the remaining *n-l* values:

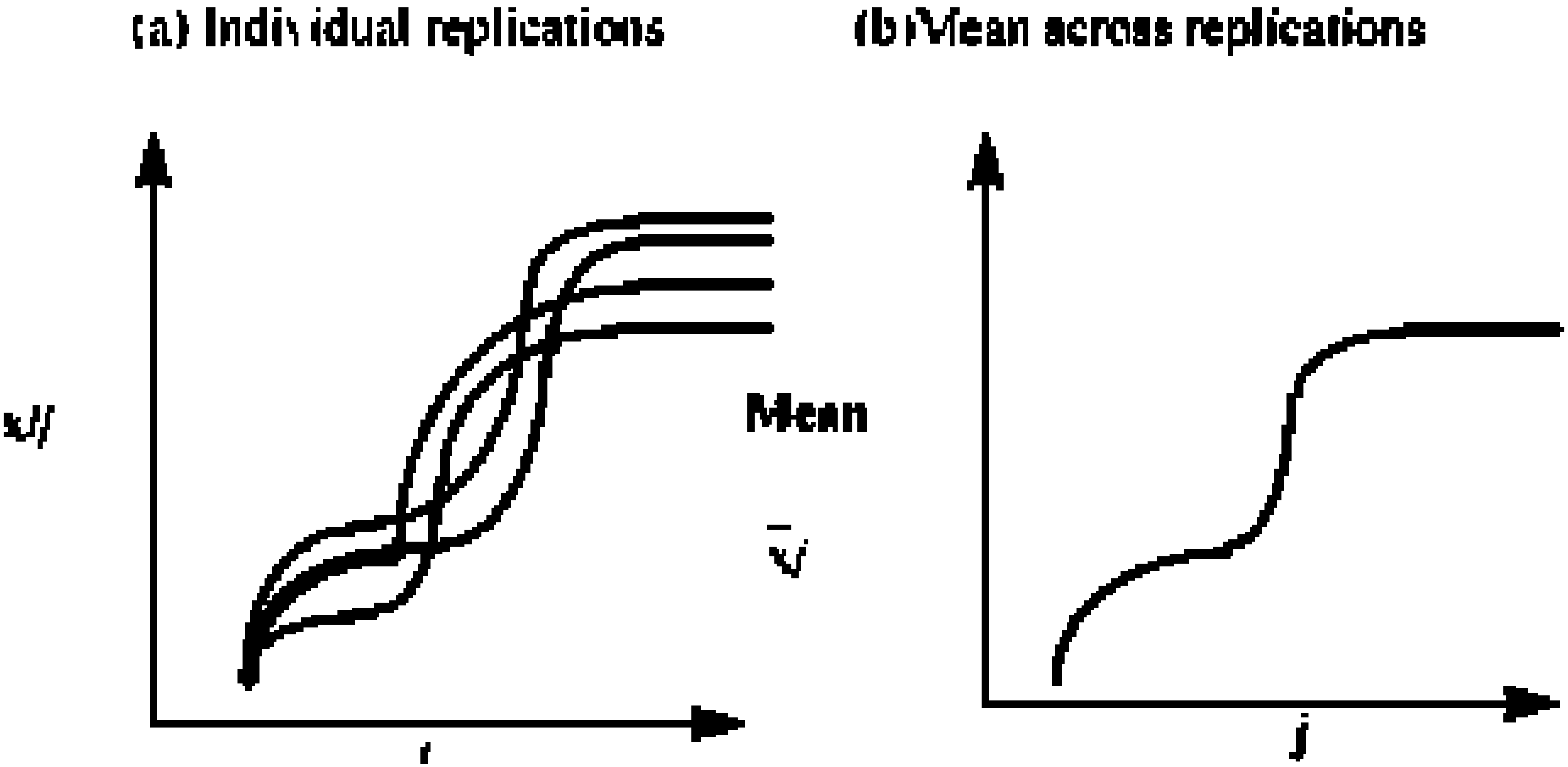
$$\bar{\bar{x}}_l = \frac{1}{n-l} \sum_{j=l+1}^{n} \bar{x}_j$$

4. Compute the relative change:

$$\text{Relative change} = \frac{\bar{\bar{x}}_l - \bar{\bar{x}}}{\bar{\bar{x}}}$$

5. Repeat steps 3 and 4 by varying *l* from 1 to *n-1*.
6. Plot the overall mean and the relative change
7. *l* at knee = length of the transient interval.

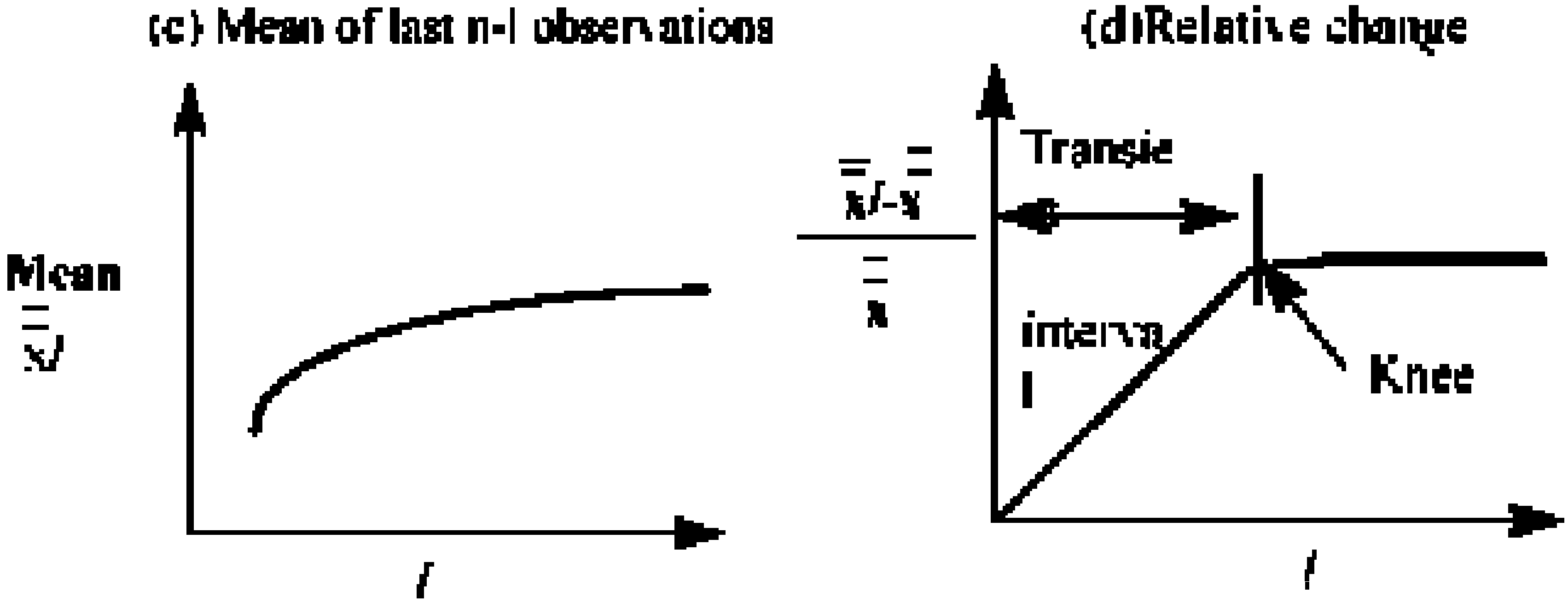
Initial Data Deletion (Cont)



(a) Individual replications

(b) Mean across replications

Initial Data Deletion (Cont)



Moving Average of Independent Replications

- Mean over a moving time interval window

1. Get a mean trajectory by averaging across replications:

$$\bar{x}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij} \quad j = 1, 2, \ldots, n$$

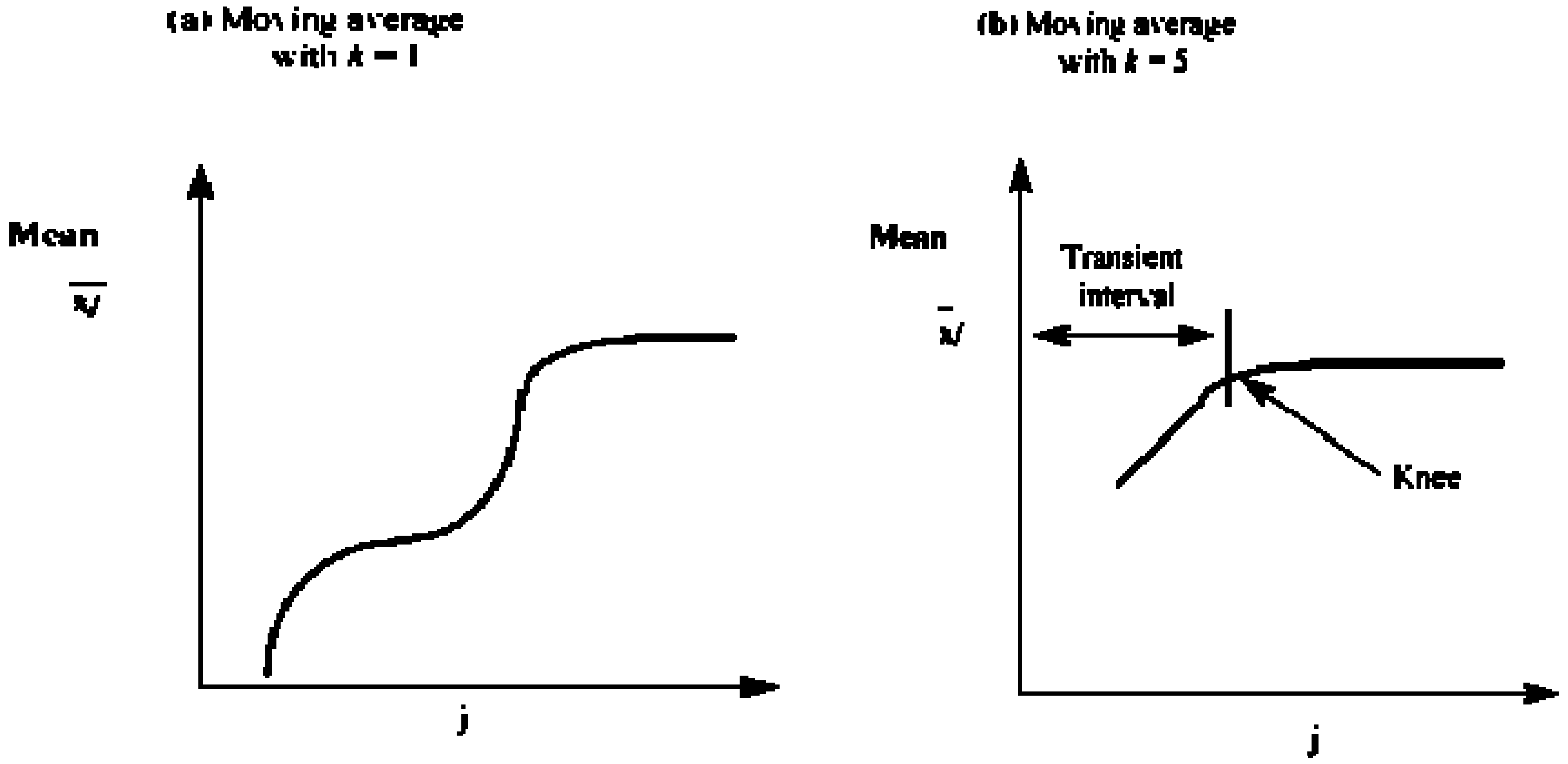
Set k = 1 and proceed to the next step.

2. Plot a trajectory of the moving average of successive 2k+1 values:

$$\bar{\bar{x}}_j = \frac{1}{2k+1}\sum_{l=-k}^{k} \bar{x}_{j+l} \quad j = k+1, k+2, ..., n-k$$

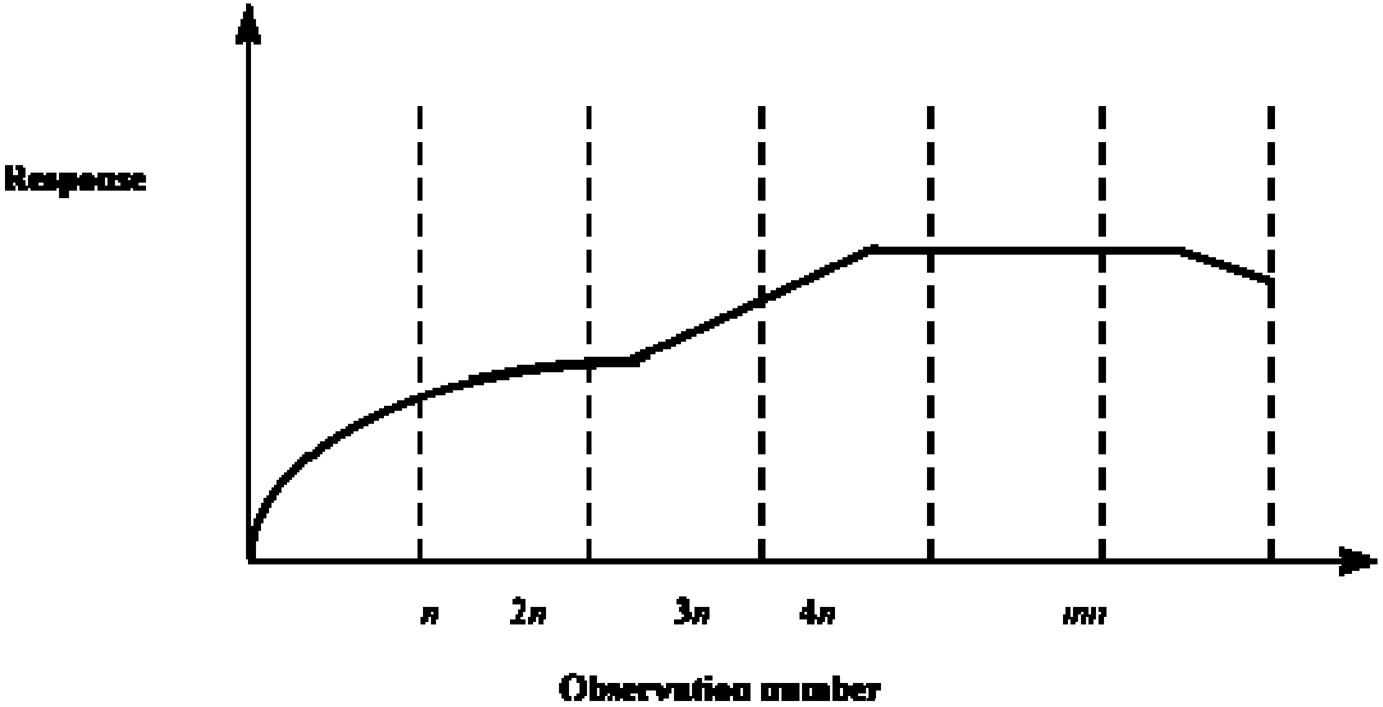
Moving Avg. of Independent Repl. (Cont)

3. Repeat step 2, with k=2, 3, and so on until the plot is smooth.
4. Value of j at the knee gives the length of the transient phase



Batch Means

- Run a long simulation and divide into equal duration part
- Part = Batch = Sub-sample
- Study variance of batch means as a function of the batch size



Batch Means (cont)

Steps:

1. For each batch, compute a batch mean:

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}, \quad i = 1, 2, \ldots m$$

2. Compute overall mean:

$$\bar{\bar{x}} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$$

3. Compute the variance of the batch means:

$$\mathrm{Var}(\bar{x}) = \frac{1}{m-1}\sum_{i=1}^{m} (\bar{x}_i - \bar{\bar{x}})^2$$

4. Repeat steps 1 and 3, for n=3, 4, 5, and so on.

Batch Means (Cont)

5. Plot the variance as a function of batch size *n*.

6. Value of *n* at which the variance definitely starts decreasing gives transient interval

7. Rationale:

   -Batch size ¿ transient

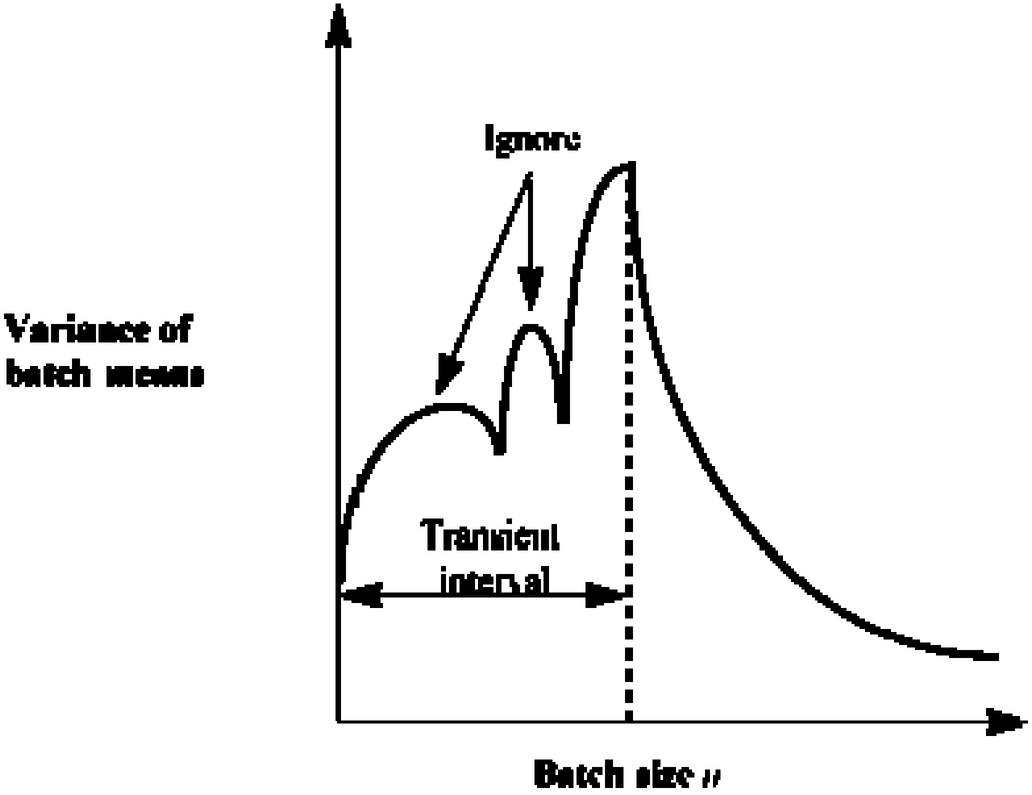
   ⇒ overall mean = initial mean ⇒ Higher variance

   -Batch size À transient

   ⇒ Overall mean = steady state mean ⇒ Lower variance

Batch Means (Cont)

- Ignore peaks followed by an upswing



Terminating Simulations

- Transient performance is of interest
  E.g., Network traffic
- System shuts down ⇒ Do not need transient removal.
- Final conditions:
  - May need to exclude the final portion from results
  - Techniques similar to transient removal

Treatment of Leftover Entities

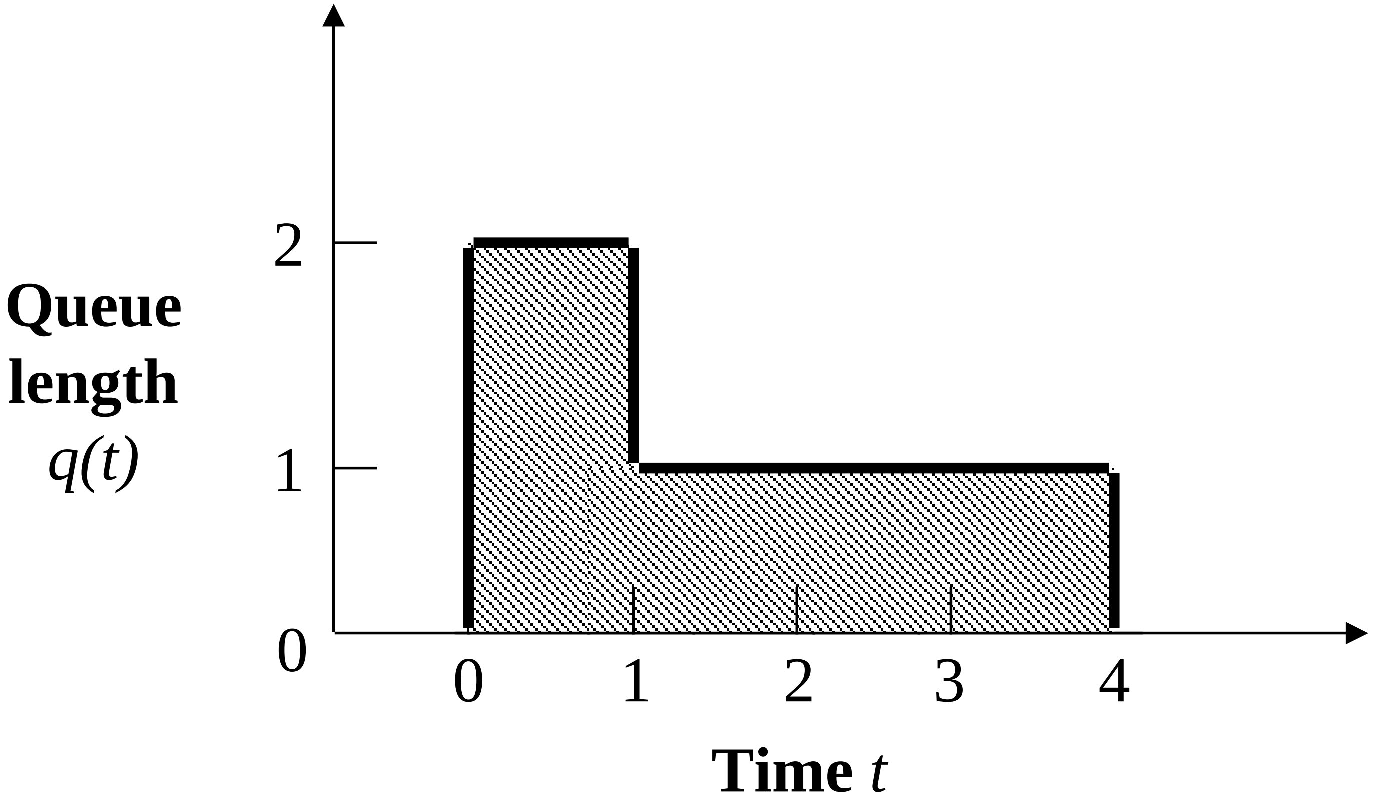
- Mean service time $= \dfrac{\text{Total service time}}{\text{Number of jobs that completed service}}$

- Mean waiting time $= \dfrac{\text{Sum of waiting time}}{\text{Number of jobs that received service}}$

- Mean Queue Length $\neq \dfrac{\sum_{j=1}^{n} \text{Queue length at event j}}{\text{Number of events} \quad n}$

$$= \frac{1}{T}\int_0^T \text{Queue\_length}(t)dt$$

Example 25.3: Treatment of Leftover Entities



- Three events: Arrival at t=0, departures at t=1 and t=4
- Q = 2, 1, 0 at these events. Avg Q ≠ (2+1+0)/3 = 1
- Avg Q = Area/4 = 5/4

Stopping Criteria: Variance Estimation

- Run until confidence interval is narrow enough

$$\bar{x} \pm z_{1-\alpha/2}\sqrt{\mathrm{Var}(\bar{x})}$$

- For Independent observations:

$$\mathrm{Var}(\bar{x}) = \frac{\mathrm{Var}(x)}{n}$$

- Independence not applicable to most simulations.
- Large waiting time for ith job
  ⇒ Large waiting time for (i+1)th job
- For correlated observations:

$$\text{Actual variance} \gg \frac{\mathrm{Var}(x)}{n}$$

Variance Estimation Methods

1. Independent Replications
2. Batch Means
3. Method of Regeneration

Independent Replications

- Assumes that means of independent replications are independent
- Conduct m replications of size n+n_0 each
  1. Compute a mean for each replication:

$$\bar{x}_i = \frac{1}{n}\sum_{j=n_0+1}^{n_0+n} x_{ij} \quad i = 1, 2, \ldots, m$$

  2. Compute an overall mean for all replications:

$$\bar{\bar{x}} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$$

Independent Replications (Cont)

3. Calculate the variance of replicate means:

$$\mathrm{Var}(\bar{x}) = \frac{1}{m-1}\sum_{i=1}^{m}(\bar{x}_i - \bar{\bar{x}})^2$$

4. Confidence interval for the mean response is:

$$\left[\bar{\bar{x}} \mp z_{1-\alpha/2}\sqrt{\mathrm{Var}(\bar{x})/m}\right]$$

- Keep replications large to avoid waste
- Ten replications generally sufficient

Batch Means

- Also called method of sub-samples
- Run a long simulation run
- Discard initial transient interval, and Divide the remaining observations run into several batches or sub-samples.

1. Compute means for each batch:

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_{ij} \quad i = 1, 2, \ldots, m$$

2. Compute an overall mean:

$$\bar{\bar{x}} = \frac{1}{m} \sum_{i=1}^{m} \bar{x}_i$$

Batch Means (Cont)

3. Calculate the variance of batch means:

$$\mathrm{Var}(\bar{x}) = \frac{1}{m-1}\sum_{i=1}^{m}(\bar{x}_i - \bar{\bar{x}})^2$$

4. Confidence interval for the mean response is:

$$\left[\bar{\bar{x}} \mp z_{1-\alpha/2}\sqrt{\mathrm{Var}(\bar{x})/m}\right]$$

- Less waste than independent replications
- Keep batches long to avoid correlation
- Check: Compute the auto-covariance of successive batch means:

$$\mathrm{Cov}(\bar{x}_i, \bar{x}_{i+1}) = \frac{1}{m-2}\sum_{i=1}^{m-1}(\bar{x}_i - \bar{\bar{x}})(\bar{x}_{i+1} - \bar{\bar{x}})$$

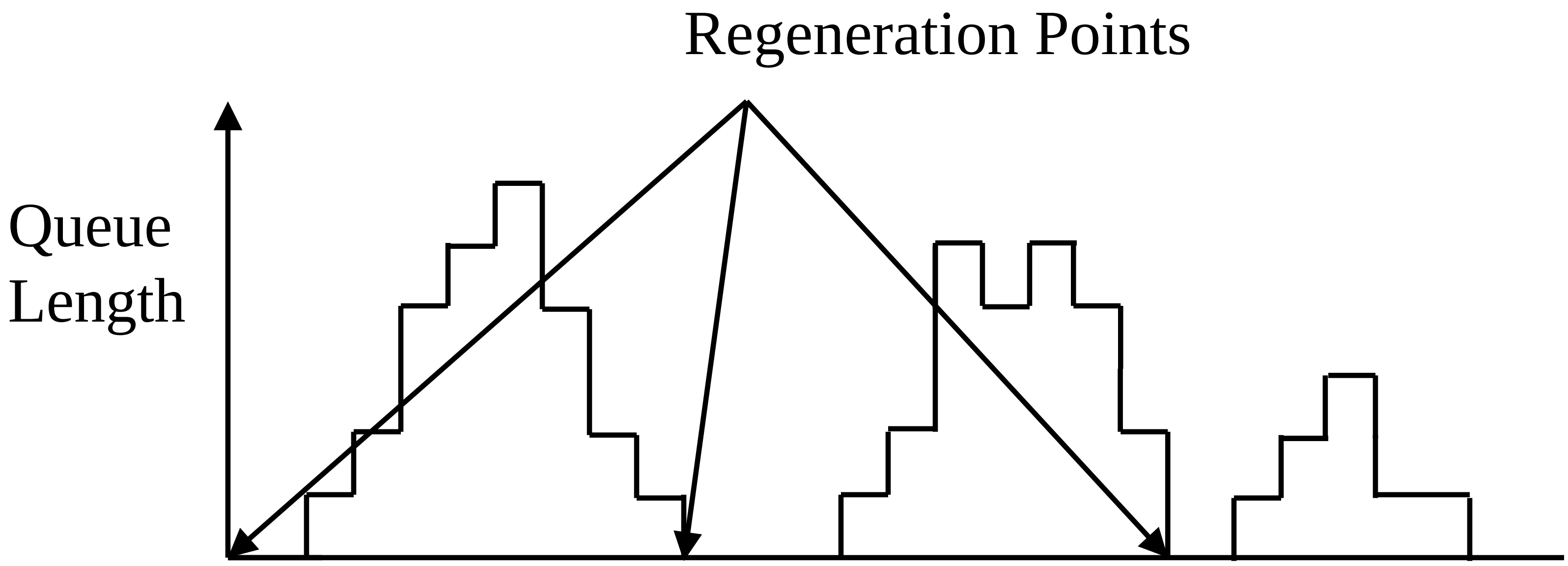
- Double n until autocovariance is small.

Case Study 25.1: Interconnection Networks

- Indirect binary n-cube networks:
  Used for processor-memory interconnection
- Two stage network with full fan out.
- At 64, autocovariance
  < 1% of sample variance

Batch Size	Autocovariance	Variance
1	-0.18792	1.79989
2	0.02643	0.81173
4	0.11024	0.42003
8	0.08979	0.26437
16	0.04001	0.17650
32	0.01108	0.10833
64	0.00010	0.06066
128	-0.00378	0.02992
256	0.00027	0.01133
512	0.00069	0.00503
1024	0.00078	0.00202

Method of Regeneration



- Behavior after idle period does not depend upon the past history
  ⇒ System takes a new birth
  ⇒ **R**egeneration point
- Note: The regeneration point are the beginning of the idle interval. (not at the ends as shown in the book).

Method of Regeneration (Cont)

- **Regeneration cycle**: Between two successive regeneration points
- Use means of regeneration cycles
- Problems:
  - Not all systems are regenerative
  - Different lengths ⇒ Computation complex
- Overall mean ≠ Average of cycle means
- Cycle means are given by:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

Method of Regeneration (Cont)

- Overall mean: $\bar{\bar{x}} \neq \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$

1. Compute cycle sums: $y_i = \sum_{j=1}^{n_i} x_{ij}$

2. Compute overall mean: $\bar{\bar{x}} = \frac{\sum_{i=1}^{m} y_i}{\sum_{i=1}^{m} n_i}$

3. Calculate the difference between expected and observed cycle sums:

$$w_i = y_i - n_i\bar{\bar{x}} \quad i = 1, 2, \ldots, m$$

Method of Regeneration (Cont)

4. Calculate the variance of the differences:

$$\mathrm{Var}(w) = s_w^2 = \frac{1}{m-1}\sum_{i=1}^{m} w_i^2$$

5. Compute mean cycle length:

$$\bar{n} = \frac{1}{m}\sum_{i=1}^{m} n_i$$

6. Confidence interval for the mean response is given by:

$$\bar{\bar{x}} \mp z_{1-\alpha/2}\frac{s_w}{\bar{n}\sqrt{m}}$$

7. No need to remove transient observations

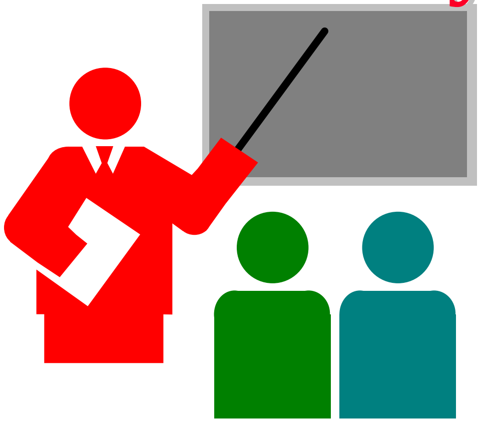
Method of Regeneration: Problems

1. The cycle lengths are unpredictable. Can't plan the simulation time beforehand.
2. Finding the regeneration point may require a lot of checking after every event.
3. Many of the variance reduction techniques can not be used due to variable length of the cycles.
4. The mean and variance estimators are biased

Variance Reduction

- Reduce variance by controlling random number streams
- Introduce correlation in successive observations
- **Problem**: Careless use may backfire and lead to increased variance.
- For statistically sophisticated analysts only
- Not recommended for beginners

Summary



1. Verification = Debugging
   ⇒ Software development techniques
2. Validation ⇒ Simulation = Real ⇒ Experts involvement
3. Transient Removal: Initial data deletion, batch means
4. Terminating Simulations = Transients are of interest
5. Stopping Criteria: Independent replications, batch means, method of regeneration
6. Variance reduction is not for novice

Exercise 25.1

Imagine that you have been called as an expert to review a simulation study. Which of the following simulation results would you consider non-intuitive and would want it carefully validated:

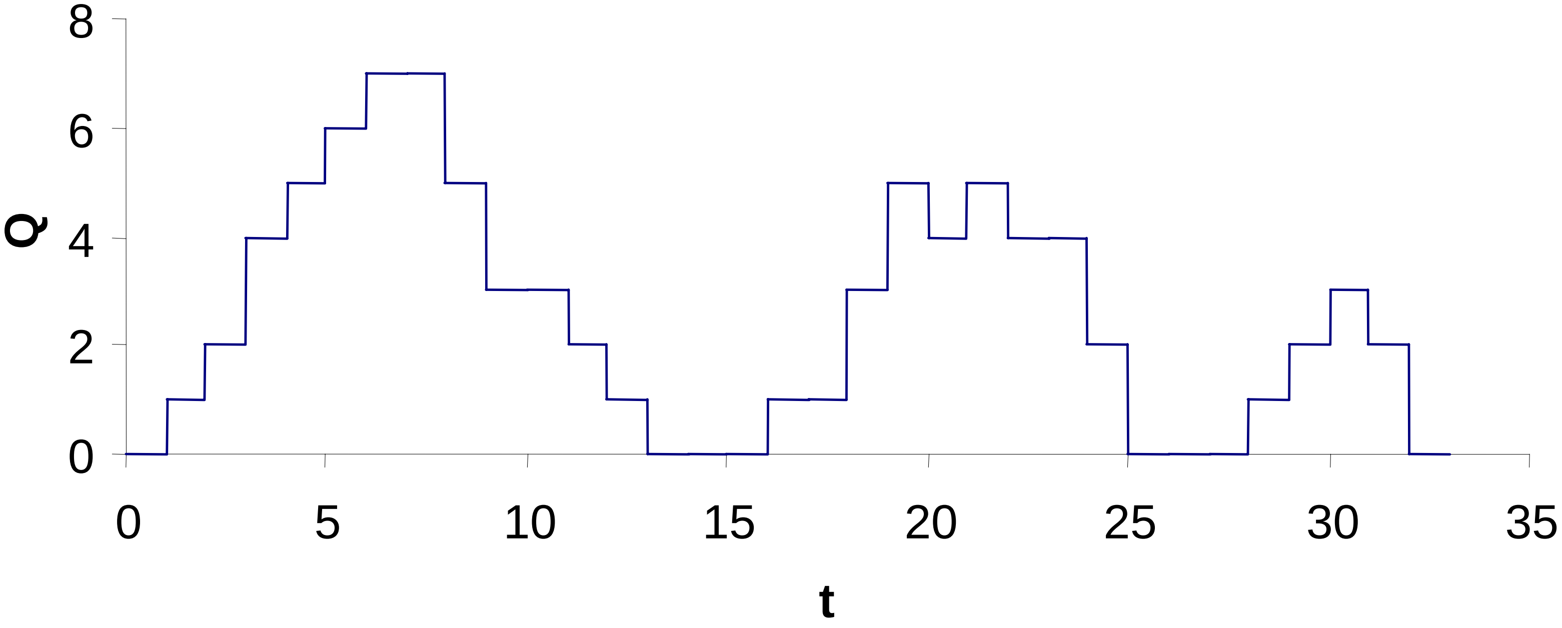
1. The throughput of a system increases as its load increases.
2. The throughput of a system decreases as its load increases.
3. The response time increases as the load increases.
4. The response time of a system decreases as its load increases.
5. The loss rate of a system decreases as the load increases.

Exercise 25.2

Find the duration of the transient interval for the following sample: 11, 4, 2, 6, 5, 7, 10, 9, 10, 9, 10, 9, 10, …, Does the method of truncation give the correct result in this case?

Homework 25

- The observed queue lengths at time t=0, 1, 2, …, 32 in a simulation are: 0, 1, 2, 4, 5, 6, 7, 7, 5, 3, 3, 2, 1, 0, 0, 0, 1, 1, 3, 5, 4, 5, 4, 4, 2, 0, 0, 0, 1, 2, 3, 2, 0. A plot of this data is shown below. Apply method of regeneration to compute the confidence interval for the mean queue length.


Q
8
6
4
2
0
0
5
10
15
20
25
30
35
t