Scheme of Evaluation
Internal Assessment Test 2 – April.2019

Sub:	System Modeling and Simulation						**Code:**	15CS834
Date:	20/04/2019	Duration:	90mins	Max Marks:	50	**Sem:** VIII	**Branch:**	ISE

Note: Answer Any Five Questions

Question #		Description	Marks Distribution		Max Marks
1		• Diagram • Explanation	4 M 6 M	10 M	10 M
2		• Diagram • Explanation	4 M 6M	10 M	10 M
3	a)	• Xi = (a.X0+c) mod m • Generating random numbers • Finding and Justifying the period	1 M 3 M 1 M	5 M	10 M
	b)	• Xi = (a.X0+c) mod m • Ri=Xi/m • Generating random numbers	1M 1M 3M	5M	
4		• Finding E value, E=N/n • Chi-Square test computational table • Justification of acceptance/rejection	2 M 7 M 1 M	10 M	10 M
5	a)	• Rank the data from smallest to largest. Finding the R_i • Finding the D^+ and D^- values • Finding D value • Justification of the data accept or reject	1 M 3 M 1M 1M	6 M	10 M
	b)	• Finding the CDF F(x) • Set F(x)=R and solving the equation in terms of X	1M 3M	4M	
6	a)	• Set n=0,p=1. • Taking the first random number and finding p value • Comparing p with e^{-a}, if p< e^{-a} accept N=0,else reject and go to step2 Repeating the above steps for all random numbers and finding acceptance/rejection	3 M 4M	7 M	10 M

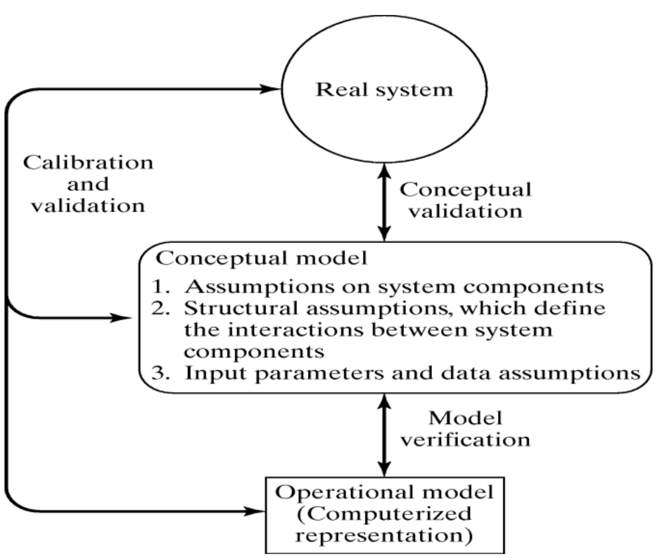
	b)	• Defining acceptance rejection technique • Explanation	1M 2M	3M	
7		Output analysis for terminating simulations. Explanation of all the below three • Statistical background • Within and across replication data • Confidence interval	10 M	10 M	10 M
8		• Finding the M value i.e. i+(M+1)l <= N • Finding the ρ_{im} value i.e. ρ_{im} =1/M+1[Ri+km *Ri+(k+1)m] -0.25 • Finding the $\sigma_{\rho im} =$ □13M+7 / 12(M+1) value • Justification of acceptance	2 M 4 M 2M 2M	10 M	

Internal Assessment Test 2 Solutions– April.2019

Sub:	System Modeling and Simulation							**Code:**	15CS834
Date**:**	20/04/2019	Duration:	90mins	Max Marks:	50	**Sem:**	VIII	**Branch:**	ISE

Note: Answer Any Five Questions

1. **Explain the components of verification and validation process. Explain with neat diagram model building, verification and validation process.**



The first step in model building consists of observing the real system and the interactions among its various components and collecting data on its behavior.

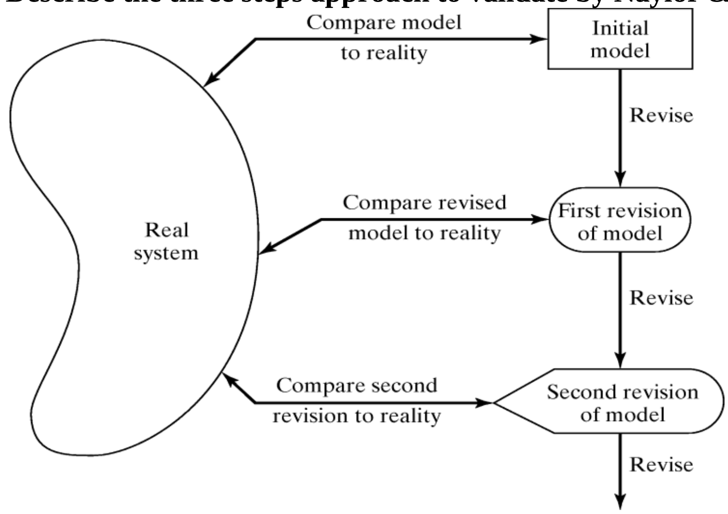
• Operators, technicians, repair and maintenance personnel, engineers, supervisors, and managers under certain aspects of the system which may be unfamiliar to others.

• As model development proceeds, new questions may arise, and the model developers will return, to this step of learning true system structure and behavior.

• The second step in model building is the construction of a conceptual model – a collection of assumptions on the components and the structure of the system, plus hypotheses on the values of model input parameters, illustrated by the following figure.

• The third step is the translation of the operational model into a computer recognizable form- the computerized model.

2. **Describe the three steps approach to validate by Naylor & Finger in the validation process.**



Naylor and Finger formulated a **three step approach** which has been widely followed:-

1. Build a model that has high face validity.
2. Validate model assumptions.
3. Compare the model input-output transformations to corresponding input-output transformations for the real system.

1. Face Validity

• The first goal of the simulation modeler is to construct a model that appears reasonable on its face to model users and others who are knowledgeable about the real system being simulated.

• The users of a model should be involved in model construction from its conceptualization to its implementation to ensure that a high degree of realism is built into the model through reasonable assumptions regarding system structure, and reliable data.

• Another advantage of user involvement is the increase in the models perceived validity or credibility without which manager will not be willing to trust simulation results as the basis for decision making.

• Sensitivity analysis can also be used to check model's face validity.

• The model user is asked if the model behaves in the expected way when one or more input variables is changed.

• Based on experience and observations on the real system the model user and model builder would probably have some notion at least of the direction of change in model output when an input variable is increased or decreased.

• The model builder must attempt to choose the most critical input variables for testing if it is too expensive or time consuming to: vary all input variables.

2. Validation of Model Assumptions

• Model assumptions fall into two general classes: structural assumptions and data assumptions.

Structural assumptions involve questions of how the system operates and usually involve simplification and abstractions of reality.

• For example, consider the customer queuing and service facility in a bank. Customers may form one line, or there may be an individual line for each teller. If there are many lines, customers may be served strictly on a first-come, first-served basis, or some customers may change lines if one is moving faster. The number of tellers may be fixed or variable. These structural assumptions should be verified by actual observation during appropriate time periods together with discussions with managers and tellers regarding bank policies and actual implementation of these policies.

• Data assumptions should be based on the collection of reliable data and correct statistical analysis of the data.

3. Validating Input-Output Transformation

In this phase of validation process the model is viewed as input –output transformation : That is, the model accepts the values of input parameters and transforms these inputs into output measure of performance. It is this correspondence that is being validated.

• Using historical input data : Instead of validating the model input-output transformation by predicting the future, the modeler may use past historical data which has been served for validation purposes that is, if one set has been used to develop calibrate the model, its recommended that a separate data test be used as final validation test.

Using Turing test: When no statistical test is readily applicable then persons knowledgeable about system behavior can be used to compare model output with system output. This type of test is called Turing test used in detecting model inadequacies and to increase the model credibility.

3.a) Given X0=1, m=64 and a=13. Using suitable technique find the series of random numbers, find the period for the same and justify whether maximum period is achieved or not.

Sol : Based on the above three conditions for finding the period, since m=64 = 2power6 so second condition satisfies.

Period = m/4 = 64/4 = 16

X0=1

X1= (13*1)mod 64 = 13

X2=(13*13)mod 64 = 41

X3=(13*41)mod 64 = 21

X4=(13*21)mod 64 = 17

………………

X16 = (13*5)mod 64 = 1

Hence maximum period is achieved at X16 since 16^{th} value is same as initial value.

3.b)Use the suitable method to generate a sequence of random numbers with X0 = 27, a= 17, c = 43, and m = 100

Sol: The sequence of Xi and subsequent Ri values is computed as follows:

X0 = 27

X1 = (a.X0+c) mod m = (17.27 + 43) mod 100 = 502 mod 100 = 2

R1=X1/m = 2⁄100=0. 02

X2 = (17 • 2 + 43) mod 100 = 77 mod 100 = 77

R2=77 ⁄100=0. 77

X3 = (17•77+ 43) mod 100 = 1352 mod 100 = 52
R3=52/100=0. 52

4. Use the Chi-square test with α=0.05,n=10,$x^2_{0.05.9}$ =16.9 to test for whether the data shown are uniformly distributed.

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.70	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.30	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.10	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.06	0.43	0.56	0.97	0.30	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.40	0.64	0.40	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.60	0.11	0.29	0.78

Sol:

- Dividing into no of intervals
- Finding Ei value

  Ei= N/n = 100/10 =10
- Chi-Square test computation table

Interval No	Interval range	Oi	Ei	Oi-Ei	(Oi-Ei)²	(Oi-Ei)²/Ei
1	0-0.1	8	10	-2	4	0.4
2	0.1-0.2	8	10	-2	4	0.4
3	0.2-0.3	10	10	0	0	0
4	0.3-0.4	9	10	-1	1	0.1
5	0.4-0.5	12	10	2	4	0.4
6	0.5-0.6	8	10	-2	4	0.4
7	0.6-0.7	10	10	0	0	0
8	0.7-.0.8	14	10	4	16	1.6
9	0.8-0.9	10	10	0	0	0
10	0.9-1.0	11	10	1	1	0.1
Total						**3.4**

- Comparing with α value and justifying for acceptance/rejection

  i.e.3.4 < 16.9 so the above values are accepted

5a) Using suitable frequency test find out whether the random numbers generated are uniformly distributed on the interval [0, 1] can be rejected. Assume α=0.05 and Dα=0.565.the random numbers are 0.44, 0.81, 0.14, 0.05, 0.93.

- Rank the data from smallest to largest. Finding the R_i

  i.e. R_i 0.05 0.14 0.44 0.81 0.93

- Finding the D^+ and D^- values

  i.e. $D^+ = i/N - R_i$

Ri =	0.05	0.14	0.44	0.81	0.93
i/N =	0.20	0.40	0.60	0.80	1.00
i/N- Ri =	0.15	0.26	0.16	-	0.07
Ri-(i-1)/N =	0.05	-	0.04	0.21	0.13

$D^+ = \max(i/N - R_i) = 0.26$
$D^- = \max(R_i - (i-1)/N) = 0.21$

- Finding D value

  i.e. $D = max(D^+, D^-) = 0.26$

- Justification of the data accept or reject

  i.e. The tabular value $D_{\alpha, n}$= 0.565.Since $D < D_\alpha$ i.e 0.26 < 0.565 the sequence of numbers given are accepted.

b)Explain inverse transform technique for exponential distribution

1.Compute the CDF, i.e. F(x) = 1-e∞x

2. Set F(x)=R in the range of X

1-e∞x = R

3. Solve F(x) in terms of X

1-e∞x = R

-∞x = ln(1-R)

X=-1/∞ln(1-R)

X is called random variate

6a)Using suitable technique generate three Poisson variates with mean 4 by taking these random numbers 0.4357, 0.4146, 0.8353, 0.9952, 0.8004, 0.7945 and 0.1530

1. Set n=0,p=1.
2. R1=0.4357,P=1*0.4357 = 0.4357
3. Since P=0.4357 > e^{-4} = 0.0183, reject n=0 and return to step 2 with n=1.
2. R2=0.4146 , p= 0.4146 * 0.4357 =0.1806
3. Since 0.1806 > 0.0183, reject n=1 and return to step 2 with n=2.
2. R3= 0.8353, p=0.8353*0.1806 = 0.1508
3. 0.1508 > 0.0183, reject n=2 and return to step 2 with n=3.
2. R4= 0.9952, p=0.9952*0.1508 = 0.1502
3. 0.1502 > 0.0183, reject n=3 and return to step 2 with n=4.
2. R5= 0.8004, p= 0.8004 * 0.1502 = 0.1202
3. 0.1202 > 0.0183,reject n=4 and return to step 2 with n=5.
2. R6= 0.7945, p=0.7945 * 0.1202 = 0.0955
3. 0.0955 > 0.0183, reject n=4 and return to step 2 with n=6.
2. R7= 0.1530, p=0.1530 * 0.0955 = 0.0146
3. 0.0146 < 0.0183, accept N=6

The variates are as follows.

n	Rn+1	p	accept/reject	Result
0	0.4357	0.4357	reject	n=1
1	0.4146	0.1806	reject	n=2
2	0.8353	0.1508	reject	n=3
3	0.9952	0.1502	reject	n=4
4	0.8004	0.1202	reject	n=5
5	0.7945	0.0955	reject	n=6
6	0.1530	0.0146	accept	N=6

b) What is acceptance-rejection technique? Explain

suppose we need a method to generate random variates b/n 1/4 and 1 the following steps has to be followed.

1. generate a random no R

2a). If $R \geq 1/4$ accept $x = R$ then go to step 3

2b) If $R < 1/4$ reject R and return to step 1

3. If another uniform random variate on $[1/4, 1]$ is needed repeat the procedure begining at step 1. If not, stop.

-Each time step 1 is executed a new random no is generated. step 2a is "acceptance" and step 2b is a "rejection" technique.

7. Explain output analysis for terminating simulations in detail.

Output analysis for terminating simulations

A terminating simulation runs over a simulated time interval $[0, T_E]$

- A common goal is to estimate

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

$$\phi = E\left[\frac{1}{T_E}\int_0^{T_E} Y(t)\,dt\right] \quad \text{for continuous o/p } Y(t),\ 0 \le t \le T_E.$$

- In general independent replications are used, each run using a different random no stream and independently choosen initial conditions.

Statistical background :-

The most confusing aspect among simulation o/p analysis is distinguishing within-replication data from across-replication data.

- For ex simulation of a manufacturing system
  - Two performance measures of that system
    - cycle time for parts (time from release into the factory until completion)
    - work in process (WIP) the total no of parts in the factory at any time.
  - Let Y_{ij} be the cycle time for the j^{th} part produced in the i^{th} replication.
  - Across-replication data are formed by summerizing within-replication data $\bar{Y}_i$
    - $\bar{Y}_i$ — sample mean of the n_i cycle times from i^{th} replication
    - S_i^2 — sample variance of the same data.

and $H_i = t_{\alpha/2, n_i-1} \frac{s_i}{\sqrt{n_i}}$ is a confidence interval half-width based on this data set.

within-in rep data	Across-Rep data
$Y_{11}\ Y_{12}\ ---\ Y_{1n_1}$	$\bar{Y}_1, -\ S_1^2, -\ H_1$
$Y_{21}\ Y_{22}\ ---\ Y_{2n_2}$	$\bar{Y}_2, -\ S_2^2, -\ H_2$
⋮	⋮
$Y_{R1}\ Y_{R2}\ ----\ Y_{Rn_R}$	$\bar{Y}_R, -\ S_R^2\ --\ H_R$
	$\bar{Y}\ --\ S^2,\ H$

within and across-rep data for cycle-time

- From the across-replication data we compute the overall statistics the avg of the daily cycle time averages.

$$\bar{Y} = \frac{1}{R}\sum_{i=1}^{R} \bar{Y_i}$$

the sample variance of the daily cycle time averages $S^2 = \frac{1}{R-1}\sum_{i=1}^{R}(\bar{Y_i} - \bar{Y})^2$

the confidence interval half width

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

-the quantity $S/\sqrt{R}$ is the standard error

within-replication

work in process is a continuous time o/p. denoted by $Y_i(t)$. the stopping time for ith replication T_{E_i} could be a random variable.

within-Rep data	Across-rep data
$Y_1(t),\ 0 \le t \le T_{E_1}$	$\bar{Y}_1,\ S_1^2,\ H_1$
$Y_2(t),\ 0 \le t \le T_{E_2}$	$\bar{Y}_2,\ S_2^2,\ H_2$
⋮	⋮
$Y_R(t),\ 0 \le t \le T_{ER}$	$\bar{Y}_R,\ S_R^2,\ H_R$
	$\bar{Y},\ S^2,\ H$

within-Rep and Across replication wip data

- the within-replication sample mean and variance are defined as

$$\bar{Y}_i = \frac{1}{T_{E_i}} \int_0^{T_{E_i}} Y_i(t)\, dt$$

$$S_i^2 = \frac{1}{T_{E_i}} \int_0^{T_{E_i}} \left(Y_i(t) - \bar{Y}_i\right)^2 dt$$

$$H_i = Z_{\alpha/2} \frac{S_i}{\sqrt{T_{E_i}}}$$

confidence intervals with 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

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

S^2 is the sample variance

R is the no of replications

- suppose that an error criterion ε is specified with probability $1-\alpha$ a sufficiently large sample size should satisfy

$$P(|\bar{Y} - \theta| < \varepsilon) \geq 1-\alpha$$

- Assume that an initial sample size of R_0 replications has been observed

8. Consider the following 60 values. Test whether $2^{nd}, 9^{th}, 16^{th}, \ldots$, numbers in the sequence are auto correlated for α=0.05.(Consider the tabular value $Z_{0.025}$ = 1.96).

0.30	0.48	0.36	0.01	0.54	0.34	0.96	0.06	0.61	0.85
0.48	0.86	0.14	0.86	0.89	0.37	0.49	0.60	0.04	0.83
0.42	0.83	0.37	0.21	0.90	0.89	0.91	0.79	0.57	0.99

0.95	0.27	0.41	0.81	0.96	0.31	0.09	0.06	0.23	0.77
0.73	0.47	0.13	0.55	0.11	0.75	0.36	0.25	0.23	0.72
0.60	0.84	0.76	0.30	0.26	0.38	0.05	0.19	0.73	0.44

Solution: Finding the autocorrelation value along with steps

- Finding the M value
  i.e. i+(M+1)m <= N

- Here i=2, m= 7 and N=60

i.e. 2+(M+1)7 <= 60.For M=8 this condition won't satisfy so the previous value is M=7.

- Finding the ρ_{im} value

i.e. ρ_{im} =
1/M+1[R2.R9+R9.R16+R16.R23+R23.R30+R30.R37+R37.R44+R44.R51+R51.R58] - 0.25

=1/8[(0.48)(0.61)+(0.61)(0.37)+(0.37)(0.37)+(0.37)(0.99)+(0.99)(0.09)+(0.09)(0.55)+(0.55)(0.60)+(0.60)(0.19)] – 0.25

=1/8(1.6043) – 0.25 = -0.494

- Finding the $\sigma_{\rho im}$ = 13M+7 / 12(M+1) value
  =13*7+7 / 12(8) = 0.1031

- Finding Z_0 value
  i.e. Z_0 = -0.0494/0.1031 = -0.4791

➢ Justification of acceptance

i.e. $-Z_{\alpha/2} <= Z_0 <= Z_{\alpha/2}$

= -1.96 < -0.4791 <1.96.Hence the sequence of numbers given are accepted