SMA6304 M2 ---Factory Planning and scheduling

Lecture Discrete Event Simulation of Manufacturing Systems

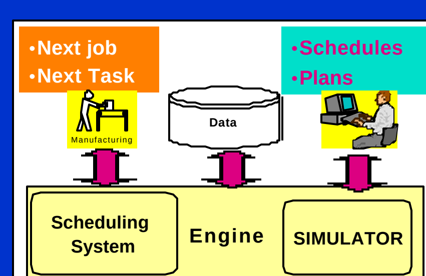
Sivakumar AI

Lecture: 12

Simulation

Simulation - A Predictive Tool

Simulation - A decision support



Simulation

Lecture (12 December 2002)

- 1. Introduction
- 2. Input probability distributions
- 3. Generating Random Numbers & Random Variates
- 4. Verification and Validation
- 5. Output data analysis
- 6. Simulation life cycle

References

- Law, A. M., Kelton, W. D. Simulation Modeling and Analysis, McGraw-Hill
- Banks, J. Handbook of Simulation, EMP

1. Introduction

The Meaning of Simulation Oxford English Dictionary

- **Simulation:**
  - The technique of imitating the behavior of some situation or system (Manufacturing, *etc*) by means of an analogous situation, model or apparatus, either to gain information more conveniently or.....
- **Simulator:**
  - An apparatus or system for reproducing the behavior of some situation or system; ....., and gives the illusion .... of behaving like the real thing.

System Modeling

Model:

- A simplified or idealized description of a system, situation, or process, often in mathematical terms, devised to facilitate calculations and predictions
- a representation of an object, system or idea in a form other than that of the entity / system itself.
- an abstraction and simplification of the real world.

System Modeling - Functions of models

- **As an analytical tool**
  - Analyze manufacturing systems
  - Evaluating equipment requirements
  - Design transportation facility
  - Ordering policy for an inventory system
- **As an aid for experimentation**
- **For planning and scheduling**
- as an aid to thought
- as an aid to communicating
- for education and training

Classification of models

- **Physical models**
  - analog models of continuous systems e.g. traffic flow.
  - iconic models e.g. pilot training simulators.
- **Analytical/Mathematical model** Most scheduling systems
  - Representing a system in terms of quantitative relationships.
- **Static simulation models**
  - Time does not play a role; e.g. Monte Carlo simulation.
- **Conventional simulation models**
  - System as it evolves over time - therefore it is dynamic
  - Have I/O and internal structure; run rather than solved.
  - Empirical
  - Stochastic, (can be deterministic for scheduling application).
- **Online simulation models**
  - As conventional; but near-real-time; useful for decision support

Classification of simulation models

- **Deterministic vs. Stochastic Simulation models**
  - If no probabilistic components then it is *deterministic*
  - If random input components used then it is *Stochastic*.
- **Continuous vs. Discrete-event Simulation models**
  - Discrete-event simulation concerns modeling a system as it evolves over time by a representation where state variables change instantaneously at the event
  - Continuous simulation covers modeling over time by a representation where state variables change continuously with respect to time (e.g. using differential equations)
- **Combined Discrete-Continuous Simulation**
  - For systems that are neither completely discrete nor completely continuos e.g arrival of a tanker and filling it.

Simulation models

Manual model generation

Conventional Simulation Model

- For analytical work
- For decisions or confirmation of decisions
- Rapid model building
- Manually intensive
- Hard to maintain

Auto model generation

Dynamic (near-real-time) Simulation Model

- Most suitable for planning and scheduling
- Uses near-real-time data
- Fully automatic
- Integrated with info. systems
- No direct maintenance

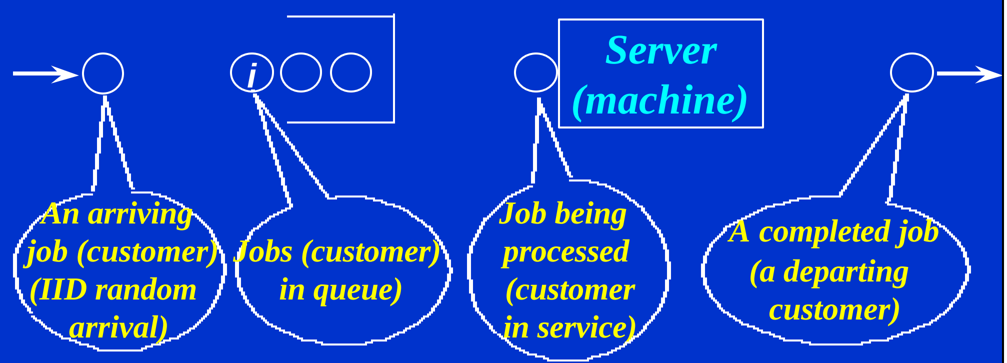
Some basic definitions

- **System state variables**
  - Collection of information needed to define what is happening in a system to a sufficient level at given point in time
- **Events**
  - *Exogenous* e.g. order arrival; *Endogenous* e.g a machine down
- **Entities and attributes**
  - *Dynamic entity* e.g. a customer, *Static entity* e.g. a machine.
  - An entity is defined by its *attributes*, e.g. quantity of a lot
- **Resources**
  - *Resource* is a static entity that provides service to dynamic entity (a lot)
- **Activity and delay**
  - *Activity* is a period whose duration is known; *Delay* is an indefinite duration caused by a combination of systems conditions

Four types of modeling structures

- **Event-Scheduling method**
  - Events are scheduled by advancing the simulation clock exactly to the time of the next event. This is one of the most accurate structures.
- **Activity Scanning method**
- **Three-Phase method**
- **Process interaction method**

Example: M/M/1 queue



Next-event time advance mechanism in Event-Scheduling method of an M/M/1 queue

t_i = Time of arrival of i^{th} customer

$A_i = t_i - t_{i-1}$ = Inter arrival time = IID random variables

S_i = Service time of i^{th} customer= IID ran. variable

D_i = Observed delay of i^{th} customer in queue

$c_i = t_i + D_i + S_i$ = Completion time of i^{th} customer

e_i = Time of occurrence of i^{th} event of any type

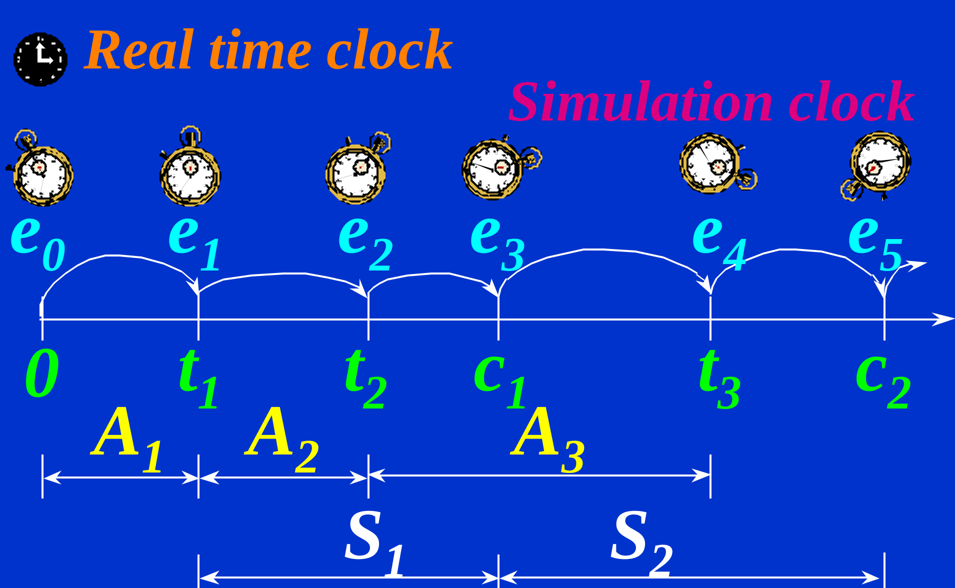
B_t (or ■) = "Busy function" defined as 1 if server is busy at time t and 0 if server is idle at time t

Next-event time advance mechanism in Event-Scheduling method of an M/M/1 queue

- Inter arrival times A_i, and service times S_i have cumulative distribution functions F_a and F_s, which are determined by collections of actual past data and fitting distributions. *(2. Input probability distributions)*

- Each value of t_i is computed using generated values of A_i, using random observations from a specific distribution. *(3. Generating Random Numbers and Random Variates)*

Next-event time advance mechanism in Event-Scheduling method of an M/M/1 queue



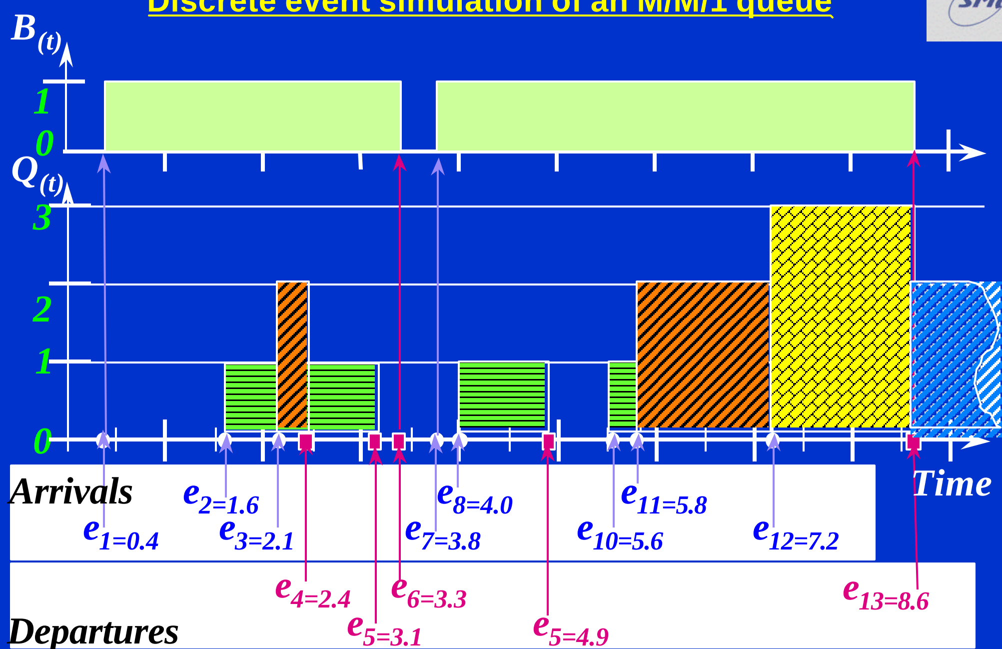
Next-event time advance mechanism in Event-Scheduling method of an M/M/1 queue

- The simulation clock is advanced from each event to the next event based on the event times of the event list.
- A machine starting to process is an ***activity*** and the end time is known from S_i & F_a
- When a customer (or a job) arrives and if the server is busy then it joins a queue and the end time is unknown. This is a ***delay***

Discrete event simulation of an M/M/1 queue

Job	Arrival t_i	Service S_i	Start c_i	End	Delay D_i
J1	0.4	2.0	0.4	2.4	0
J2	1.6	0.7	2.4	3.1	0.8
J3	2.1	0.2	3.1	3.3	1.0
J4	3.8	1.1	3.8	4.9	0
J5	4.0	3.7	4.9	8.6	0.9
J6	5.6	2.8			
J7	5.8	1.6			
J8	7.2	3.1			

Discrete event simulation of an M/M/1 queue



 19

Expected average delay in M/M/1 queue

From a single run of simulation of n jobs (customers), a point estimate for $\overline{D}(n)$, expected average delay in queue of the n jobs (customers) is

$$\overline{D}(n) = \frac{\sum_{i=1}^{n} D_i}{n}$$

 20

Expected average number of jobs in M/M/1 queue

T_i = Total time during the simulation in which the queue of customers(jobs) is observed as length i

T_n = Time to observe n delays in queue

$Q_{(t)}$ = Number of customers in queue at time t

q_n = Average number of customers(jobs) in queue during n observations

$$q(n) = \frac{\sum_{i=1}^{\infty} iT_i}{T_n}$$

Expected average number of jobs in M/M/1 queue

Since

$$\sum_{i=1}^{\infty} iT_i = \int_0^{T(n)} Q(t)dt$$

We get

$$q(n) = \frac{\int_0^{T(n)} Q(t)dt}{T_n}$$

Expected utilization of the machine in M/M/1 system

u_n = Expected proportion of time the server (machine) is busy during n observations

Since B_t is always either 0 or 1

$$u(n) = \frac{\int_0^{T(n)} B(t)\,dt}{T\,n}$$

Simple output from discrete event simulation

Let take a case of the earlier table of arrival and service with the first 5 observation of completion (T_n =8.6)

$$\sum_{i=1}^{\infty} i T_i = (0 \times 3.2) + (1 \times 2.3) + (2 \times 1.7) + (3 \times 1.4) = 9.9$$

$T_i = 0$ for $i >= 4$

and therefore $q(n)$ = 9.9/8.6 = 1.15

and

$u(n)$ = [(3.3 -0.4) + (8.6-3.8)]/8.6 = 0.90

Discrete event simulation

- These values are simple illustrations of statistics of discrete event simulation
  - Discrete-time statistics (e.g. average delay in queue) or
  - Continuous-time statistics (e.g. proportion of server busy time)
- A very large number of other useful statistics could be obtained from each simulation run.
- In addition very complex manufacturing systems could be modeled using this simple approach.

Discrete event simulation

- However model ***MUST*** be verified and validated to add credibility to the results. *(4. verification and validation)*
- Experimental runs should then be carried out using the validated model.

Discrete event simulation

- However values of each experimental run are based on "sample" size of 1 (one complete simulation run) and sample size of 1 is not statistically useful.
- Multiple replications and confidence interval are therefore essential elements of simulation output data analysis. *(5. Output data analysis)*

Probability & statistics and Simulation

- Probability and statistics are integral part of simulation study
- Need to understand how to model a probabilistic system
- Validate a simulation model
- Input probability distributions,
- Generate / use random samples from these distributions
- Perform statistical analysis of output data
- Design the simulation experiments.

2. Input probability distributions

Randomness in Manufacturing

- Process time
- MTTR
- MTTF
- Inter arrival time
- Job types or part mix
- Yield
- Rework
- Transport time
- Setup time
- and so on

Using past data

- Use past data directly in simulation. This is ***trace driven simulation***. *(Effective for model validation)*
- Use the sample data to define an ***empirical distribution function*** and sample the required input data from this distribution.
- Use a standard technique (e.g. regression) to fit a ***theoretical distribution form to*** the sample data (e.g. exponential etc.), and sample the required data from it.

Steps in selecting Input probability distributions

- Assess sample independence:
  - Must confirm the observations X1,X2,…Xn are independent using techniques such as correlation plot.
- Hypothesizing families of distributions:
  - Without concern for specific parameters, we must select general family e.g. normal, exponential etc.
- Estimation of parameters:
  - Use the past numerical data to *estimate* parameters.
- Determine the best fit:
  - Use a technique such as probability plot or chi-square test and identify the most suitable distribution function
- **The last three steps are integral part of available software and therefore we may not have to manually carryout these steps**

3. Generating Random Numbers and Random Variates

Status

- Early simulation studies required random number generation and generation of random variates from the distributions, often manually coded in computers.
- Most of the current simulation languages and simulators have built in features for this.

Random number generation for simulation

- Built in feature should have the following:
- Generate random numbers, uniformly distributed on U[0,1] that do not exhibit any correlation with each other.
- Reproduce a given *stream* of random numbers exactly (i.e. Identical random numbers) for verification etc.
- Have ability to generate large number of streams for *multiple replications* (i.e. Different streams are separate and independent generators.)

Random variate generation for simulation

- Built in feature should also have the following:
- Generate random variates.
- This means: Produce observations for each selected variable (e.g. MTTR) from the parameters of the desired distribution function (e.g. Gamma) using the IID U(0,1) random numbers with computational efficiency.

4. Verification and Validation

Definition

- **Model verification**: Building model right
  - Correct translation of conceptual simulation model in to a working program
  - Debugging
- **Model validation**: Building the right model
  - Determine if the conceptual simulation model is an accurate representation.
- **Credible Model**: Objectives fulfilled using model
  - When the model is accepted by user / client and used for the purpose it was built

Verification

- Errors arise from data, conceptual model, computer model, even computer system.
- Test sub-models first, then complete model.
- Common techniques
  - *Static*: a structured, walk-through technique.
  - *Dynamic*: run program under different conditions, then check if the output is reasonable.
  - *Trace*: Identify selected state variables (event list) after each event and check with manual calculations
  - *Animation*: observe animation

What is Validation?

- Valid if it's output behavior is sufficiently accurate to fulfill the purpose. Absolute accuracy is not essential and it is too time-consuming and expensive.
  - check underlying theories, assumptions, approximation.
  - check model structure and logic, math and relationships (by tracing entities in all sub-models and main model).
  - Model should be validated relative to the measures in the objective.

Validation

- **Data have to be validated:**
  - –difficult, time-consuming and costly to obtain relevant, sufficient and most importantly consistent and accurate factory data.

A three step Validation process

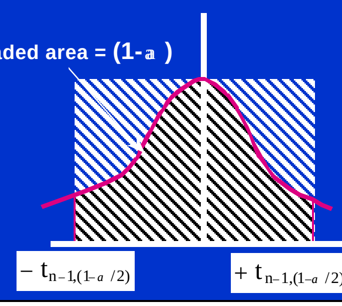
- **Conventional simulation studies :**
  - –Step 1. Face validation: ask people knowledgeable / experienced about the system under study
  - –Step 2. Empirically Test & compare with other models, e.g. analytical models
  - –Step 3. Detail output data validation
    - –(a) Confidence Intervals
    - –(b) Correlated Inspection Approach

Confidence Intervals

- Let $Y_1, Y_2, \ldots Y_n$, be IID random variables with mean $\bar{y}_n$ and sample variance $s^2_{(n)}$
- It can be shown using ***central limit theorem*** that, when n is "sufficiently large", that an approximate $100(1-\alpha)$ percent confidence interval (assuming a *t* distribution with (n-1) degrees of freedom) is given by the following:

$$l(n,\alpha) = \bar{Y}_{(n)} - t_{n-1,(1-\alpha/2)}\sqrt{\frac{S^2(n)}{n}}$$

$$u(n,\alpha) = \bar{Y}_{(n)} + t_{n-1,(1-\alpha/2)}\sqrt{\frac{S^2(n)}{n}}$$



Validation: Confidence Intervals approach

- Assume we collect *m* independent sets of data from the ***system*** and *n* independent sets of data from the ***model***.
- Let X_j be the average of the observations of a desired variable (e.g. throughput) in the j^{th} set of system data and U_j be the average of the observations in the j^{th} set of model data of the same variable.

Validation: Confidence Intervals approach

- If we assume the *m* sets of system data are homogeneous, then X_j's are IID random variables with the mean μ_x.
- If the *n* sets of simulation model data were generated using ***independent replications*** then U_j's are IID random variables with the mean μ_y.

Validation: Confidence Intervals approach

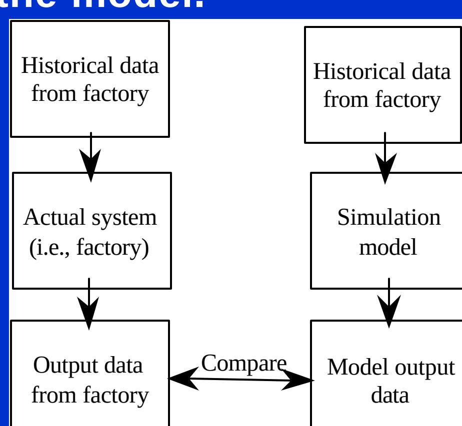
- One of the methods to compare the model with the system is by constructing a confidence interval for x where $x = m_x - m_y$
- Let $l(n, a)$ and $u(n, a)$ correspond to lower & upper confidence interval end points of x .
- If $0 \notin [l(a), u(a)]$, then the observed difference between m_x and m_y is said to be statistically significant at level a

Validation: Confidence Intervals approach

- However, even though the difference is statistically significant, the model may still be a valid representation of the system to fulfill the objectives of the simulation.
- On the other hand if $0 \in [l(a), u(a)]$, then the observed difference between m_x and m_y is said to be statistically not significant at level a and may be explained by sampling fluctuations.

Validation : Correlated Inspection Approach

- Statistics of the desired variable from the systems is compared with corresponding statistics of the model.



Validation : Correlated Inspection Approach

- Suppose we want to validate the cycle time (e.g.makespan / duration from arrival to completion).
- We make a number of observation of the factory cycle time X_j's and for example, the inter arrival time of the jobs.
- We then use the observed values of the inter arrival time of the factory to drive the simulation model and obtain cycle time values U_j's from it.

Validation : Correlated Inspection Approach

- We compare j^{th} set of system data X_j and simulation model data U_j where $X_j - Y_j$ n estimate for $m_x - m_y$
- We can look at the sample mean and sample variance of all to $m_x - m_y$ udgment on the validity of the simulation model to fulfill the objectives.

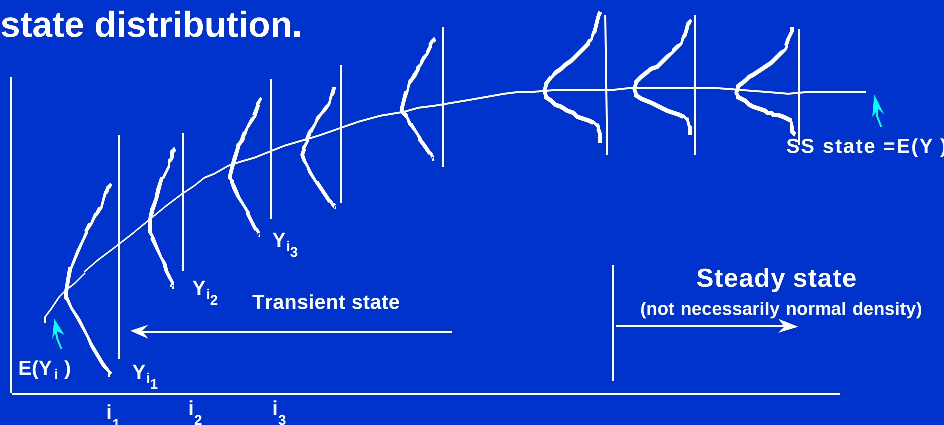
5. Output data analysis

Transient and steady state behavior

- Experimental observations for output analysis should be made at steady state, i.e. after a transient state of the stochastic simulation run.
- Consider the output of the random variable U_i for i=1,2,…m.
- Let $F_i(y|I) = P(Y_i \le y|I)$ re y is a real number and I represents ***initial*** condition.
- The ***transient distribution*** $F_i(y|I)$ screte time i is different for each i and each set of I.

Transient and steady state behavior

- If $F_i(y|I) \to F(y)$ $i \to \infty$ ll y and any Initial condition I then is sai $F(y)$ steady state distribution.



53

Random nature of simulation output

- Let U_1, U_2, ... U_i ... U_m, be output of stochastic process from a ***single simulation*** run (e.g. U_i is throughput at i^{th} hour, *m* = number of observations).
- U_i's are random variables & generally <u>not IID</u>.
- If we carry out *n* independent replication (using *n* different streams) then we realize a set of random observations as shown below:

y_{11}	,	.	.	.	,	y_{1i}	,	.	.	.	y_{1m}
y_{21}	,	.	.	.	,	y_{2i}	,	.	.	.	y_{2m}
.	.	.	.	.	.	.	.	.	.	.	.
y_{n1}	,	.	.	.	,	y_{ni}	,	.	.	.	y_{nm}

54

Output analysis

- Observations from a single replication (row) are not IID.
- However $y_{1i}, y_{2i}, \ldots y_{ni}$ from *n* replications (column) are IID observations of the random variable, U_i for i=1,2,…m.
- This is the basis of statistical analysis of the observations of y_{ji}. For example an unbiased estimate of $E(Y_i)$

$$\bar{y}_i(n) = \frac{\sum_{j=1}^{n} y_{ji}}{n}$$

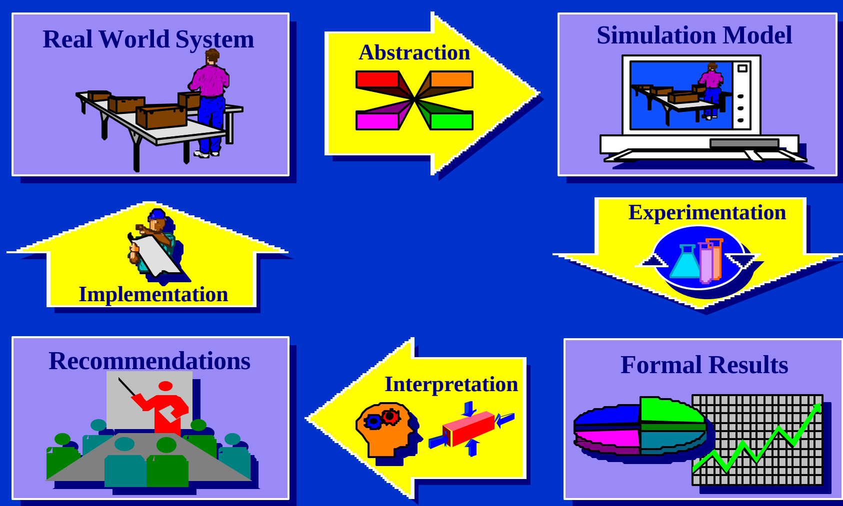
Confidence Interval of Output analysis - an example

- Comparing two systems on a given measure of performance is done by forming a confidence interval for the difference in two expected values of for example $E(Y_i)$ and check if $0 \in$.
- If the number of observations n1=n2=n then we pair Y_{1j} with Y_{2j} and define $Z_j = Y_{1j} - Y_{2j}$ and Z_j's are the IID random variables. Approximate $100(1-a)$ percent confidence interval

$$\bar{Z}_{(n)} \pm t_{n-1,(1-a/2)} \sqrt{\frac{S^2(n)}{n}}$$

Paired-t confidence interval

Steps in conventional simulation



Credibility Assessment in Simulation projects

Credibility Assessment Stages

Quality Control of processes between phases

- Credibility assessment will always be subjective because
  - modeling is an art
  - credibility assessment is situation-dependent
- Accuracy of assessment always relative to the objective of simulation, never absolute.

Credibility Assessment Stages

Quality Control of processes between phases

- Peer Assessment
  - Panel of persons who are
    - experts of the system under study
    - expert modelers
    - simulation analysts
    - familiar with simulation projects

Credibility Assessment Stages

Quality Control of processes between phases

- Verify formulated problem
  - to make sure it faithfully reflects the real problem.
- Feasibility of simulation
  - is data available? Easy or costly to get?
  - are resources for simulation available?
  - cost-benefit: any time limit imposed to complete study?
- The real system
  - are system's boundaries well-defined?
  - have objectives of simulation changed with time?
  - counter-intuitive behavior accounted for?
  - any drift to low performance?

Credibility Assessment Stages

Quality Control of processes between phases

- Qualifying the conceptual model
  - are assumptions explicitly defined, appropriate?
- Verifying the communicative model
  - use techniques such as walk-through, structural analysis, data-flow analysis (Whitner & Balci, 1986).
- Verifying the programmed model
  - use standard software verification techniques.

Credibility Assessment Stages

Quality Control of processes between phases

- Verify experimental design
  - is random number generator accurate and true?
  - are statistical techniques for design and analysis of experiments appropriate?
  - initial transients accounted for?
  - have you ensured identical experimental conditions for each alternative operating policy?
- Data validation (of model parameters and input data)
  - are they appropriate? current? unbiased? inter-dependent? complete? accurate?
  - are instruments for data measurement and collection accurate?

Credibility Assessment Stages

Quality Control of processes between phases

- Validating the Experimental Model
  - always compare the behavior of the model and real system under *identical input conditions*.
  - subjective and statistical validation techniques applicable only when data are completely *observable*.
- Interpretation of simulation results
  - interpret numerical results based on the objective of the study. Judgment involved.
- Documentation
  - Embed documentation into model development cycle.

Credibility Assessment Stages

Quality Control of processes between phases

- **Presentation**
  - communicating simulation results
    - translate the jargon so non-simulation people & decision-makers can understand
  - presentation techniques
    - integrate simulation results with a DSS, so decision-maker can appreciate the significance of the simulation results.

Wrap-up

- We have looked at the simulation of a M/M/1 queue
- We have discussed Input probability distributions
- We talked about the Random Numbers and Random Variates
- We discussed Validation techniques
- We outlined output data analysis and confidence intervals
- Life cycle and Credibility Assessment of simulation models