unit 1:Introduction to simulation

1. Simulation:

- **Simulation is the imitation of the operation of a real world process or system over time.**
- **Simulation models help us to study the behavior of system as it evolves**
- **models keeps the set of assumption concerning the operation of the system**
- **Assumptions are expressed in terms of mathematical, logical and symbolic relationship between the entities or object of interest of the system.**
- **Simulation modeling can be used both as an analysis tools to predict the performance of the new system and also predict the effect of changes to existing system.**
- **simulation can be done by hand or computer its keeps the history of system**
- **Simulation produce the set of data is used to estimate the measures of performance of system.**

1.1 When Simulation is the Appropriate Tool:

- **Study of and experimentation** with the internal interactions of a complex system, or of a subsystem within a complex system.
- Informational, organizational and environmental changes can be simulated and **the model's behavior can be observer.**
- The knowledge gained in designing a simulation model can be of great **value toward suggesting improvement in the system under investigation.**
- By changing simulation inputs and observing the resulting outputs, valuable insight may be obtained into which variables are most important and how variables interact.
- Simulation can be used as a **pedagogical (teaching) device** to reinforce analytic solution methodologies.
- Can be used to experiment **with new designs or policies prior to implementation,** so as to prepare for what may happen.
- Can be used to **verify analytic solutions.**
- By simulating different capabilities for a machine, requirements can be determined.
- Simulation models **designed for training, allow learning without the cost and disruption of on-the-job instructions.**
- **Animation** shows a system in simulated operation so that **the plan can be visualized**.
- **The modern system (factory, water fabrication plant, service organization, etc) is so complex** that the interactions can be treated only through simulation

1.2 When Simulation is Not Appropriate

õ Simulation should **not be used when the problem can be solved using common sense.**
õ Simulation should **not be used** if the problem can be **solved analytically.**
õ Simulation should **not be used** if it is easier to perform **direct experiments.**
õ Simulation should **not be used**, if the **costs exceeds** savings.
õ Simulation should **not be used** if the **resources or time are not available.**
õ No data is available, not even estimate simulation is not advised.

- õ If there is not enough time or the people are not available, simulation is not appropriate.
- õ If managers have unreasonable expectation say, too much soon – or the power of simulation is over estimated, simulation may not be appropriate.
- õ **If system behavior is too complex or cannot be defined,** simulation is not appropriate

1.3Advantages of Simulation

1. **New policies, operating procedures, decision rules, information flow, etc can be explored** without disrupting the ongoing operations of the real system.
2. **New hardware designs, physical layouts, transportation systems** can be tested without committing resources for their acquisition.
3. **Hypotheses** about how or why certain phenomena occur can be **tested for feasibility.**
4. **Time can be compressed or expanded allowing for a speedup or slowdown** of the phenomena under investigation.
5. Insight can be obtained about the **interaction of variables.**
6. Insight can be obtained about **the importance of variables to the performance of the system.**
7. **Bottleneck analysis** can be performed indication where work-in process, information materials and so on are being excessively delayed.
8. A **simulation study can help in understanding how the system operates** rather than how individuals think the system operates.
9. "what-if" questions can be answered. **Useful in the design of new systems.**

1.4Disadvantages of simulation

1. Model building **requires special training**. It is an art **that is learned over time and through experience.**
2. If two models are **constructed by two competent individuals,** they may have similarities, but it is highly unlikely that they will be the same.
3. Simulation results may be **difficult to interpret.** Since most simulation outputs are essentially random variables (they are usually based on random inputs), it may be hard to determine whether an observation is a result of system interrelationships or randomness.
4. Simulation modeling and analysis can be **time consuming and expensive**. Skimping on resources for modeling and analysis may result in a simulation model or analysis that is not sufficient for the task.
5. Simulation is used in some cases when an analytical solution is possible, or even preferable. This might be particularly true in the simulation of some waiting lines where closed-form queueing models are available.

1.5Applications of Simulation

- **Manufacturing application**
- Semiconductor manufacturing
- construction engineering
- military application
- Business process simulation
- Human system

1. **Manufacturing Applications**
   - Analysis of electronics assembly operations
   - Design and evaluation of a selective assembly station for high-precision scroll compressor shells
   - Comparison of dispatching rules for semiconductor manufacturing using large-facility models
   - Evaluation of cluster tool throughput for thin-film head production
   - Determining optimal lot size for a semiconductor back-end factory
   - Optimization of cycle time and utilization in semiconductor test manufacturing
   - Analysis of storage and retrieval strategies in a warehouse
   - Investigation of dynamics in a service-oriented supply chain
   - Model for an Army chemical munitions disposal facility
2. **Semiconductor Manufacturing**
   - Comparison of dispatching rules using large-facility models
   - The corrupting influence of variability
   - A new lot-release rule for wafer fabs
   - Assessment of potential gains in productivity due to proactive retile management
   - Comparison of a 200-mm and 300-mm X-ray lithography cell
   - Capacity planning with time constraints between operations
   - 300-mm logistic system risk reduction
3. **Construction Engineering**
   - Construction of a dam embankment
   - Trenchless renewal of underground urban infrastructures
   - Activity scheduling in a dynamic, multi project setting
   - Investigation of the structural steel erection process
   - Special-purpose template for utility tunnel construction
4. **Military Application**
   - Modeling leadership effects and recruit type in an Army recruiting station
   - Design and test of an intelligent controller for autonomous underwater vehicles
   - Modeling military requirements for non war fighting operations
   - Using adaptive agent in U.S Air Force pilot retention
5. **Logistics, Transportation, and Distribution Applications**
   - Evaluating the potential benefits of a rail-traffic planning algorithm
   - Evaluating strategies to improve railroad performance
   - Parametric modeling in rail-capacity planning
   - Analysis of passenger flows in an airport terminal
   - Proactive flight-schedule evaluation
   - Logistics issues in autonomous food production systems for extended-duration space exploration
   - Sizing industrial rail-car fleets
   - Product distribution in the newspaper industry
   - Design of a toll plaza

- Choosing between rental-car locations
- Quick-response replenishment

6. **Business Process Simulation**
   - Impact of connection bank redesign on airport gate assignment
   - Product development program planning
   - Reconciliation of business and systems modeling
   - Personnel forecasting and strategic workforce planning
7. **Human Systems and Healthcare**
   - Modeling human performance in complex systems
   - Studying the human element in air traffic control
   - Modeling front office and patient care in ambulatory health care practices
   - Evaluating hospital operations b/n the emergency department and a medical
   - Estimating maximum capacity in an emergency room and reducing length of stay in that room.

1.6 Systems and System Environment

System:

System is defined as a group of object that are joined together in some regular interaction or interdependence toward the accomplishment of same.

System environment:

A system is often affected by changes occurring outside the system,Such changes are said to occure in the system environment.

1.7 Components of a System

1) **Entity**: An entity is an object of interest in a system.
   Ex: In the factory system, departments, orders, parts and products are the entities.
2) **Attribute**: An attribute denotes the property of an entity.
   Ex: Quantities for each order, type of part, or number of machines in a department are attributes of factory system.
3) **Activity**: Represent a time period of specified length
   Ex: Manufacturing process of the department.
4) **State of the System**: The state of a system is defined as the collection of variables necessary to describe a system at any time, relative to the objective of study.
5) **Event**: An event is defined as an instantaneous occurrence that may change the state of the system.

   Endogenous : IS used to descried activites and events occurring with in the system

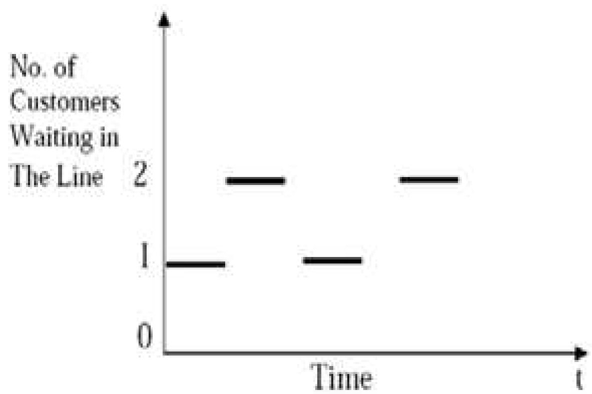
   Exogenous: Is used to descried activites and events in the environment that affect the system.

Examples of system and components					
System	**Entities**	**Attributes**	**Activities**	**Events**	**State variables**
Banking	Customers	Checking-account balance	Making deposits	Arrival; departure	No. of busy tellers; no. of customers waiting.
Rapid rail	Riders	Origination; destination	Traveling	Arrival at station; arrival at destination	No. of riders waiting at each station; No. of riders in transit
Production	Machines	Speed; capacity; breakdown rate length	Welding; stamping	Breakdown	Status of machines (busy, idle or down)
Inventory	Warehouse	Capacity	Withdrawing	Demand	Levels of inventory; backlogged demands

1.8 Discrete and Continuous Systems

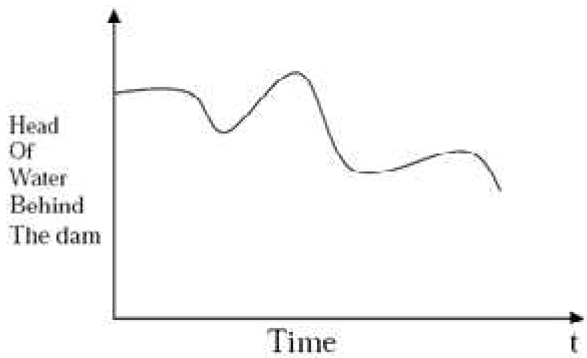
Discrete System:

- **Is one in which the state variable change only at a discrete set of points in time.**
- **The bank is an example, since the state variable the number of customer in the bank changes only when a customer arrives or when the service provided a customer is completed.**



Continuous system:

- **Is one in which the state variable change continuous over time.**
- **head of water behind a dam, during and for some time after a rain storm water flow into the lake behind the dam.**



1.9 Model of a system

- A **model** is defined as a representation of a system for the purpose of studying the system.
- It is necessary to consider only those aspects of the system that affect the problem under investigation.
- These aspects are represented in a model, and by definition it is a simplification of the system.

Types of Models:

- **Mathematical or physical model**
- **Static and dynamic model**
- **deterministic and stochastic model**
- **discrete and continuous model**

1.Mathematical or physical model:

Mathematical model uses symbolic notation and equations to represents a system

2.Static model:

A static simulation models represent a system at a particular point in time it is also called as monte carlo simulation.

3.dynamic model:

A dynamic simulation models represent system as the change over time. simulation of a bank from 9 to 4 is an example

4.Deterministic model:

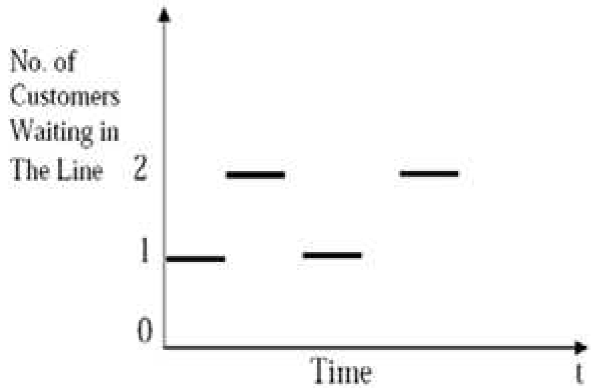
A simulation variable that contain no random variable, have a set of known input which will result in a unique set of output.

5.Stochastic model:

A stochastic simulation model has one or more random **variable** as input. Random input lead to random output.Since the output are random they can be consider only as estimates of the true characteristics of a model.

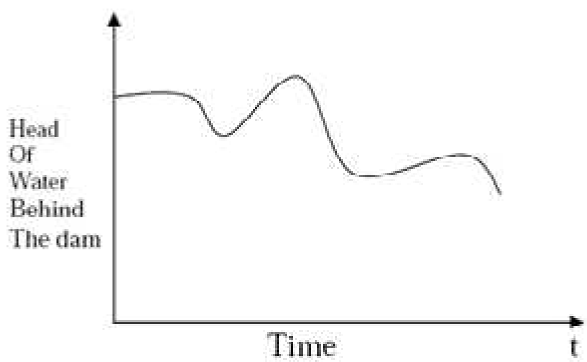
6.Discrete System:

- Is one in which the state variable change only at a discrete set of points in time.
- The bank is an example, since the state variable the number of customer in the bank changes only when a customer arrives or when the service provided a customer is completed.



7.Continuous system:

- Is one in which the state variable change continuous over time.
- head of water behind a dam, during and for some time after a rain storm water flow into the lake behind the dam.



1.10 Discrete event system simulation:

- The model of system in which state variable changes only at a discrete set of points in times
- The simulation models are analyzed by numerical rather than by analytical methods.
- Analytical methods employ the deductive reasoning of mathematics to solve the model. E.g.: Differential calculus can be used to determine the minimum cost policy for some inventory models.
- Numerical methods use computational procedures and are 'runs', which is generated based on the model assumptions and observations are collected to be analyzed and to estimate the true system performance measures.
- Real-world simulation is so vast, whose runs are conducted with the help of computer. Much insight can be obtained by simulation manually which is applicable for small

systems.

1.11Steps in a simulation study:

1. Problem formulation
2. Setting of objectives and overall project plan
3. model conceptualization
4. data Collection
5. model translation
6. verified
7. validated
8. Experimental design
9. production runs and analysis
10. more runs
11. documentation and reporting
12. Implementation

1. Problem formulation:

- Every study should begin with a statement of the problem.
- If the statement is provided by the policy makers or those that have the problem, The analyst must ensure that the problem being described is clearly understood
- If the problem statement is being developed by the analyst, it is important that the policy makers understand and agree with the formulation.

2. Setting of objective and overall project plan:

- The objectives indicate the questions to be answered by simulation.
- At this point a determination should be made concerning whether simulation is the appropriate methodology. Assuming that it is appropriate,
- the overall project plan should include the study in terms of
    - A statement of the alternative systems
    - A method for evaluating the effectiveness of these alternatives
    - Plans for the study in terms of the number of people involved
    - Cost of the study
    - The number of days required to accomplish each phase of the work with the anticipated results.

3. Model Conceptualization:

- The construction of a model of a system is probably as much art as science.
- The art of modeling is enhanced by ability to have following:
    - To abstract the essential features of a problem.
    - To select and modify basic assumptions that characterizes the system.
    - To enrich and elaborate the model until a useful approximation results.

4. Data Collection:

- There is a constant interplay between the construction of the model and the collection of the needed input data.
- As complexity of the model changes the required data elements may also change.
- Since data collection takes such a large portion of the total time required to perform a simulation it is necessary to begin it as early as possible.

5. Model Translation:

- Since most real world system result in model that require a great deal of information storage and computation, the model must be entered into a computer recognizable format.
- we use term program even though it is possible to accomplish the desired result in many instances with little or no actual coding.

6.Varified:

- It pertains to the computer program and checking the performance.
- If the input parameters and logical structure and correctly represented, verification is completed.

7.Validated:

- validation is the determination that a model is an accurate representation of the real system.
- Is usually achieved through the calibration of the model an iterative process of comparing the model to actual system behavior and using the discrepancy between the two and the insights gained to improve the model.
- This process is repeated until model accuracy is judges acceptable.

8.**Experimental Design:**

- The alternatives that are to be simulated must be determined. For each system design, decisions need to be made concerning

  a. Length of the initialization period
  b. Length of simulation runs
  c. Number of replication to be made of each run

9.Production runs and analysis:

- They are used to estimate measures of performance for the system designs that are being simulated.

10.More runs:

- Based on the analysis of runs that have been completed. The analyst determines if additional runs are needed and what design those additional experiments should follow.

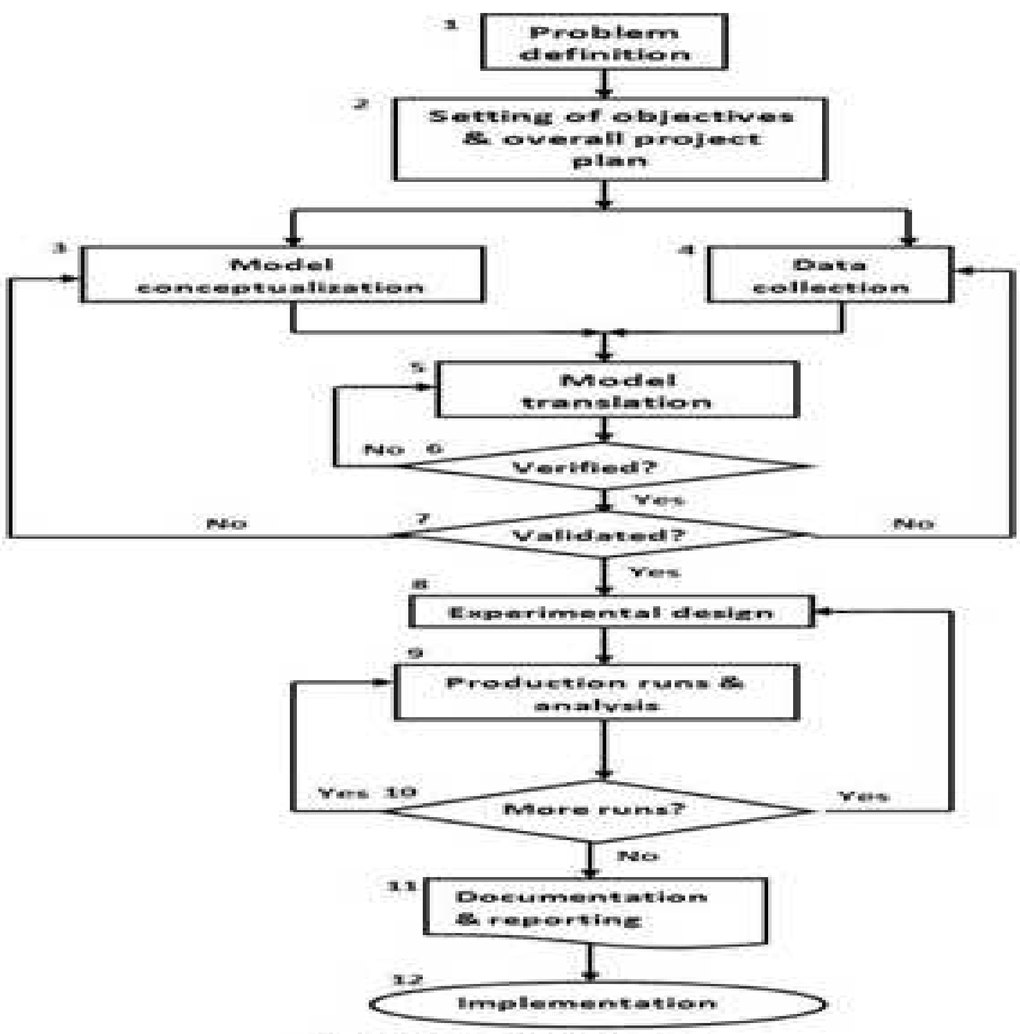
11.Documentation and reporting:

Two types of documentation. Program documentation and Process documentation

- **Program documentation:** Can be used again by the same or different analysts to understand how the program operates
- **Process documentation:** This enable to review the final formulation and alternatives, results of the experiments and the recommended solution to the problem. The final report provides a vehicle of certification.

12.Implementation:

Success depends on the previous steps. If the model user has been thoroughly involved and understands the nature of the model and its outputs, likelihood of a vigorous implementation is enhanced.



Steps in a simulation study

1.12 Simulation of queuing systems

A **Queuing system** is described by **its calling population, the nature of its arrivals, the service mechanism, the system capacity, and queuing discipline.**

Simulation is often used in the analysis of queuing models. In a simple typical queuing model, shown in

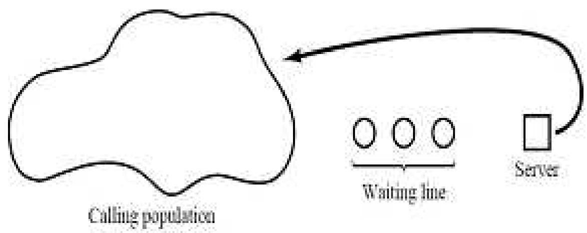


Figure 2 Queueing system.

- In the single-channel queue, the calling population is **infinite**; that is, if a **unit leaves the calling population and joins the waiting line or enters service, there is no change in the arrival rate of other units that may need service.**
- **Arrivals for service occur one at a time in a random fashion**; once they **join the waiting line,** they are eventually served.
- The **system capacity has no limit**, meaning that any number of units can wait in line. Finally, **units are served in the order of their arrival (often called FIFO: first in, first out) by a single server or channel**.
- **Arrivals and services** are defined by the **distributions of the time between arrivals** and the **distribution of service times,** respectively.
- The **state of the system:** the number of units in the system and the status of the server, busy or idle.
- **An event**: a set of circumstances that cause an instantaneous change in the state of the system. In a single-channel queueing system there are only two possible events that can affect the state of the system.
- The **simulation clock** is used to track simulated time.

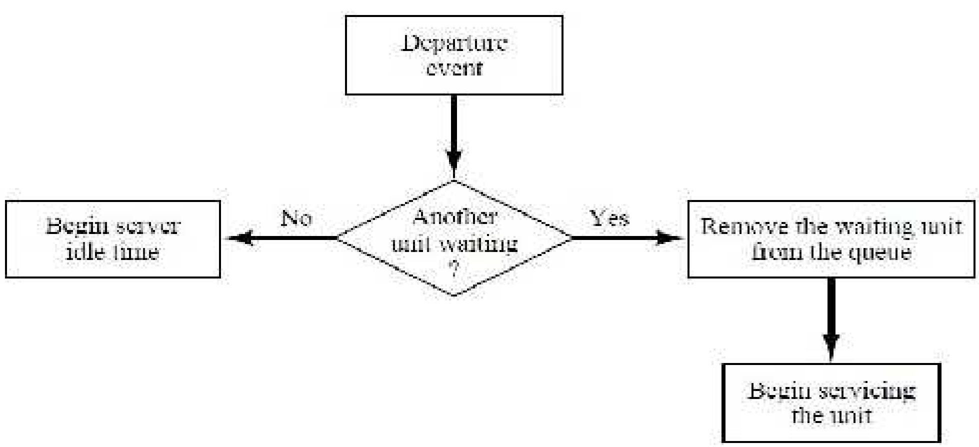


Figure 2.2 Service-just-completed flow diagram.

- The arrival event occurs when a unit enters the system. **The flow diagram for the arrival event is shown in**

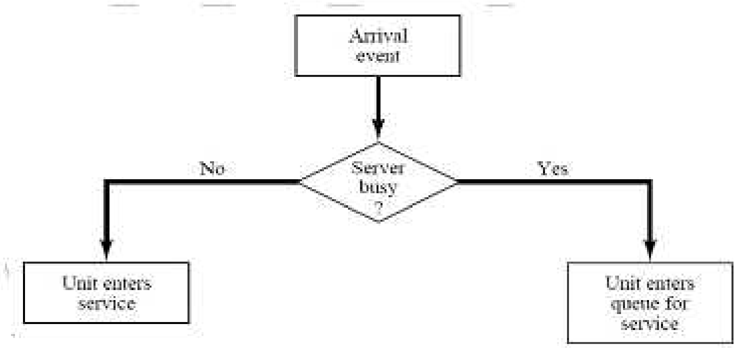


Figure Unit-entering-system flow diagram.

- The unit may find the **server either idle or busy;** therefore, either the unit begins service immediately, or it enters the queue for the server. The unit follows the course of action shown in fig 2.4.
- If the server is busy, the unit enters the queue. If the server is idle and the queue is empty, the unit begins service. It is not possible for the server to be idle and the queue to be nonempty.

		Queue status	
		Not empty	Empty
Server status	Busy	Enter queue	Enter queue
	Idle	Impossible	Enter service

Figure 2.4 Potential unit actions upon arrival.

- After the completion of a service the service may become idle or remain busy with the next unit. The relationship of these two outcomes to the status of the queue is shown in fig 2.5. If the queue is not empty, another unit will enter the server and it will be busy

		Queue status	
		Not empty	Empty
Server outcomes	Busy		Impossible
	Idle	Impossible	

Figure 2.5 Server outcomes after service completion.

Problems:

Single channel queuing system problem formulas:

1. Time Customer wait in queue= Time service begin – Arrival Time
2. Time Service End= Service time + Time service begin
3. Time customer Spend In system= Time service end-Arrival Time
4. Idel Time of Server=Time service Begin(N)-Time Service end(N-1)

Standard Formulas:

1.**Average waiting time(i.e customer wait)**=total time customer wait in queue / Total number of customer

2.**Probability(Wait i.e customer wait)=**Number of Customer who wait / Total number of customer

3.**Probability of idle server(idle time of server)=**total idle time of server / total run time of simulation

4.**average service time**=total service time/total number of customer

5.**average times between arrivals**=sum of all times between arrival/number of arrivals-1

6.**Average waiting time those who wait in queue**=total time customer wait in queue/total number of customer who wait

7.**Average time customer spend In the system**=Total time customer spend in system/total number of customer

① Single channel and Multi channel (Able Baker prob.) Queuing system problem :

UNIT-1

$$\text{Average Waiting time (of customer)} = \frac{\text{Total time customer waiting in queue}}{\text{Total no. of customers}}$$

$$\text{Probability wait (of customer)} = \frac{\text{No. of customers to wait}}{\text{Total no. of customers}}$$

$$\text{Probability of idle server (Idle time of server)} = \frac{\text{Total idle time of server}}{\text{Total run time of simulation (TSE last value)}}$$

$$\text{Average Service Time} = \frac{\text{Total service time}}{\text{Total no. of customers}}$$

$$\text{Average time b/w arrival} = \frac{\text{Sum of all time b/w arrival}}{\text{No. of arrival} - 1}$$

$$\text{Average waiting time those who wait in queue} = \frac{\text{Total time customer wait in queu}}{\text{Total no. of customer who wait}}$$

$$\text{Average time customer spends in the system} = \frac{\text{Total time customer spends in S/m}}{\text{Total no. of customers}}$$

AT → Arrival Time

IAT → Inter Arrival Time

ST → Service Time

C.No → Customer Number

RD → Random Digit

RDA → Random Digit Assessment

TSB → Time Service Begin

TSE → Time Service End

CP → Cumulative Probability

P → Probability

1. A small grocery store has only one check out counter. The customer arrives at this check out counter at random from 1 to 8 min apart. Each possible value of service time has same probability of occurance. The service time varies from 1 to 6 mins apart. Each possible value of service time has same probability of occurance. Develop simulation distribution table for 8 customers.

Random digit for arrival time :

913 727 015 948 309 922 753 235 302

Service Time : (Random digit)

84 10 74 53 17 79 91 67 89 38

i) Determine Inter Arrival Time distribution table

C.No	P(1/8)	Cumulative Probability	Random Digit Assessment
1	0.125	0.125	001 – 125
2	0.125	0.250	126 – 250
3	0.125	0.375	251 – 375
4	0.125	0.500	376 – 500
5	0.125	0.625	501 – 625
6	0.125	0.750	626 – 750
7	0.125	0.875	751 – 875
8	0.125	1.000	876 – 000

(2). ii] Compute Arrival time from IAT

C.No	Random Digit	IAT	Arrival Time
1	–	–	0
2	913	8	8
3	727	6	14
4	015	1	15
5	948	8	23
6	309	3	26
7	922	8	34
8	753	7	41
9	235	2	43
10	302	3	46

iii] Determine Service Time Distribution Table

S.No	P(1/6)	Cumulative probability	Random Digit Assess
1	0.16	0.16	01–16
2	0.16	0.32	17–32
3	0.16	0.48	33–48
4	0.16	0.64	49–64
5	0.16	0.80	65–80
6	0.16	0.96	81–96

iv] Compute Service Time

S.No	Random digit	Service Time
1	84	6
2	10	1
3	74	5
4	53	4
5	17	2
6	79	5
7	91	6
8	67	5
9	89	6
10	38	3

v) Simulation Table for 10 customers

					TSE − AT	TSB(n) − TSE(n−1)	TSB − AT
C.No	Arrival Time	Service Time	Time Service Begin	Time Service End	Customer Time spent in System	Idle Server Time	Customer Time wait in queue
1	0	6	0	6	6 (6−0)	0	0
2	8	1	8	9	1 (9−8)	2 (8−6)	0 (8−8)
3	14	5	14	19	5 (19−14)	5 (14−9)	0 (14−14)
4	15	4	19	23	8	0	4 (19−15)
5	23	2	23	25	2	0	0
6	26	5	26	31	5	1	0
7	34	6	34	40	6	3	0
8	41	5	41	46	5	1	0
9	43	6	46	52	9	0	3
10	46	3	52	55	9	0	6

$$\text{Average waiting Time} = \frac{\text{Total time customer waiting in queue}}{\text{Total no. of customers}} = \frac{4+3+6}{10} = \frac{13}{10} = 1.3 \text{ min}$$

$$\text{Probability wait} = \frac{\text{No. of customers to wait}}{\text{Total no. of customers}} = \frac{3}{10} = 0.3 \text{ min}$$

$$\text{Probability of idle server} = \frac{\text{Total idle time of server}}{\text{Total run time of simulat}^n} = \frac{2+5+1+3+1}{55} = \frac{12}{55} = 0.218 \text{ min}$$

(TSE last value)

(3) Average Service Time = $\frac{\text{Total Service Time}}{\text{Total no. of customers}}$

$$= \frac{6+1+5+4+2+5+6+5+6+3}{10} = \frac{43}{10} = 4.3\text{m}$$

Average Time b/w arrival = $\frac{\text{Sum of all time b/w arrival}}{\text{No. of arrival} - 1}$

$$= \frac{46}{10-1} = \frac{46}{9} = 5.11 \text{ min}$$

Average waiting time those who wait in queue = $\frac{\text{Total time customer wait in queue}}{\text{Total no. of customer who wait}}$

$$= \frac{4+3+6}{3} = \frac{13}{3} = 4.3 \text{ min}$$

Average time customer spent in system = $\frac{\text{Total time customer spends in s/r}}{\text{Total no. of customers}}$

$$= \frac{6+1+5+8+2+5+6+5+9+9}{10}$$

$$= \frac{56}{10} = 5.6 \text{ min}$$

---

2. A small grocery store has only one check out counte at random from 1-6 min apart. Each possible value of IAT has the same probability of occurance. The service time vary from 1 to 6 mins with probability shown below.

ST :	1	2	3	4	5	6
P :	0.10	0.20	0.25	0.30	0.10	0.05

Develop a simulation Table for 10 customers.

Take a random digit for arrival :

13 27 15 48 9 22 53 35 2

Random digit for Service Time

84 10 74 53 17 91 79 67 38 89 sequentially.

Calc. Avg Service Time, probability of idle Service time, average b/w arrivals & avg. time customer spent in System.

i] Determine IAT Distribution Table

C.No	Probability (1/6)	Cumulative probability	Random digit assessment
1	0.16	0.16	01-16
2	0.16	0.32	17-32
3	0.16	0.48	33-48
4	0.16	0.64	49-64
5	0.16	0.80	65-80
6	0.16	0.96	81-96

ii] Compute Arrival time from IAT

C.No	RD	IAT	AT
1	–	–	0
2	13	1	1
3	27	2	3
4	15	1	4
5	48	3	7
6	9	1	8
7	22	2	10
8	53	4	14
9	35	3	17
10	2	1	18

iii] Determine ST Dist. Table

S.No	Probability	CP	RDA
1	0.10	0.10	01-10
2	0.20	0.30	11-30
3	0.25	0.55	31-55
4	0.30	0.85	56-85
5	[illegible]	0.95	86-95

iv) Compute Service Time from Distribution Table

S.No	RD	ST
1	84	4
2	10	1
3	74	4
4	53	3
5	17	2
6	91	5
7	79	4
8	67	4
9	38	3
10	89	5

imulation Table for 10 customers

o	Arrival Time	Service Time	Time Service Begin	Time Service End	Cust. time spent in s/m	Idle Server Time	Cust. time wait in queue
	0	4	0	4	4	0	0
	1	1	4	5	4	0	3
	3	4	5	9	6	0	2
	4	3	9	12	8	0	5
	7	2	12	14	7	0	5
	8	5	14	19	11	0	6
	10	4	19	23	13	0	9
	14	4	23	27	13	0	9
	17	3	27	30	13	0	10
	18	5	30	35	17	0	12

erage Service Time = $\frac{35}{10} = \underline{\underline{3.5 \text{ min}}}$

bability of Idle Server Time = $\frac{0}{35} = \underline{\underline{0}}$

erage Time b/w arrivals = $\frac{18}{9} = \underline{\underline{2 \text{ min}}}$

erage time Customer spent in system = $\frac{61}{10} = \underline{\underline{6.1 \text{ min}}}$

3. Consider a store with one checkout counter. Prepare simulation table & find out average waiting time of customer in waiting queue, probability of idle server, average service time

* IAT : 3 2 6 4 4 5 8 7

ST : 4 5 5 8 4 6 2 3 4

Assume 1^{st} customer arrives at $t=0$.

i] Inter arrival distribution Table

C.No	IAT	AT
1	–	0
2	3	3
3	2	5
4	6	11
5	4	15
6	4	19
7	5	24
8	8	32
9	7	39

ii] Simulation table for 9 customers

C.No	AT	ST	TSB	TSE	Cust. spent in s/m.	Idle time of server	Cust. time wait in que
1	0	4	0	4	4	0	0
2	3	5	4	9	6	0	1
3	5	5	9	14	9	0	4
4	11	8	14	22	11	0	3
5	15	4	22	26	11	0	7
6	19	6	26	32	13	0	7
7	24	2	32	34	10	0	8
8	32	3	34	37	5	0	2
9	39	4	39	43	4	2	0

⑤ Average waiting time of customer in waiting queue $= \frac{32}{7} = 4.57$ min

Probability of Idle Server $= \frac{2}{10} = 0.2$ min

Average Service Time $= \frac{41}{9} = 4.56$ min

Multichannel Problems :

Type I

4 Consider a simulation with a restaurant system where car ~~$~~ hope takes order & brings an item to the car. The car arrives in the manner :

Time b/w arrival :	1	2	3	4
Probability :	0.25	0.40	0.20	0.15

Consider 2 persons Able & Baker. Able is better & bit faster than Baker.

Able Service Time :

ST :	2	3	4	5
Prob :	0.30	0.28	0.25	0.17

Baker Service Time :

ST :	3	4	5	6
Prob :	0.35	0.25	0.20	0.20

Take a random digit for arrival :

26 98 90 26 42 74 80 68 22

Random digit for service time :

95 21 51 92 89 38 13 61 50 49

i) Determine Inter Arrival distribution Table

C·No	Probability	Cumulative Probability	Random Digit Assessment
1	0·25	0·25	01 – 25
2	0·40	0·65	26 – 65
3	0·20	0·85	66 – 85
4	0·15	1·00	86 – 00

ii) Compute AT from IAT distribution Table

C·No	RD	IAT	AT
1	–	–	0
2	26	2	2
3	98	4	6
4	90	4	10
5	26	2	12
6	42	2	14
7	74	3	17
8	80	3	20
9	68	3	23
10	22	1	24

iii) Able Service Time Distribution Table

S·No	Probability	CP	RDA
2	0·30	0·30	01 – 30
3	0·28	0·58	31–58
4	0·25	0·83	59 – 83
5	0·17	1·00	84 – 00

iv) Baker Service Time Distribution Table

S·No	Probability	CP	RDA
3	0·35	0·35	01 – 35
4	0·25	0·60	36 – 60
5	0·20	0·80	61 – 80
6	0·20	1·00	81 – 00

v 6 Simulation table for 10 customers

C.No	AT	RD for service	ST	When Able is Available	When Baker is Available	Server Choosen	Able TSB	Able TSE	Baker TSB	Baker TSE	Cust in s/m	Idle Server time	wa in que
1	0	95	5	0	0	A	0	5	–	–	5	0	0
2	2	21	3	5	0	B	–	–	2	5	3	2	0
3	6	51	3	5	5	A	6	9	–	–	3	1	0
4	10	92	5	9	5	A	10	15	–	–	5	1	0
5	12	89	6	15	5	B	–	–	12	18	6	7	0
6	14	38	3	15	18	A	15	18	–	–	4	0	1
7	17	13	2	18	18	A	18	20	–	–	3	0	1
8	20	61	4	20	18	A	20	24	–	–	4	0	0
9	23	50	4	24	18	B	–	–	23	27	4	5	0
10	24	49	3	24	27	B	24	27	–	–	3	0	0

5. Consider Simulation table for Able & Baker problem where time b/w arrival are :

Time b/w arrival :	1	2	3	4	5
Probability :	0.20	0.15	0.05	0.20	0.40

Baker is faster than Able.

Service Time for Able :

ST :	3	4	5	6	2
Prob :	0.20	0.05	0.15	0.20	0.40

Service Time for Baker :

ST :	2	3	5	6	1
Prob :	0.15	0.20	0.05	0.20	0.40

Random Digit for Arrival :

98 90 42 80 22 26 74 26 68

Random Digit for Service Time :

49 50 61 13 38 89 92 51 21 95

i) Determine IAT distribution Table

C.No	Probability	CP	RDA
1	0.20	0.20	01-20
2	0.15	0.35	21-35
3	0.05	0.40	36-40
4	0.20	0.60	41-60
5	0.40	1.00	61-00

ii) Compute Arrival Time from IAT dist. table.

C.No	RD	IAT	AT
1	-	-	0
2	98	5	5
3	90	5	10
4	42	4	14
5	80	5	19
6	22	2	21
7	26	2	23
8	74	5	28
9	26	2	30
10	68	5	35

iii) Able Service Time Distribution Table

S.No	P	CP	RDA
3	0.20	0.20	01-20
4	0.05	0.25	21-25
5	0.15	0.40	26-40
6	0.20	0.60	41-60
2	0.40	1.00	61-00

iv) Baker Service Time Distribution Table

S.No	P	CP	RDA
2	0.15	0.15	01-15
3	0.20	0.35	16-35
5	0.05	0.40	36-40
6	0.20	0.60	41-60
1	0.40	1.00	61-00

(RD - Random Digit)

v] Simulation table for 10 customers

C.No	Arrival Time	RD of service	Service Time	When Able is Available	When Baker is Available	Server Choosen	Able		Baker		Cust. time spent in s/m	Idle time of Server	Cust. t wait queu
							TSB	TSE	TSB	TSE			
1	0	49	6	0	0	B	—	—	0	6	6	0	0
2	5	50	6	0	6	A	5	11	—	—	6	5	0
3	10	61	1	11	6	B	—	—	10	11	1	4	0
4	14	13	2	11	11	B	—	—	14	16	2	3	0
5	19	38	5	11	16	B	—	—	19	24	5	3	0
6	21	89	2	11	24	A	21	23	—	—	2	10	0
7	23	92	2	23	24	A	23	25	—	—	2	0	0
8	28	51	6	25	24	B	—	—	28	34	6	4	0
9	30	21	4	25	34	B	30	34	—	—	4	5	0
10	35	95	1	34	34	A	—	—	35	36	1	1	

Customer time Spent in s/m = TSE - AT

Idle time of Server = TSB(n) - TSE(n-1)

Customer time waiting in queue = TSB - AT

6. Consider Able Baker Dist. table, IAT having equal probability ratio 1 to 6 min apart.

* Able Service Time have equal probability ratio 1 - 5 min

Baker Service Time have equal probability ratio 1 - 4 min.
1 - 6

Random Digit for time b/w arrival :

62 89 9 62 24 47 8 86 22

Random Digit for Service Time :

59 12 15 29 28 83 13 16 5 94

Able is faster than Baker.

i) Determine IAT Distribution Table

C.No	P	CP	RDA
1	0.16	0.16	01 - 16
2	0.16	0.32	17 - 32
3	0.16	0.48	33 - 48
4	0.16	0.64	49 - 64
5	0.16	0.80	65 - 80
6	0.16	0.96	81 - 96

ii) Compute Arrival Time from IAT Dist. Table

C.No	RD	IAT	AT
1	-	-	0
2	62	4	4
3	89	6	10
4	9	1	11
5	62	4	15
6	24	2	17
7	47	3	20
8	8	1	21
9	86	6	27
10	22	2	29

Distribution Table

S·No	P	CP	RDA
1	0·20	0·20	01–20
2	0·20	0·40	21–40
3	0·20	0·60	41–60
4	0·20	0·80	61–80
5	0·20	1·00	81–00

Distribution Table

S·No	P	CP	RDA
1	0·25	0·25	01–25
2	0·25	0·50	26–50
3	0·25	0·75	51–75
4	0·25	1·00	76–00

v) Simulation Table for 10 customers

C·No	AT	RD of service	ST	When Able is Available	When Baker is Available	Server Choosen	Able		Baker		Cust. spent in s/m	Idle time of server	Cust. time wait in queue
							TSB	TSE	TSB	TSE			
1	0	59	3	0	0	A	0	3	–	–	3	0	0
2	4	12	1	3	0	A	4	5	–	–	1	1	0
3	10	15	1	5	0	A	10	11	–	–	1	5	0
4	11	29	2	11	0	A	11	13	–	–	2	0	0
5	15	98	15	13	0	A	15	20	–	–	2	2	0
6	17	83	4	20	0	B	–	–	17	21	2	17	0
7	20	13	1	20	21	A	20	21	–	–	1	0	0
8	21	16	1	21	21	A	21	22	–	–	1	0	0
9	27	5	1	22	21	A	27	28	–	–	2	5	0
10	29	94	5	28	21	A	29	34	–	–	5	1	0

7. Develop a simulation table for Able & Baker problem.

Inter Arrival Times : 3 2 1 5 4 6

Service Time (Able) : 3 5 2 6 2 1 7

Service Time (Baker) : 2 1 4 5 7 6 2

Able is much faster than Baker

i) Compute AT from IAT

C. No	IAT	AT
1	–	0
2	3	3
3	2	5
4	1	6
5	5	11
6	4	15
7	6	21

ii) Simulation Table for 7 customers

C.No	AT	ST	When Able is Available	When Baker is Available	Server choosen	Able TSB	Able TSE	Baker TSB	Baker TSE	Cust. spent in s/m	Idle time of Server	Cust. time wait in queue
1	0	3	0	0	A	0	3	–	–	0	0	3
2	3	5	3	0	A	3	8	–	–	0	0	5
3	5	4	8	0	B	–	–	5	9	0	5	4
4	6	6	8	9	A	8	14	–	–	2	0	8
5	11	7	14	9	B	–	–	11	18	0	2	7
6	15	1	14	18	A	15	16	–	–	0	1	1
7	21	7	16	18	A	21	28	–	–	0	5	7

(9) having equal probability ratio 1 to 7 min apart.

Service Time for Able :

ST :	1	2	3	4	5
Prob :	0.20	0.10	0.30	0.20	0.20

Service Time for Baker :

ST :	1	2	3	4	5
Prob. :	0.10	0.20	0.20	0.30	0.20

Random Visit for Arrival : 95 60 35 40 52 54 10

Random visit for Service : 60 95 35 40 24 54 10 25

Baker is faster than Able.

i) Determine Inter Arrival Distribution Table.

C.No	P	CP	RDA
1	0.14	0.14	01-14
2	0.14	0.28	15-28
3	0.14	0.42	29-42
4	0.14	0.56	43-56
5	0.14	0.70	57-70
6	0.14	0.84	71-84
7	0.14	0.98	85-98

ii) Compute AT from IADT.

C.No	RD	IAT	AT
1	-	-	0
2	95	7	7
3	60	5	12
4	35	3	15
5	40	3	18
6	52	4	22
7	54	4	26
8	10	1	27

iii) Determine Able Service Time Distribution Table

S.No	P	CP	RDA
1	0.20	0.20	01-20
2	0.10	0.30	21-30
3	0.30	0.60	31-60
4	0.20	0.80	61-80
5	0.20	1.00	81-00

iv) Determine Baker Service Time Distribution table

S.No	P	CP	RDA
1	0.10	0.10	1-10
2	0.20	0.30	11-30
3	0.20	0.50	31-50
4	0.30	0.80	51-80
5	0.20	1.00	81-00

v] Simulation Table for 8 customers.

C.No	Arrival Time	RD of Service	Service Time	When Able is Available	When Baker is Available	Server Choosen	Able		Baker		Cust. spent in s/m	Idle Server Time	Cust. W in que
							TSB	TSE	TSB	TSE			
1	0	60	4	0	0	B	–	–	0	4	4	0	0
2	7	95	5	0	4	B	–	–	7	12	5	3	0
3	12	35	3	0	12	B	–	–	12	15	3	0	0
4	15	40	3	0	15	B	–	–	15	18	3	0	0
5	18	24	2	0	18	B	–	–	18	20	2	0	0
6	22	54	4	0	20	B	–	–	22	26	4	2	0
7	26	10	1	0	26	B	–	–	26	27	1	0	0
8	27	25	2	0	27	B	–	–	27	29	2	0	0

General Principles

1. Discrete-event simulation

- The basic building blocks of all discrete-event simulation models: entities and attributes, activities and events.
- A system is modeled in terms of
  - Its state at each point in time
  - The entities that pass through the system and the entities that represent system resources
  - The activities and events that cause system state to change.
- Discrete-event models are appropriate for those systems for which changes in system state occur only at discrete points in time.
- This chapter deals exclusively with dynamic, stochastic systems (i.e., involving time and containing random elements) which change in a discrete manner.

Concepts in Discrete-Event Simulation(components of discrete event Simulation)

1. **System:** A collection of entities (e.g., people and machines) that together over time to accomplish one or more goals.
2. **Model:** An abstract representation of a system, usually containing structural, logical, or mathematical relationships which describe a system in terms of state, entities and their attributes, sets, processes, events, activities, and delays.
3. **System state:** A collection of variables that contain all the information necessary to describe the system at any time.
4. **Entity:** Any object or component in the system which requires explicit representation in the model (e.g., a server, a customer, a machine).
5. **Attributes:** The properties of a given entity (e.g., the priority of a v customer, the routing of a job through a job shop).
6. **List:** A collection of (permanently or temporarily) associated entities ordered in some logical fashion (such as all customers currently in a waiting line, ordered by first come, first served, or by priority).
7. **Event:** An instantaneous occurrence that changes the state of a system as an arrival of a new customer).
8. **Event notice:** A record of an event to occur at the current or some future time, along with any associated data necessary to execute the event; at a minimum, the record includes the event type and the event time.
9. **Event list:** A list of event notices for future events, ordered by time of occurrence; also known as

the future event list (FEL).

10. **Activity:** A duration of time of specified length (e.g., a service time or arrival time), which is known when it begins (although it may be defined in terms of a statistical distribution).
11. **Delay:** A duration of time of unspecified indefinite length, which is not known until it ends (e.g., a customer's delay in a last-in, first-out waiting line which, when it begins, depends on future arrivals).
12. **Clock:** A variable representing simulated time.

The Event-Scheduling/Time-Advance Algorithm

- **The mechanism for advancing simulation time and guaranteeing that all events occur in correct chronological order is based on the future event list (FEL).**
- **Future Event List (FEL)**
  - **To contain all event notices for events that have been scheduled to occur at a future time.**
  - **To be ordered by event time,** meaning that the events are arranged chronologically; that is, the event times satisfy.
  - **Scheduling a future event** means that at the instant an activity begins, its duration is computed or drawn as a sample from a statistical distribution and the end-activity event, together with its event time, is placed on the future event list.

The sequence of actions which a simulator must perform to advance the clock system snapshot is called the event- scheduling/time-advance algorithm.

The system snapshot at time t=0 and t=t1 (VIP VTU question)

ClK	System State		Future Event List	
T	(5,1,6)		(3, t1)— Type 3 event to occur at timet1 (1, t2)— Type 1 event to occur at time t2 (1, t3)- Type 1 event to occur at time t3 (2, tn)— Type 2 event to occur at time tn	

Event-scheduling/time-advance algorithm

Step 1. Remove the event notice for the imminent event

(event 3, time t\) from FEL

Step 2. Advance CLOCK to imminent event time

(i.e., advance CLOCK from r to t1).

Step 3. Execute imminent event: update system state, change entity attributes, and set membership as needed.

Step 4. Generate future events (if necessary) and place their event notices on PEL ranked by event time.

(Example: Event 4 to occur at time t*, where t2 < t* < t3.)

Step 5. Update cumulative statistics and counters.

New system snapshot at time *t1*

OCK	System		Future Event List	
T1	(5,1,5)		(1, t2)— Type 1 event to occur at time t1 (4, t*)— Type 4 event to occur at time t* (1, t3)— Type 1 event to occur at time t3 (2, tn)—Type 2 event to occur at time tn	

2.Manual Simulation Using Event Scheduling

In an event-scheduling simulation, a simulation table is used to record the successive system snapshots as time advances.

Let us consider the example of a grocery shop which has only one checkout counter. **(Single-Channel Queue)**

The system consists of those customers in the waiting line plus the one (if any) checking out. The model has the following components:

System state (LQ (t), LS (t)), where LQ (t) is the number of customers in the waiting line, and LS (t) is the number being served (0 or 1) at time t.

Entities: The server and customers are not explicitly modeled, except in terms of the state variables above.

Events

Arrival(A)

Departure(D)

Stopping event (E), scheduled to occur at time 60.

Event notices

(A, t). Representing an arrival event to occur at future time t

(D, t), representing a customer departure at future time t

(E, 60), representing the simulation-stop event at future time 60

Activities

Interarrival time, Service time,

Delay Customer time spent in waiting line.

In this model, the FEL will always contain either two or three event notices.

Flow Chart for execution of arrival and departure event using time advance /Event scheduling algorithm (vtu Question)

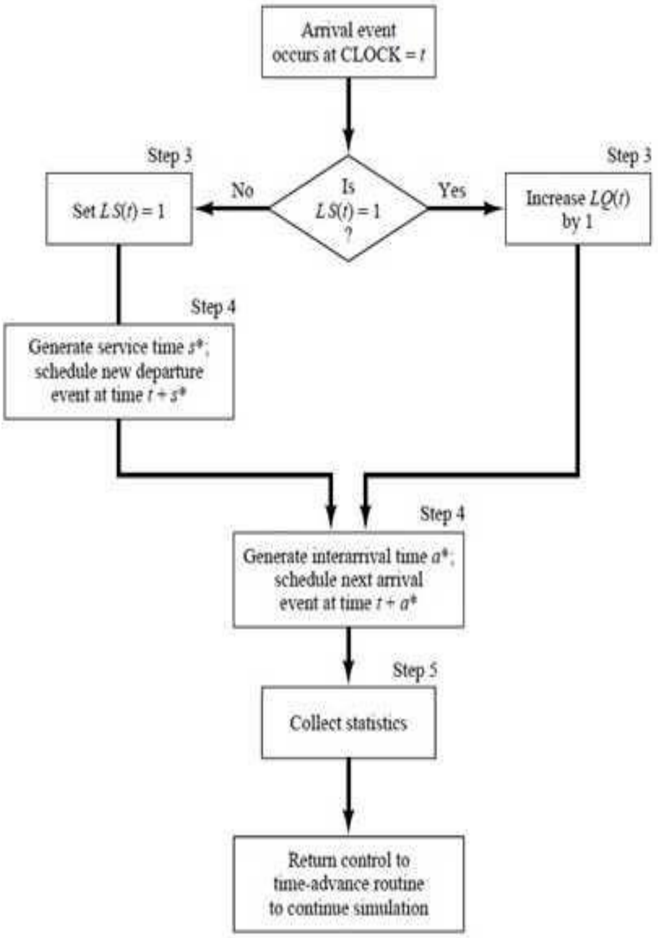


Figure Execution of the arrival event.

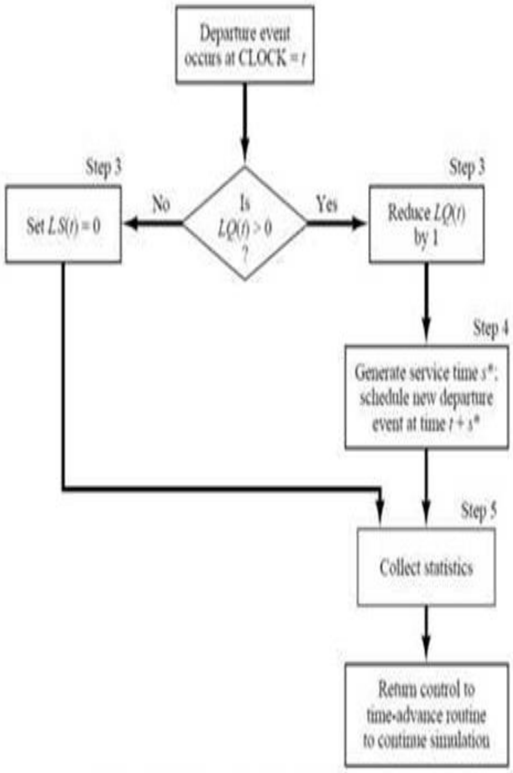


Figure Execution of the departure event.

Question Bank

1. When the simulation is appropriate tool & when it is not.
2. Advantages & disadvantages of simulation.
3. Components of systems & model and it types.
4. Steps in simulation study.
5. Examples (single server channel queue refer 2015, 2014, 2013 question paper, & class problem.
6. Examples Able & Bakes call center problem (two channel server problem)
7. Explain the terms used in discrete event simulation with an example(Ex. Able & Baker)
8. Explain the event scheduling algorithm by generating system snapshots at clock =t and clock=t1.
9. Explain the event scheduling algorithm with an example (single-channel-queue – execution of arrival event & execution of departure event).

UNIT-2 [NOTE : Check out time column should be included in prob. No 1 to 5. It is included in prob. 6 to 8.]

Problems on Event Scheduling / Time advanced Algorithm

1. Prepare a simulation table for 1 channel queuing s/m using ES/TA algorithm. Stopping event is at 30.

IAT : 8 6 1 8 2 8

ST : 4 1 4 3 2 4

LQ → Load Queue

LS → Load Service

Step 1 : Compute Departure Time

C.No	Inter Arrival Time	Arrival Time	Service Time	Departure Time
1	–	0	4	4
2	8	8	1	9 (8+1)
3	6	14	4	18 (14+4)
4	1	15	3	21 (18+3)
5	8	23	2	25 (23+2)
6	2	25	4	29 (25+4)
7	8	33	–	–

Step 2 : Simulation Table for 6 customers

Event Type	Clock	System State		Future Event List	Cumulative Statistics	
		LQ(E)	LS(E)		Busy	Maximum Queue
A_1	0	0	1	(D, 4) (A, 8) (E, 30)	0	0
D_1	4	0	0	(A, 8) (D, 9) (E, 30)	4	0
A_2	8	0	1	(D, 9) (A, 14) (E, 30)	4	0
D_2	9	0	0	(A, 14) (D, 18) (E, 30)	5	0
A_3	14	0	1	(A, 15) (D, 18) (E, 30)	5	0
A_4	15	1	1	(D, 18) (A, 23) (E, 30)	6	1
D_3	18	0	1	(D, 21) (A, 23) (E, 30)	9	1
D_4	21	0	0	(A, 23) (D, 25) (E, 30)	12	1
A_5	23	0	1	(D, 25) (A, 25) (E, 30)	12	1

A_6/D_5	25	0	1	(A, 33) (D, 29) (E, 30)	14	1
D_6	29	0	0	(E, 30) (A, 33)	14	1
~~D_6~~	~~29~~	~~0~~	~~0~~	~~(E, 30)~~	~~18~~	~~1~~

2. Stopping Time = 30

IAT : 2 4 3 1 2 5 6

ST : 4 3 2 5 2 1

i] Compute Departure Time

C. No	IAT	AT	ST	DT
1	–	0	4	4
2	2	2	3	7
3	4	6	2	9
4	3	9	5	14
5	1	10	2	16
6	2	12	1	17
7	5	17	–	–
8	6	23	–	–

ii] Simulation Table for 6 customers

Event Type	Clock	System State		Future Event List	CS	
		LQ(t)	LS(t)		B	MQ
A_1	0	0	1	(A, 2) (D, 4) (E, 30)	0	0
A_2	2	1	1	(D, 4) (A, 6) (E, 30)	2	1
D_1	4	0	1	(A, 6) (D, 7) (E, 30)	4	1
A_3	6	1	1	(D, 7) (A, 9) (E, 30)	6	1
D_2	7	0	1	(A, 9) (D, 9) (E, 30)	7	1
A_4/D_3	9	0	1	(A, 10) (D, 14) (E, 30)	9	1
~~A_4~~	[illegible]	[illegible]	[illegible]	~~(A, 10) (D, 14) (E, 30)~~	[illegible]	[illegible]

(2) A_6	12	2	1	(D, 14) (A, 17) (E, 30)	12	2
D_4	14	1	1	(A, 17) (D, 16) (E, 30)	14	2
D_5	16	1	1	(A, 17) (D, 17) (E, 30)	16	2
A_7/D_6	17	0	1	(A, 23) (E, 30)	17	2

3. Stopping Time = 60

IAT : 1 1 6 3 7 5 2 4 1

ST : 4 2 5 4 1 5 4 1 4

i) Compute Departure Time

C. No	IAT	AT	ST	DT
1	–	0	4	4
2	1	1	2	6
3	1	2	5	11
4	6	8	4	15
5	3	11	1	16
6	7	18	5	23
7	5	23	4	27
8	2	25	1	28
9	4	29	4	33
10	1	–	–	–

ii) Simulation Table for 9 customers

Event Type	Clock	S/m state LQ(t)	S/m state LS(t)	Future Event List	CS B	CS MQ
A_1	0	0	1	(A, 1) (D, 4) (E, 60)	0	0
A_2	1	1	1	(A, 2) (D, 4) (E, 60)	1	1
A_3	2	2	1	(D, 4) (A, 8) (E, 60)	2	2
D_1	4	1	1	(D, 6) (A, 8) (E, 60)	4	2
D_2	6	0	1	(A, 8) (D, 11) (E, 60)	6	2
A_4	8	1	1	(A, 11) (D, 11) (E, 60)	8	2

D_3/A_5	11	1	1	(D,15) (A,18) (E,60)	11	2
~~D_3~~	[illegible]	[illegible]	[illegible]	~~(D,15) (A,18) (E,60)~~	[illegible]	[illegible]
D_4	15	0	1	(D,16) (A,18) (E,60)	15	2
D_5	16	0	0	(A,18) (D,23) (E,60)	16	2
A_6	18	0	1	(D,23) (A,23) (E,60)	16	2
D_6	23	0	1	(A,23) (D,27) (E,60)	21	2
A_7	23	0	1	(D,27) (A,25) (E,60)	21	2
A_8	25	1	1	(D,27) (A,29) (E,60)	23	2
D_7	27	0	1	(A,29) (A,28) (E,60)	25	2
D_8	28	0	0	(A,29) (D,33) (E,60)	26	2
A_9	29	0	1	(D,33) (E,60)	26	2
D_9	33	0	0	(E,60)	30	2

4. Prepare a simulation table using ES/TA algorithm until the clock reaches time 23 using IAT & ST given :
Stopping Time is 30.

IAT : 5 1 2 3 4 9 5 8 6 1

ST : 4 7 8 1 4 2 5 3 1 4

i) Compute Departure Time

C. No	IAT	AT	ST	DT
1	–	0	4	4
2	5	5	7	12
3	1	6	8	20
4	2	8	1	21
5	3	11	4	25
6	4	15	2	27
7	9	24	5	32
8	5	29	3	35

C. No	IAT	AT	ST	DT
9	8	37	1	38
10	6	43	4	47
11	1	–	–	–

(3) ii] Simulation Table for 9 customers.

Event Type	Clock	S/m state		Future Event List	CS	
		$LQ(t)$	$LS(t)$		B	MQ
A_1	0	0	1	(A, 1) (D, 4)	0	0
A_2	1	1	1	(D, 4) (A, 7)	1	1
D_1	4	0	1	(D, 6) (A, 7)	4	1
D_2	6	0	0	(A, 7) (D, 12)	6	1
A_3	7	0	1	(A, 8) (D, 12)	6	1
A_4	8	1	1	(A, 11) (D, 12)	7	1
A_5	11	2	1	(D, 12) (A, 16)	10	2
D_3	12	1	1	(A, 16) (D, 17)	11	2
A_6	16	2	1	(D, 17) (A, 23)	15	2
D_4	17	1	1	(D, 18) (A, 23)	16	2
D_5	18	0	1	(D, 22) (A, 23)	17	2
D_6	22	0	0	(A, 23) (D, 27)	21	2
A_7	23	0	1	(A, 24) (D, 27)	21	2
A_8	24	1	1	(D, 27) (A, 28)	22	2
D_7	27	0	1	(A, 28) (D, 31)	25	2
A_9	28	1	1	(D, 31) (D, 32)	26	2
D_8	31	0	1	(D, 32)	29	2
D_9	32	1	0	–	30	2

S → Customer who taken response time from system

N_D → No. of customer departured

F → customer who spent more than given time (Eg: 4 min)

S = Response Time + Current Departure

Response Time = Clock – Current Arrival

ii) Simulation Table

Event Type	Clock	S/m state LQ(t)	LS(t)	Future Event List	CS B	MQ
A_1	0	0	1	(D, 4) (A, 5) (E, 30)	0	0
D_1	4	0	0	(A, 5) (D, 12) (E, 30)	4	0
A_2	5	0	1	(A, 6) (D, 12) (E, 30)	4	0
A_3	6	1	1	(A, 8) (D, 12) (E, 30)	5	1
A_4	8	2	1	(A, 11) (D, 12) (E, 30)	7	2
A_5	11	3	1	(D, 12) (A, 15) (E, 30)	10	3
D_2	12	2	1	(A, 15) (D, 20) (E, 30)	11	3
A_6	15	3	1	(D, 20) (A, 24) (E, 30)	14	3
D_3	20	2	1	(D, 21) (A, 24) (E, 30)	19	3
D_4	21	1	1	(A, 24) (D, 25) (E, 30)	20	3

5. Prepare a simulation table using Time Advanced Algorithm:

IAT : 1 6 1 3 5 7 1 4

ST : 4 2 5 5 1 4 4 4 1

i) Compute Departure Time

C. No	IAT	AT	ST	DT
1	–	0	4	4
2	1	1	2	6
3	6	7	5	12
4	1	8	5	17
5	3	11	1	18
6	5	16	4	22
7	7	23	4	27
8	1	24	4	31
9	4	28	1	32

6. Prepare a simulation table using Time advanced algorithm with IAT : 1 1 8 6 8 (4)

ST : 4 2 3 4 1 2

Find customers who spent more than 4 min.

i) Compute Departure Time

C. No	IAT	AT	ST	DT
1	-	0	4	4
2	1	1	2	6
3	1	2	3	9
4	8	10	4	14
5	6	16	1	17
6	8	24	2	26

ii) Simulation table of 6 customers.

Event type	Clock	S/m state		Check out time	Future Event List	CS				
		LQ(t)	LS(t)			S	N_D	F	B	MQ
A_1	0	0	1	$(C_1, 0)$	(A,1) (D,4)	0	0	0	0	0
A_2	1	1	1	$(C_1, 0)$ $(C_2, 1)$	(A,2) (D,4)	0	0	0	1	1
A_3	2	2	1	$(C_1, 0)$ $(C_2, 1)$ $(C_3, 2)$	(D,4) (A,10)	0	0	0	2	2
D_1	4	1	1	$(C_2, 1)$ $(C_3, 2)$	(D,6) (A,10)	4	1	0	4	2
D_2	6	0	1	$(C_3, 2)$	(D,9) (A,10)	9	2	1	6	2
D_3	9	0	0	-	(A,10) (D,14)	16	3	2	9	2
A_4	10	0	1	$(C_4, 10)$	(D,14) (A,16)	16	3	2	9	2
D_4	14	0	0	-	(A,16) (D,17)	20	4	2	13	2
A_5	16	0	1	$(C_5, 16)$	(D,17) (A,24)	20	4	2	13	2
D_5	17	0	0	-	(A,24) (D,26)	21	5	2	14	2
A_6	24	0	1	$(C_6, 24)$	(D,26)	21	5	2	14	2
D_6	26	0	0	-	-	23	6	2	16	2

7. Consider single server queue with one checkout counter using ES/TA algorithm

IAT : 4 2 8 1 8 3 6 8

ST : 4 6 5 2 3 4 4 1

Find the no. of customers who spent 4 or more ~~min~~ in the system. Stopping time = 32

i) Compute Arrival & Departure time

C·NO	IAT	AT	ST	DT
1	–	0	4	4
2	4	4	6	10
3	2	6	5	15
4	8	14	2	17
5	1	15	3	20
6	8	23	4	27
7	3	26	4	31
8	6	32	1	33
9	8	40	–	–

ii) Simulation table

Event type	clock	S/m state		checkout time	Future Event List	CS				
		LQ(t)	LS(t)			S	N_D	F	B	MQ
A_1	0	0	1	$(C_1,1)$	(A,4) (D,4) (E,32)	0	0	0	0	0
D_1/A_2	4	0	1	$(C_2,4)$	(A,6) (D,10) (E,32)	4	1	1	4	0
A_3	6	1	1	$(C_2,4)$ $(C_3,6)$	(D,10) (A,14) (E,32)	4	1	1	6	1
D_2	10	0	1	$(C_3,6)$	(A,14) (D,15) (E,32)	10	2	2	10	1
A_4	14	1	1	$(C_3,6)$ $(C_4,14)$	(A,15) (D,15) (E,32)	10	2	2	14	1
A_5/D_3	15	1	1	$(C_4,14)$ $(C_5,15)$	(D,17) (A,23) (E,32)	19	3	3	15	1
D_4	17	0	1	$(C_5,15)$	(D,20) (A,23) (E,32)	22	4	3	17	1
D_5	20	0	0	–	(A,23) (D,27) (E,32)	27	5	4	20	1
A_6	23	0	1	$(C_6,23)$	(A,26) (D,27) (E,32)	27	5	4	20	1

(5)	D_6	27	0	1	$(C_7, 26)$	(D,31) (A,32) (E,32)	31	6	5	24	1
	D_7	31	0	0	-	(A,32) (D,33) (E,32)	36	7	6	28	1
	A_8	32	0	1	$(C_8, 32)$	(A,40) (D,33) (E,32)	36	7	6	28	1

8. Develop a simulation table for single server queue with one check out counter using TA algorithm. Find busy time of server, maximum queue length, Total no. of customer who spent 3 min or more in system, Total number of departure.

IAT : 1 6 8 8 3 8 4 2 8

ST : 4 1 4 4 2 3 5 6 4

i) Compute Arrival and Departure time

C.No	IAT	AT	ST	DT
1	-	0	4	4
2	1	1	1	5
3	6	7	4	11
4	8	15	4	19
5	8	23	2	25
6	3	26	3	29
7	8	34	5	39
8	4	38	6	45
9	2	40	4	49
10	8	48	-	-

Event type	Clock	S/m state $LQ(t)$	S/m state $LS(t)$	Check Out Time	Future Event List	CS: S	CS: N_D	CS: F	CS: B	CS: MQ
A_1	0	0	1	$(C_1, 0)$	(A, 1) (D, 4)	0	0	0	0	0
A_2	1	1	1	$(C_1, 0)(C_2, 1)$	(A, 7) (D, 4)	0	0	0	1	1
D_1	4	0	1	$(C_2, 1)$	(A, 7) (D, 5)	4	1	1	4	1
D_2	5	0	0	–	(A, 7) (D, 11)	8	2	2	5	1
A_3	7	0	1	$(C_3, 7)$	(A, 15) (D, 11)	8	2	2	5	1
D_3	11	0	0	–	(A, 15) (D, 19)	12	3	3	9	1
A_4	15	0	1	$(C_4, 15)$	(D, 19) (A, 23)	12	3	3	9	1
D_4	19	0	0	–	(A, 23) (D, 25)	16	4	4	13	1
A_5	23	0	1	$(C_5, 23)$	(D, 25) (A, 26)	16	4	4	13	1
D_5	25	0	0	–	(A, 26) (D, 29)	18	5	4	15	1
A_6	26	0	1	$(C_6, 26)$	(D, 29) (A, 34)	18	5	4	15	1
D_6	29	0	0	–	(A, 34) (D, 39)	21	6	5	18	1
A_7	34	0	1	$(C_7, 34)$	(A, 38) (D, 39)	21	6	5	18	1
A_8	38	1	1	$(C_7, 34)(C_8, 38)$	(D, 39) (A, 40)	21	6	5	22	1
D_7	39	0	1	$(C_8, 38)$	(A, 40) (D, 45)	26	7	6	23	1
A_9	40	1	1	$(C_8, 38)(C_9, 40)$	~~(A, 48)~~ (D, 45)	26	7	6	24	1
D_8	45	0	1	$(C_9, 40)$	~~(A, 48)~~ (D, 49)	33	8	7	29	1
[illegible]	[illegible]	[illegible]	[illegible]	~~$(C_9, 40)(C_{10}, 48)$~~	~~(D, 49)~~	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
D_9	49	0	0	$(C_{10}, 48)$	–	42	9	8	33	1

Busy time of server = 33 min

Maximum Queue Length = 1

Total no. of customer who spent 3 or more in System = 8

Total no. of departure = 9

Travel Times as given below: Consider stopping time 32 clock cycle.
(OR) The stopping event will be completion of 2 weighings (2 A_{QW})

Loading Times	5	15	10	15	15	10	5
Weighing Times	17	15	15	15	15	12	
Travel Times	40	60	60	60	80		

Solution:
Simulation table for Dump-truck Operation.

Clock t	System State				Lists		Future Event List	Cumulative Statistics	
	LQ(t)	L(t)	WQ(t)	W(t)	Loader Queue	Weigh Queue		B_L	B_S
0	4	2	0	1	DT_4 DT_5 DT_6 DT_7	–	$(E_L, 5, DT_2)$* $(E_L, 15, DT_3)$ $(E_W, 17, DT_1)$	0	0
5	3	2	1	1	DT_5 DT_6 DT_7	DT_2	$(E_L, 15, DT_3)$* $(E_W, 17, DT_1)$ $(E_L, \underset{(5+10)}{15}, DT_4)$	10 $(5-0)*2+0$	5 $(5-0)+0$
15	2	2	2	1	DT_6 DT_7	DT_2 DT_3	$(E_W, 17, DT_1)$ $(E_L, 15, DT_4)$* $(E_L, \underset{(15+15)}{30}, DT_5)$	30 $(15-5)*2+10$	15 $(15-5)+5$
15	1	2	3	1	DT_7	DT_2 DT_3 DT_4	$(E_W, 17, DT_1)$* $(E_L, 30, DT_5)$ $(E_L, 30, DT_6)$	30	15
17	1	2	2	1	DT_7	DT_3 DT_4	$(E_L, 30, DT_5)$* $(E_L, 30, DT_6)$ $(A_Q, 57, DT_1)$ (17+40) $(E_W, 32, DT_2)$ (17+15)	34	17
30	0	2	3	1	–	DT_3 DT_4 DT_5	$(E_L, 30, DT_6)$* $(A_Q, 57, DT_1)$ $(E_W, 32, DT_2)$ $(E_L, 40, DT_7)$	60	30
30	0	1	4	1	–	DT_3 DT_4 DT_5 DT_6	$(A_{QL}, 57, DT_1)$ $(E_W, 32, DT_2)$* $(E_L, 40, DT_7)$	60	30
32	0	1	3	1	–	DT_4 DT_5 DT_6	$(A_{QL}, 57, DT_1)$ $(E_L, 40, DT_7)$ $(A_{QL}, 92, DT_2)$ $(E_W, 47, DT_3)$	62	32

2. Consider 6 Dump-trucks with loading times, weighing time & Traveling times are given below,

Loading Times	5	5	10	15	20	5	5
Weighing Times	12	15	20	12	15	15	
Travel Time	40	60	20	80			

- Until clock cycle 52
- Calculate
  i) Avg loader utilization
  ii) Avg scale utilization.

Solution:

Simulation table for Dump truck operations.

Clock t	System State LQ(t)	L(t)	WQ(t)	W(t)	Lists: Loader queue	Weigh queue	Future Event List	Cumulative Statistics B_L	B_S
0	3	2	0	1	DT_4 DT_5 DT_6	–	$(E_L, 5, DT_2)^*$ $(E_L, 5, DT_3)$ $(E_W, 12, DT_1)$	0	0
5	2	2	1	1	DT_5 DT_6	DT_2	$(E_L, 5, DT_3)^*$ $(E_W, 12, DT_1)$ $(E_L, 15, DT_4)$	10	5
5	1	2	2	1	DT_6	DT_2 DT_3	$(E_W, 12, DT_1)^*$ $(E_L, 15, DT_4)$ $(E_L, 20, DT_5)$	10	5
12	1	2	1	1	DT_6	DT_3	$(E_L, 15, DT_4)^*$ $(E_L, 20, DT_5)$ $(AQL, 52, DT_1)$ $(E_W, 27, DT_2)$	24	12
15	0	2	2	1	–	DT_3 DT_4	$(E_L, 20, DT_5)^*$ $(AQL, 52, DT_1)$ $(E_W, 27, DT_2)$ $(E_L, 35, DT_6)$	30	15
20	0	1	3	1	–	DT_3 DT_4 DT_5	$(AQL, 52, DT_1)$ $(E_W, 27, DT_2)^*$ $(E_L, 35, DT_6)$	40	20
27	0	1	2	1	–	DT_4 DT_5	$(AQL, 52, DT_1)$ $(E_L, 35, DT_6)^*$ $(AQL, 87, DT_2)$ $(E_W, 47, DT_3)$	47	27
35	0	0	3	1	–	DT_4 DT_5 DT_6	$(AQL, 52, DT_1)$ $(AQL, 87, DT_2)$ $(E_W, 47, DT_3)^*$	55	35

Clock t	System State				Lists		Future event List	Cumulative statistics	
	LQ(t)	L(t)	WQ(t)	W(t)	Loader queue	weigh queue		B_L	B_S
47	0	0	2	1	–	DT_5 DT_6	$(AQL, 52, DT_1)$ * $(AQL, 87, DT_2)$ $(AQL, 67, DT_3)$ $(E_W, 59, DT_4)$	55	47
52	0	1	2	1	–	DT_5 DT_6	$(AQL, 87, DT_2)$ $(AQL, 67, DT_3)$ $(E_L, 57, DT_1)$ [52+5] $(E_W, 59, DT_4)$	$\boxed{55}$	$\boxed{52}$

i) Average loader utilization $= \dfrac{55/2}{52} = \underline{0.529}$

ii) Average scale utilization $= \dfrac{52}{52} = \underline{1.00}$

---

3. Consider Loading Times, weighing Times & Travel times of 6 Dump trucks. Loading Times & weighing Times are based on FIFO, Until clock cycle 60 (or) entire system runs for 1 hours (or) estimation time is 60 mins.

Loading Times	10	10	15	5	10	5
weighing Times	10	15	5	10	10	
Travel Times	40	60	80	100		

* if all the trucks have equal value use FIFO method.

Solution:

Simulation table for Dump-truck operation.

Clock t	System State				Lists		Future event List	Cumulative Statistics	
	LQ(t)	L(t)	WQ(t)	W(t)	loaders queue	weigh queue		B_L	B_S
0	3	2	0	1	DT_4 DT_5 DT_6	–	$(E_L, 10, DT_2)$ $(E_L, 10, DT_3)$ $(E_W, 10, DT_1)$*	0	0
10	2	2	0	1	DT_5 DT_6	–	$(E_L, 10, DT_3)$* $(AQL, 50, DT_1)$ [10+40] $(E_W, 25, DT_2)$ [10+15] $(E_L, 25, DT_4)$ [10+15]	20	10
10	1	2	1	1	DT_6	DT_3	$(AQL, 50, DT_1)$ $(E_W, 25, DT_2)$ $(E_L, 25, DT_4)$ $(E_L, 15, DT_5)$* [10+5]	20	10

Clock	System State				Lists		Future Event List	Cumulative Statistics	
t	LQ(t)	L(t)	WQ(t)	W(t)	Loader Queue	Weigh Queue		B_L	B_S
15	0	2	2	1	–	DT_3 DT_5	$(A_{QL}, 50, DT_1)$ $(E_W, 25, DT_2)$* $(E_L, 25, DT_4)$ $(E_L, 25, DT_6)$	30	15
25	0	2	1	1	–	DT_5	$(A_{QL}, 50, DT_1)$ $(E_L, 25, DT_4)$* $(E_L, 25, DT_6)$ $(A_{QL}, 85, DT_2)$ $(E_W, 30, DT_3)$	50	25
25	0	1	2	1	–	DT_5 DT_4	$(A_{QL}, 50, DT_1)$ $(E_L, 25, DT_6)$* $(A_{QL}, 85, DT_2)$ $(E_W, 30, DT_3)$	50	25
25	0	0	3	1	–	DT_5 DT_4 DT_6	$(A_{QL}, 50, DT_1)$ $(A_{QL}, 85, DT_2)$ $(E_W, 30, DT_3)$*	50	25
30	0	0	2	1	–	DT_4 DT_6	$(A_{QL}, 50, DT_1)$ $(A_{QL}, 85, DT_2)$ $(A_{QL}, 110, DT_3)$ $(E_W, 40, DT_5)$*	50	30
40	0	0	1	1	–	DT_6	$(A_{QL}, 50, DT_1)$* $(A_{QL}, 85, DT_2)$ $(A_{QL}, 110, DT_3)$ $(A_{QL}, 140, DT_5)$ $(E_W, 50, DT_4)$	50	40
50	0	1	1	1	–	DT_6	$(A_{QL}, 85, DT_2)$ $(A_{QL}, 110, DT_3)$ $(E_L, 55, DT_1)$ $(E_W, 50, DT_4)$* $(A_{QL}, 140, DT_5)$	50	50
50	0	1	0	1	–	–	$(A_{QL}, 85, DT_2)$ $(A_{QL}, 110, DT_3)$ $(A_{QL}, 140, DT_5)$ $(A_{QL}, 90, DT_4)$ $(E_W, 60, DT_6)$ $(E_L, 55, DT_1)$*	50	50
55	0	0	1	1	–	DT_1	$(A_{QL}, 85, DT_2)$ $(A_{QL}, 110, DT_3)$ $(A_{QL}, 140, DT_5)$ $(A_{QL}, 90, DT_4)$ $(E_W, 60, DT_6)$*	55	55

Clock	System State				Lists		Future Event List	Cumulative Statistics	
t	$LQ(t)$	$L(t)$	$WQ(t)$	$W(t)$	Loader Queue	Weigh queue		B_L	B_S
60	0	0	0	1	—	—	$(ALQ, 85, DT_2)$ $(ALQ, 110, DT_3)$ $(ALQ, 140, DT_2)$ $(ALQ, 90, DT_4)$ $(ALQ, 120, DT_6)$ $(EW, 70, DT_1)$	55	60

UNIT 4: QUEUEING MODELS

4.1 Characteristics of Queueing System

- The key element's of queuing system are the **"customer and servers".**
- **Term Customer:** Can refer to people, trucks, mechanics, airplanes or anything that arrives at a facility and requires services.
- **Term Server:** Refer to receptionists, repairperson, medical personal, retrieval machines that provides the requested services.

4.1.1 Calling Population

- The population of potential customers referred to as the **"calling population".**
- The calling population may be assumed to be finite or infinite.
- The calling population is finite and consists
- In system with a large population of potential customers, the calling population is usually assumed to be infinite.
- The main difference between finite and infinite population models is how the arrival rate is defined.
- In an infinite population model, arrival rate is not affected by the number of customer who have left the calling population and joined the queueing.

4.1.2 System Capacity

- In many queueing system , there is a limit to the number of customers that may be in the waiting line or system.
- An arriving customer who finds the system full does not enter but returns immediately to the calling population.

4.1.3 Arrival Process

- The arrival process for "**Infinite population"** models is usually characterized in terms of interarrival time of successive customers.
- Arrivals may occur at scheduled times or at random times.
- When random times , the interarrival times are usually characterized by a probability distribution.
- Customer may arrive one at a time or in batches, the batches may be of constant size or random size.
- The second important class of arrivals is scheduled arrivals such as scheduled airline flight arrivals to an input.
- Third situation occurs when one at customer is assumed to always be present in the queue. So that the server is never idle because of a lack of customer.
- For finite population model, the arrivals process is characterized in a completely different fashion.
- Define customer as pending when that customer is outside the queueing system and a member of the calling population

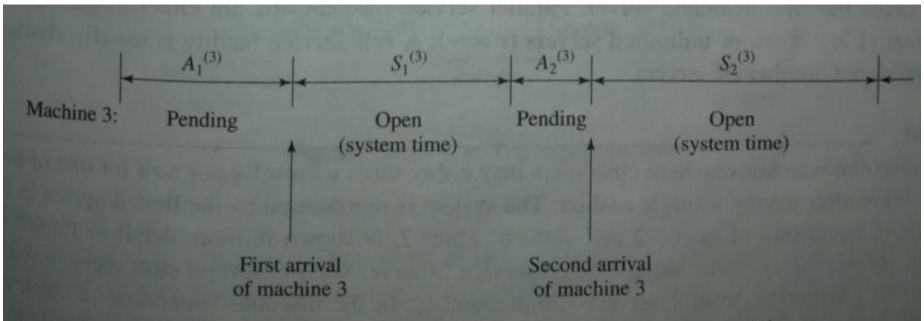


Figure 6.2 Arrival process for a finite-population model.

4.1.4 Queue Behavior and Queue Discipline

- It refers to the actions of customers while in a queue waiting for the service to begin.
- In some situations, there is a possibility that incoming customers will balk(leave when they see that the line is too long) , renege(leave after being in the line when they see that the line is moving slowly) , or jockey(move from one line to another if they think they have chosen a slow line).
- Queue discipline refers to the logical ordering of the customers in a queue and determines which customer will be chosen for service when a server becomes free.
- Common queue disciplines include FIFO, LIFO, service in random order(SIRO), shortest processing time first(SPT) and service according to priority (PR).

4.1.5 Service Times and Service Mechanism

- The service times of successive arrivals are denoted by s1, s2, sn.. They may be constant or of random duration.
- When {s1,s2,sn} is usually characterized as a sequence of independent and identically distributed random variables.
- The exponential, weibull, gamma, lognormal and truncated normal distribution have all been used successively as models of service times in different situations.
- A queueing system consists of a number of service centers and inter connecting queues. Each service center consists of some number of servers c, working in parallel.
- That is upon getting to the head of the line of customer takes the first available server.
- Parallel Service mechanisms are either single server or multiple server($1<c<\infty$) are unlimited servers($c=\infty$).
- A self service facility is usually characterized as having an unlimited number of servers.

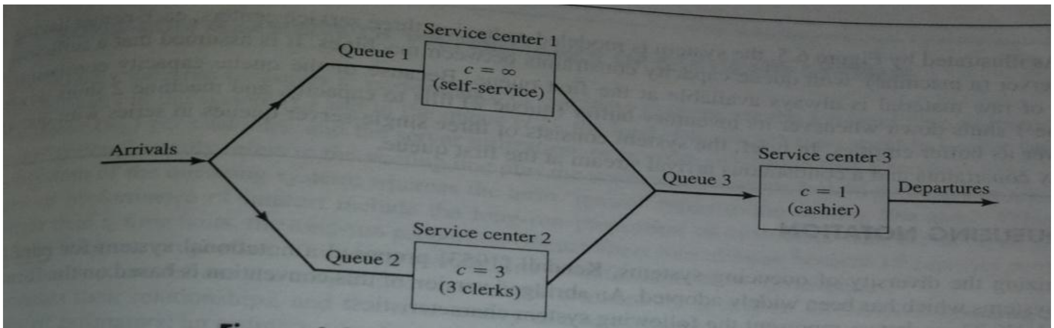


Figure 6.3 Discount warehouse with three service centers.

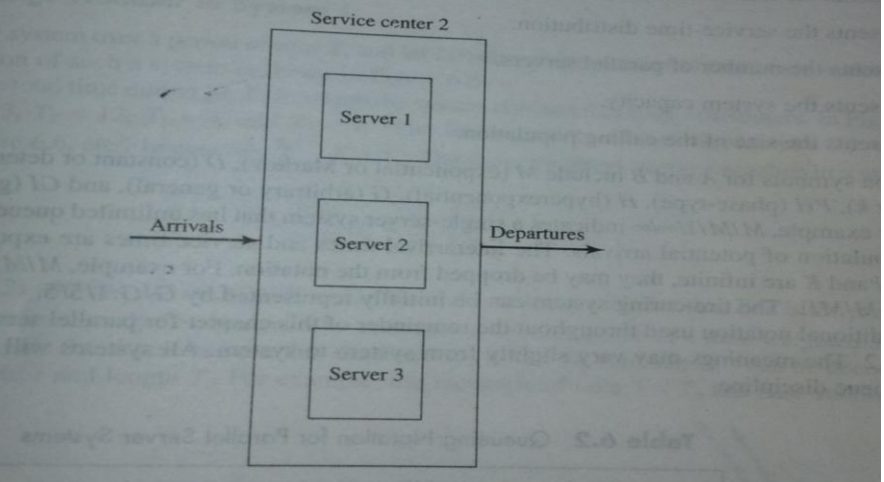


Figure 6.4 Service center 2, with $c = 3$ parallel servers.

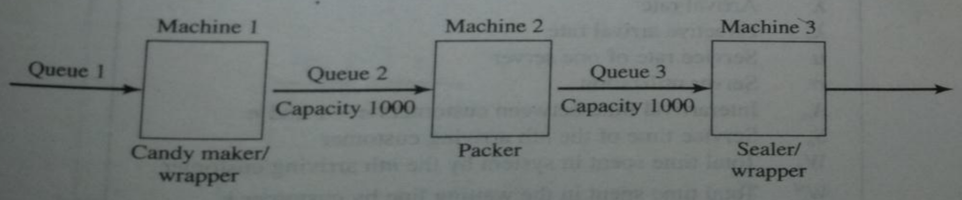


Figure 6.5 Candy-production line.

4.2 <u>Queueing Notation(Kendal's Notation)</u>

- Kendal's proposal a notational s/m for parallel server s/m which has been widely adopted.
- An a bridge version of this convention is based on format A|B|C|N|K
- These letters represent the following s/m characteristics:

A-Represents the InterArrival Time distribution
B-Represents the service time distribution
C-Represents the number of parallel servers
N-Represents the s/m capacity
K-Represents the size of the calling populations

Common symbols for A & B include M(exponential or Markov), D(constant or deterministic), E_k (Erlang of order k), PH (phase-type), H(hyperexponential), G(arbitrary or general), & GI(general independent).

- For eg, M|M|1|∞|∞ indicates a single server s/m that has unlimited queue capacity & an infinite population of potential arrivals
- The interarrival tmes & service times are exponentially distributed when N & K are infinite, they may be dropped from the notation.
- For eg, , M|M|1|∞|∞ is often short ended to M|M|1. The tire-curing s/m can be initially represented by G|G|1|5|5.

- Additional notation used for parallel server queueing s/m are as follows:

Table 6.2 Queueing Notation for Parallel Server Systems

P_n	Steady-state probability of having n customers in system
$P_n(t)$	Probability of n customers in system at time t
λ	Arrival rate
λ_e	Effective arrival rate
μ	Service rate of one server
ρ	Server utilization
A_n	Interarrival time between customers $n-1$ and n
S_n	Service time of the nth arriving customer
W_n	Total time spent in system by the nth arriving customer
W_n^Q	Total time spent in the waiting line by customer n
$L(t)$	The number of customers in system at time t
$L_Q(t)$	The number of customers in queue at time t
L	Long-run time-average number of customers in system
L_Q	Long-run time-average number of customers in queue
w	Long-run average time spent in system per customer
w_Q	Long-run average time spent in queue per customer

4.3 Long-run Measures of performance of queueing systems

- The primary long run measures of performance of queueing system are the long run time average number of customer in s/m(L) & queue(L_Q)
- The long run average time spent in s/m(w) & in the queue(w_Q) per customer
- Server utilization or population of time that a server is busy (p).

4.3.1 Time average Number in s/m (L):

- Consider a queueing s/m over a period of time T & let L(t) denote the number of customer I the s/m at time t.
- Let Ti denote the total time during[0,T] in which the s/m contained exactly I customers.

$$\boxed{\hat{L} = \sum_{i=0}^{\infty} i\left(\frac{T_i}{T}\right)}$$

- where $\hat{L}$ is the time weighted average number in a System. i

- Consider an example of Queueing s/m with line Segment 3, 12, 4, 1. Compute the time weighted - average number in a s/m.

Sol^n

$$\hat{L} = \sum_{i=0}^{\infty} i \left(\frac{T_i}{T}\right)$$

$$\hat{L} = \left[\, 0(3) + 1(12) + 2(4) + 3(1) \,\right] / 20$$

$$= 23/20$$

$$= 1.15 \text{ customers.}$$

4.3.2 Average Time spent in s/m per customer (w):

- Average s/m time is given as:

$$\boxed{\hat{w} = \frac{1}{N} \sum_{i=1}^{N} w_i} \quad \text{--- } (1)$$

where,

N - is the number of arrivals during [0,T]

w_i - is Customer spend in the s/m during [0,T]

- For stable s/m N-> ∞

$$\boxed{\hat{w} \rightarrow w} \quad \text{-- } (2)$$

With probability 1, where w is called the long-run average s/m time.

- Considering the equation 1 & 2 are written as,

$$\boxed{\hat{w}_Q = \frac{1}{N}\sum_{i=1}^{N} w_i^Q \to w_Q}$$

where,

w_i^Q = is the total time customer i spends waiting in queue.

$\hat{w}_Q$ - is the observed average time spent in queue.

w_Q - is the long run average delay per customer

example :- Consider the Queueing s/m with $N=5$ Customer arrive at $w_1 = 2$ & $w_5 = 20-16 = 4$ but w_2, w_3 & w_4 cannot be computed unless more is Know about the s/m. Arrival occur at times 0, 3, 5, 7 & 16 & departures occur at time 2, 8, 10 & 14.

Sol^n

$$\boxed{\hat{w} = \frac{1}{N}\sum_{i=1}^{N} w_i}$$

$w_1 = 2, \quad w_5 = 4$

$w_2 = 8 - 3 = 5$

$w_3 = 10 - 5 = 5$

$w_4 = 14 - 7 = 7$

$$\hat{w} = \frac{2 + 5 + 5 + 7 + 4}{5} = \frac{23}{5} = 4.6 \text{ time units.}$$

4.3.3 Server utilization:

- Server utilization is defined as the population of time server is busy
- Server utilization is denoted by ƥ is defined over a specified time interval[01]
- Long run server utilization is denoted by p

$\boxed{\text{Ƥ} \rightarrow P}$ as $T \rightarrow \infty$

- **Server utilization in G|G|C|∞|∞ queues**
  - Consider a queuing s/m with c identical servers in parallel
  - If arriving customer finds more than one server idle the customer choose a server without favoring any particular server.
  - The average number of busy servers say Ls is given by,

$\boxed{L_s = \lambda / \mu}$ $0 <= L_s <= C$

- The long run average server utilization is defined by

$$\boxed{P = \frac{L_s}{c} = \frac{\lambda}{c\mu}} \quad \therefore 0 \le P \le 1$$

- The utilization P can be interpreted as the proportion of time an arbitrary server is busy in the long run example :

Customer arrive at random to a license bureau at a rate of $\lambda = 50$ customer per hour. Currently there are 20 clerks, each serving $\mu = 5$ customers per hour on the average. Compute long-run or steady state average utilization of a server & average number of busy server.

Sol^n

Average utilization of server:

$$\boxed{\rho = \frac{\lambda}{c\mu}}$$

$$\rho = \frac{50}{20(5)} = 0.5$$

Average number of busy servers is:

$$\boxed{L_s = \frac{\lambda}{\mu}}$$

$$L_s = \frac{50}{5} = 10$$

4.4 STEADY-STATE BEHAVIOUR OF INFINITE-POPULATION MARKOVIAN MODLES

- For the infinite population models, the arrivals are assumed to follow a poisson process with rate λ arrivals per time unit
- The interarrival times are assumed to be exponentially distributed with mean 1/λ
- Service times may be exponentially distributed(M) or arbitrary(G)
- The queue discipline will be FIFO because of the exponential distributed assumptions on the arrival process, these model are called "MARKOVIAN MODEL".
- The steady-state parameter L, the time average number of customers in the s/m can be computed as

$$L = \sum_{n=0}^{\infty} nPn$$

Where Pn are the steady state probability of finding n customers in the s/m

- Other steady state parameters can be computed readily from little equation to whole system & to queue alone

$$w = L/\lambda$$
$$w_Q = w - (1/\mu)$$
$$L_Q = \lambda w_Q$$

Where λ is the arrival rate & μ is the service rate per server

4.4.1 SINGLE-SERVER QUEUE WITH POISSON ARRIVALS & UNLIMITED CAPACITY: M|G|1

- Suppose that service times have mean $1/\mu$ & variance σ^2 & that there is one server
- If $P = \lambda / \mu < 1$, then the M|G|1 queue has a steady state probability distribution with steady state characteristics
- The quantity $P = \lambda / \mu$ is the server utilization or lon run proportion of time the server is busy
- Steady state parameters of the M|G|1 are:

Notation	Description
① $P = \frac{\lambda}{\mu}$	• P is server ultilization • λ is arrival rate • μ is service rate
② $L = P + \frac{P^2(1+\sigma^2\mu^2)}{2(1-P)}$	• L is long run time average number of customers in s/m • σ is the mean service time
③ $w = \frac{1}{\mu} + \frac{\lambda(1/\mu^2+\sigma^2)}{2(1-P)}$	• w is long run average time spent in s/m per customer
④ $w_Q = \frac{\lambda(1/\mu^2+\sigma^2)}{2(1-P)}$	• wQ is long run average time spent in queue per customer
⑤ $L_Q = \frac{P^2(1+\sigma^2\mu^2)}{2(1-P)}$	• LQ is long run time avg no. of customer in queue
⑥ $P_0 = 1-P$	• P_0 is steady state probability of customers in s/m

Example : Consider a candy factory for making a candy at rate $\lambda = 1.5$ per hour. Observation over several months has found by the single m/c. It's mean service time $\sigma = 1/2$ hour, service rate is $\mu = 2$. Compute long run time average number of customer in s/m, long run time average number of customer in queue & long run average time spent in queue per customer.

Soln

↳ long run time average number of customer in s/m

$$\boxed{L = \rho + \frac{\rho^2 (1 + \sigma^2 \mu^2)}{2(1-\rho)}}$$

$$\rho = \frac{\lambda}{\mu} = 1.5/2 = 0.75$$

$$L = 0.75 + \frac{0.75 (1 + (0.5)^2 (2)^2)}{2(1 - 0.75)}$$

$$= 3.75$$

↳ long run time average number of customer in queue

$$\boxed{L_Q = \frac{\rho^2 (1 + \sigma^2 \mu^2)}{2(1-\rho)}}$$

$$L_Q = \frac{(0.75)^2 (1 + (0.5)^2 (2)^2)}{2(1 - 0.75)}$$

$$= 2.25$$

↳ long run average time spent in queue per customer:

$$w_Q = \frac{\lambda (1/\mu^2 + \sigma^2)}{2(1-\rho)}$$

$$w_Q = \frac{1.5 (1/(2)^2 + (0.5)^2)}{2(1 - 0.75)}$$

$$= 1.5$$

- Steady state parameters of the m/m/1 queue

Notation	Description
$L = \frac{\rho}{1-\rho}$	• L is long run time average number of customer in s/m • ρ is server utilization
$\omega = \frac{1}{\mu(1-\rho)}$	• ω is long run average time spent is s/m per customer • μ is service rate
$\omega_q = \frac{\rho}{\mu(1-\rho)}$	• ωq is long run average time spent in queue per customer
$L_q = \frac{\rho^2}{1-\rho}$	• Lq is long run time average number of customer in queue
$P_n = (1-\rho)\rho^n$	• Pn is steady state probability of n customer in s/m

4.4 2 MULTISERVER QUEUE: M|M|C|∞|∞

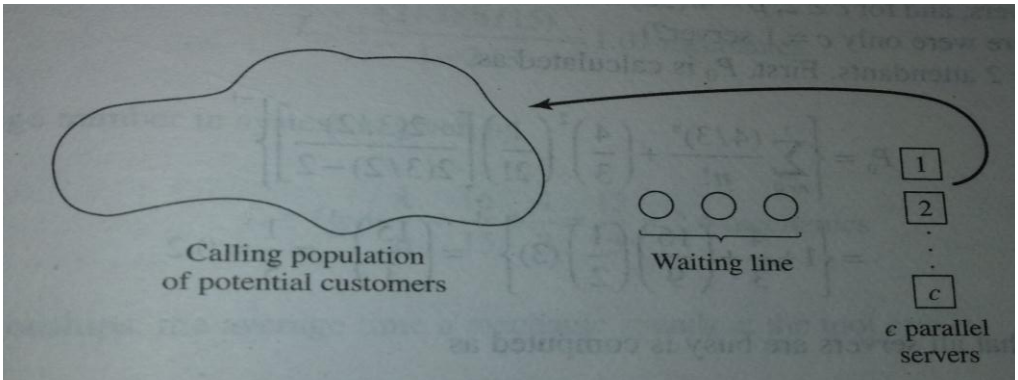


Figure 6.13 Multiserver queueing system.

- Suppose that there are c channels operating in parallel
- Each of these channels has an independent & identical exponential service time distribution with mean 1/μ
- The arrival process is poisson with rate λ. Arrival will join a single queue & enter the first available service channel

- For the M|M|C queue to have statistical equilibrium the offered load must satisfy $\lambda/\mu < c$ in which case $\lambda/(c\mu) = P$ the server utilization.

· The steady state parameters for the m/m/c queue

Notation	Description
· $\rho = \frac{\lambda}{\mu}$	· ρ is server utilization · λ arrival rate · μ service rate
· $P_0 = \left\{ \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} \right] + \left[(c\rho)^c \left(\frac{1}{c!} \right) \left(\frac{1}{1-\rho} \right) \right] \right\}^{-1}$	· is steady state for probability of customer in s/m
· $L = c\rho + \frac{\rho P(L(\infty) \geq c)}{(1-\rho)}$	· L is long run time average number of customer in s/m
$w = \frac{L}{\lambda}$	· w is long run average time spent is s/m per customer
$w_q = w - \frac{1}{\mu}$	· w_q is long run average time spent in queue per customer
$L_q = \frac{\rho P(L(\infty) \geq c)}{(1-\rho)}$	· L_q is long run time average number of customer in queue
$L - L_q = c\rho$	· 1

WHEN THE NUMBER OF SERVERS IS INFINITE (M|c|∞|∞)

- There are at least three situations in which it is appropriate to treat the number of server as infinite
    1. When each customer is its own server in other words in a self service s/m
    2. When service capacity far exceeds service demand as in a so called ample server s/m
    3. When wee want to know how many servers are required so that customer will rarely be delayed.

Steady state parameter for the M/G/∞ queue

Notation	description
$P_0 = e^{-\lambda/\mu}$	P_0 - probability of customer in s/m
$w = \frac{1}{\mu}$	w - long run average time spent in s/m
$w_q = 0$	w_q - long run average time spent in queue
$L = \lambda/\mu$	L - long run time average no of customer in s/m
$L_q = 0$	
$P_n = \frac{e^{-\lambda/\mu}(\lambda/\mu)^n}{n!}$	

4.5 STEADY STATE BEHAVIOR OF FINITE POPULATION MODELS (M|M|C|K|K)

- In many practical problems, the assumption of an infinite calling population leads to invalid results because the calling population is, in fact small.
- When the calling population is small, the presence of one or more customers in the system have a strong effect on the distribution of future arrivals and the use of an infinite population model can be misleading.
- Consider a finite calling population model with k customers. The time between the end of one service visit and the next call for service for each member of the population is assumed to be exponentially distributed with mean 1/ λ time units.
- Service times are also exponentially distributed, with mean 1/ μ time units. There are c parallel servers and system capacity is so that all arrivals remain for service. Such a system is shown in figure.

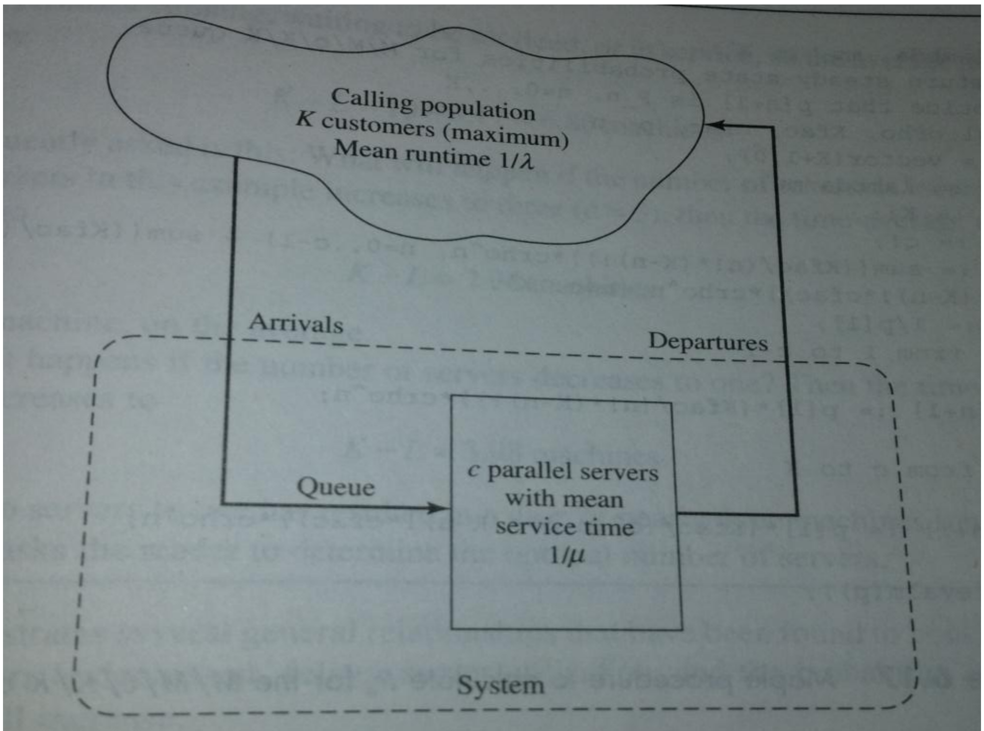


Figure 6.16 Finite-population queueing model.

The effective arrival rate λ_e has several valid interpretations:

Λ_e = long-run effective arrival rate of customers to queue
= long-run effective arrival rate of customers entering service
= long-run rate at which customers exit from service
= long-run rate at which customers enter the calling population
=long-run rate at which customers exit from the calling population.

Table 6.8 Steady-State Parameters for the M/M/c/K/K Queue

P_0	$\left[\sum_{n=0}^{c-1}\binom{K}{n}\left(\frac{\lambda}{\mu}\right)^n + \sum_{n=c}^{K}\frac{K!}{(K-n)!c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n\right]^{-1}$
P_n	$\begin{cases}\binom{K}{n}\left(\frac{\lambda}{\mu}\right)^n P_0, & n=0,1,\ldots,c-1\\ \frac{K!}{(K-n)!c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n P_0, & n=c,c+1,\ldots,K\end{cases}$
L	$\sum_{n=c}^{K} nP_n$
L_Q	$\sum_{n=c+1}^{K}(n-c)P_n$
λ_e	$\sum_{n=0}^{K}(K-n)\lambda P_n$
w	L/λ_e
w_Q	L_Q/λ_e
ρ	$\frac{L-L_Q}{c}=\frac{\lambda_e}{c\mu}$

4.6 NETWORKS OF QUEUE

- Many systems are naturally modeled as networks of single queues in which customer departing from one queue may be routed to another
- The following results assume a stable system with infinite calling population and no limit on system capacity.

1) Provided that no customers are created or destroyed in the queue,then the departure rate out of a queue is the same as the arrival rate into the queue over the long run.
2) If customers arrive to queue i at rate λi and a fraction $0 \leq p_{ij} \leq 1$ of them are routed to queue j upon departure, then the arrival rate from queue i to queue j is $\lambda_i p_{ij}$ is over long run
3) The overall arrival rate into queue j,λ_i is the sum of the arrival rate from all source.If customers arrive from outside the network at rate a_i then

$$\boxed{\lambda_j = a_j + \sum_{\text{all } i} \lambda_i P_{ij}}$$

4) If queue j has $ci < \infty$ parallel servers, each working at rate μ ,then the long run utilization of each server is

$$\boxed{\rho_j = \frac{\lambda_j}{c_j \mu_j}}$$

& Pj<1 is required for queue to be stable

5) If, for each queue j ,arrivals from outside the network form a poisson process with rate a and if there are ci identical services delivering exponentially distributed service times with mean $1/\mu$ then in steady state queue j behaves like a M|M|C; queue with arrival rate

$$\boxed{\lambda_j = a_j + \sum_{\text{all } i} \lambda_i P_{ij}}$$

UNIT 5:Random number generation And Variation Generation

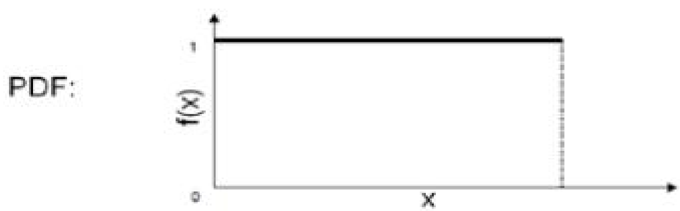
RANDOM-NUMBER GENERATION Random numbers are a necessary basic ingredient in the simulation of almost all discrete systems. Most computer languages have a subroutine, object, or function that will generate a random number. Similarly simulation languages generate random numbers that are used to generate event times and other random variables.

5.1 Properties of Random Numbers A sequence of random numbers, R1, R2... must have two important statistical properties, uniformity and independence. Each random number *Ri*, is an independent sample drawn from a continuous uniform distribution between zero and 1.

That is, the pdf is given by

$$\text{pdf:}\quad f(x) = \begin{cases} 1\,,\ 0 \le x \le 1 \\ 0\,,\ \text{otherwise} \end{cases}$$

The density function is shown below:



The expected value of Ri, is

$$E(R) = \int_0^1 x\,dx = [x^2/2]_0^1 = 1/2$$

The variance is given by 0

$$V(R) = \int_0^1 x^2 dx - [E(R)]^2$$

$$= [x^3/3]_0^1 - (1/2)^2 = 1/3 - 1/4$$

$$= 1/12$$

Some consequences of the uniformity and independence properties are the following:

1. If the interval (0, 1) is divided into n classes, or subintervals of equal length, the expected number of observations m each interval ii N/n where A' is the total number of observations.

2. The probability of observing a value in a particular interval is of the previous values drawn.

5.2 Generation of Pseudo-Random Numbers

Pseudo means false, so false random numbers are being generated. The goal of any generation scheme, is to produce a sequence of numbers between zero and 1 which simulates, or initiates, the ideal properties of uniform distribution and independence as closely as possible. When generating pseudo-random numbers, certain problems or errors can occur. These errors, or departures from ideal randomness, are all related to the properties stated previously. **Some examples include the following**

1) The generated numbers may not be uniformly distributed.

2) The generated numbers may be discrete -valued instead continuous valued

3) The mean of the generated numbers may be too high or too low.

4) The variance of the generated numbers may be too high or low

5) There may be dependence.

The following are examples:

a) Autocorrelation between numbers.

b) Numbers successively higher or lower than adjacent numbers.

c) Several numbers above the mean followed by several numbers below the mean.

Usually, random numbers are generated by a digital computer as part of the simulation. Numerous methods can be used to generate the values. In selecting among these methods, or routines, there are a number of important considerations.

1. The routine should be **fast.** The total cost can be managed by selecting a computationally efficient method of random-number generation.

2. The routine should be **portable** to different computers, and ideally to different programming languages .This is desirable so that the simulation program produces the same results wherever it is executed.

3. The routine should have a sufficiently **long cycle.** The cycle length, or period, represents the length of the random-number sequence before previous numbers begin to repeat themselves in an earlier order. Thus, if 10,000 events are to be generated, the period should be many times that long.

A special case cycling is degenerating. A routine degenerates when the same random numbers appear repeatedly. Such an occurrence is certainly unacceptable. This can happen rapidly with some methods.

4. The random numbers should be **replicable.** Given the starting point (or conditions), it should be possible to generate the same set of random numbers, completely independent of the system that is being simulated. This is helpful for debugging purpose and is a means of facilitating comparisons between systems.

5. Most important, and as indicated previously, the generated random numbers should closely approximate the ideal statistical properties of **uniformity and independences**

5.3 Techniques for Generating Random Numbers

5.3.1 The linear congruential method

It widely used technique, initially proposed by Lehmer [1951], produces a sequence of integers, X1, X2,... between zero and m — 1 according to the following recursive relationship:

$$X_{i+1} = (aX_i + c) \bmod m, \quad i = 0, 1, 2.... \quad (7.1)$$

The initial value **X0** is called the seed, **a** is called the constant multiplier, **c** is the increment, and **m** is the modulus.

If $c \neq 0$ in Equation (7.1), the form is called the **mixed congruential method.** When $c = 0$, the form is known as the **multiplicative congruential method**.

The selection of the values for a, c, m and X0 drastically affects the statistical properties and the cycle length. An example will illustrate how this technique operates.

EXAMPLE 1 Use the linear congruential method to generate a sequence of random numbers with *X0* = 27, *a*= 17, *c* = 43, and *m* = 100.

Here, the integer values generated will all be between zero and 99 because of the value of the modulus. These random integers should appear to be uniformly distributed the integers zero to 99.

Random numbers between zero and 1 can be generated by

Ri =Xi/m, i= 1,2,…… (7.2)

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

X0 = 27

X1 = (17*27 + 43) mod 100 = 502 mod 100 = 2 R1=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

Second, to help achieve maximum density, and to avoid cycling (i.e., recurrence of the same sequence of generated numbers) in practical applications, the generator should have the largest possible period. Maximal period can be achieved by the proper choice of a, c, m, and X0.

The max period (P) is:

- **For m a power of 2, say m = 2b, and c ¹≠0, the longest possible period is P = m = 2b,** which is achieved provided that c is relatively prime to m (that is, the greatest common factor of c and m is 1), and a = 1 + 4k, where k is an integer.
- **For m a power of 2, say m = 2b, and c = 0, the longest possible period is P = m / 4 = 2b-2,** which is achieved provided that the seed X0 is odd and the multiplier, a, is given by a = 3 + 8k or a = 5 + 8k, for some k = 0, 1,...
- **For m a prime number and c = 0, the longest possible period is P = m - 1**, which is achieved provided that the multiplier, a, has the property that the smallest integer k such that ak - 1 is divisible by m is
  k = m – 1.

Multiplicative Congruential Method:

Basic Relationship:

Xi+1 = a Xi (mod m), where a ≥ 0 and m ≥ 0 … (7.3)

Most natural choice for **m** is one that equals to the capacity of a computer word. m = 2b (binary machine), where b is the number of bits in the computer word.

m = 10d (decimal machine), where d is the number of digits in the computer word.

EXAMPLE 1: Let m = 102 = 100, a = 19, c = 0, and X0 = 63, and generate a sequence c random integers using Equation

Xi+1 = (aXi + c) mod m, i = 0, 1, 2....

X0 = 63 X1 = (19)(63) mod 100 = 1197 mod 100 = 97

X2 = (19) (97) mod 100 = 1843 mod 100 = 43

X3 = (19) (43) mod 100 = 817 mod 100 = 17 . . . .

When m is a power of 10, say m = 10b, the modulo operation is accomplished by saving the b rightmost (decimal) digits.

5.3.2 Combined Linear Congruential Generators

As computing power has increased, the complexity of the systems that we are able to simulate has also increased. One fruitful approach is to combine two or more multiplicative congruential generators in such a way that the combined generator has good statistical properties and a longer period. The following result from L'Ecuyer [1988] suggests how this can be done: If $W_{i,1}$, $W_{i,2}$,... , $W_{i,k}$ are any independent, discrete-valued random variables (not necessarily identically distributed), but one of them, say $W_{i,1}$, is uniformly distributed on the integers 0 to m_i — 2, then

$$W_i = \left(\sum_{j=1}^{k} (-1)^{j-1} W_{i,j} \right) \bmod m_1 - 1$$

is uniformly distributed on the integers 0 to m_i — 2. To see how this result can be used to form combined generators, let $X_{i,1}$, $X_{i,2}$,..., $X_{i,k}$ be the i *th output from k different multiplicative congruential generators, where the* j *th generator has prime modulus m_j, and the multiplier a_j is chosen so that the period is m_j — 1. Then the j'th generator is producing integers $X_{i,j}$ that are approximately uniformly distributed on 1 to m_j - 1, and $W_{i,j} = X_{i,j} - 1$ is approximately uniformly distributed on 0 to m_j - 2. L'Ecuyer [1988] therefore suggests combined generators of the form*

$$Xi = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) \bmod m_1 - 1$$

$$Ri = \begin{cases} \dfrac{X_i}{m_1}, X_i > 0 \\ \dfrac{m_1 - 1}{m_1}, X_i = 0 \end{cases}$$

Notice that the "$(-1)^{j-1}$" coefficient implicitly performs the subtraction $X_{i,1}-1$; for example, if k = 2, then

$$(-1)^0(X_{i,1}-1)-(-1)^1(X_{i,2}-1)=\sum_{j=1}^{2}(-1)^{j-1}X_{i,j}$$

The maximum possible period for such a generator is

$$p = \frac{(m_1-1)(m_2-1)\ldots(m_k-1)}{2^{k-1}}$$

5.4 Tests for Random Numbers

1. ***Frequency test***. Uses the Kolmogorov-Smirnov or the chi-square test to compare the distribution of the set of numbers generated to a uniform distribution.
2. ***Autocorrelation test***. Tests the correlation between numbers and compares the sample correlation to the expected correlation of zero.

5.4.1 Frequency Tests

A basic test that should always be performed to validate a new generator is the test of uniformity. Two different methods of testing are available.

1. Kolmogorov-Smirnov(KS test) and

2. Chi-square test.

• Both of these tests measure the degree of agreement between the distribution of a sample of generated random numbers and the theoretical uniform distribution.

• Both tests are on the null hypothesis of no significant difference between the sample distribution and the theoretical distribution.

1. The Kolmogorov-Smirnov test. This test compares the continuous cdf, F(X), of the uniform distribution to the empirical cdf, SN(x), of the sample of N observations. By definition,

F(x) = x, 0 ≤ x ≤ 1

If the sample from the random-number generator is R1 R2, ,..., RN, then the empirical cdf, SN(x), is defined by

$$S_n(x)\frac{\text{number of R1, R2, ..., Rn which are} \leq \text{x}}{N}$$

The Kolmogorov-Smirnov test is based on the largest absolute deviation between F(x) and SN(X) over the range of the random variable. That is. it is based on the statistic ***D = max |F(x) -SN(x)|*** *For testing against a uniform cdf, the test procedure follows these steps:*

Step 1: *Rank the data from smallest to largest. Let R (i) denote the i th smallest observation, so that*

$$R(1) \leq R(2) \leq \ldots \leq R(N)$$

Step 2: *Compute*

$$D^{+} = \max_{1 \le i \le n} \left\{ \frac{i}{N} - R_{(i)} \right\}$$

$$D^{-} = \max_{1 \le i \le n} \left\{ R_{(i)} - \frac{i-1}{N} \right\}$$

Step 3: *Compute D = max (D+, D-).*

Step 4: *Determine the critical value,* ***Dα,*** *from* ***Table A.8*** *for the specified significance level α and the given sample size N.*

Step 5:

$D \le D_\alpha$ Accept: No Difference between **S_N(x) and F(x)**

$D > D_\alpha$ Reject: No Difference between **S_N(x) and F(x)**

We conclude that no difference has been detected between the true distribution of {R1, R2,... RN} and the uniform distribution.

EXAMPLE 6: *Suppose that the five numbers* ***0.44, 0.81, 0.14, 0.05, 0.93*** *were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance* ***α of 0.05.***

Step 1: *Rank the data from smallest to largest. 0.05, 0.14, 0.44, 0.81, 0.93*

Step 2: *Compute* ***D+ and D-***

	R_i	$\frac{i}{N}$	$D^{+} = \max_{1 \le i \le n} \left\{ \frac{i}{N} - R_{(i)} \right\}$	$D^{-} = \max_{1 \le i \le n} \left\{ R_{(i)} - \frac{i-1}{N} \right\}$
1	0.05	0.20	0.15	0.05
2	0.14	0.40	0.26	~
3	0.44	0.60	0.16	0.04
4	0.81	0.80	~	0.21
5	0.93	1.00	0.07	0.13

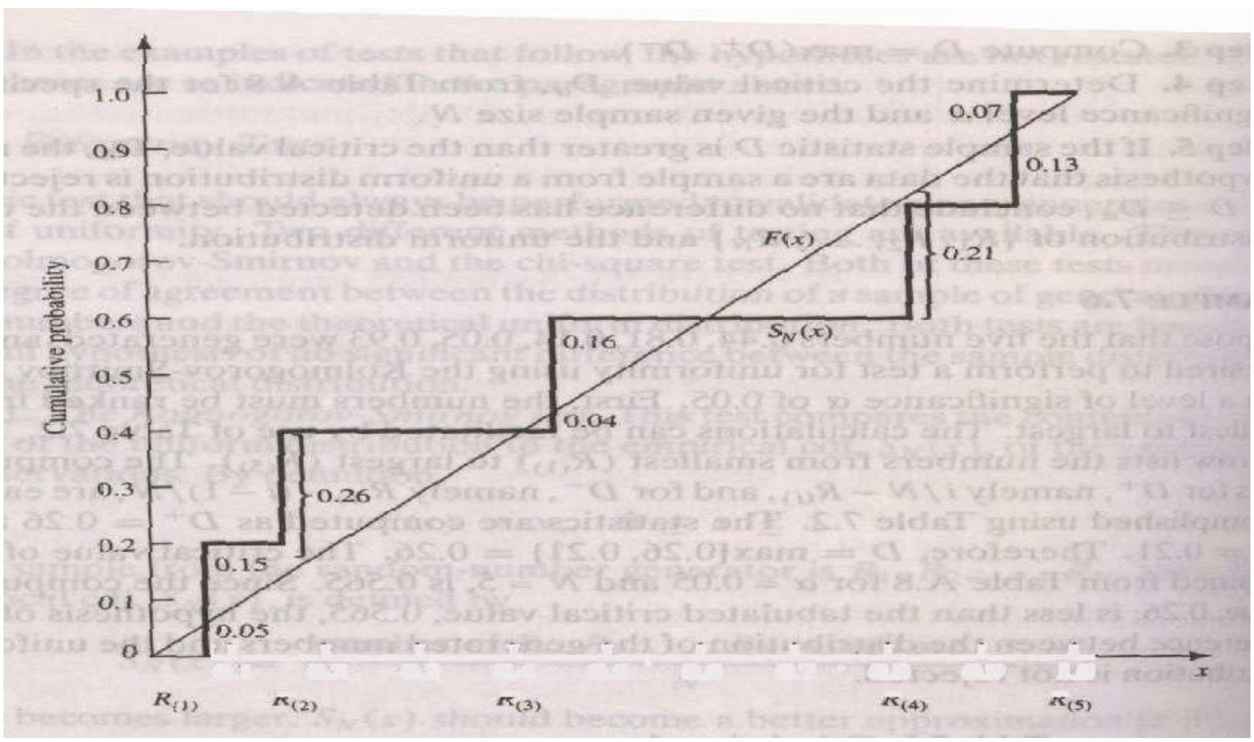
Step3: Compute D = max (D+, D-)

. D=max (0.26, 0.21) = 0.26

Step 4: Determine the critical value, Dα, from Table A.8 for the specified significance level α and the given sample size N. **Here α=0.05, N=5 then value of Dα = 0.565**

Step 5: Since the computed value, 0.26 is less than the tabulated critical value, 0.565,

the hypothesis of no difference between the distribution of the generated numbers and the uniform distribution is not rejected.



compare F(x) with Sn(X)

2. The chi-square test.

The chi-square test uses the sample statistic

$$\chi_0^2 = \sum_{i=0}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Where, Oi is observed number in the i th class

Ei is expected number in the i th class,

$$E_i = \frac{N}{n}$$

N – No. of observation

n – No. of classes

Note: sampling distribution of χ_0^2 approximately the chi square has n-1 degrees of freedom

Example 7: *Use the chi-square test with α = 0.05 to test whether the data shown below are uniformly distributed. The test uses n = 10 intervals of equal length, namely [0, 0.1), [0.1, 0.2)... [0.9, 1.0).* ***(REFER TABLE A.6)***

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
0.72	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

Interval	Range	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	0.0-0.1	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	0.0
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	0.0
8	0.7-0.8	14	10	4	16	1.6
9	0.8-0.9	10	10	0	0	0.0
10	0.9-1.0	11	10	1	1	0.1
		100	100	0		3.4

The value of χ_0^2 is 3.4. This is compared with the critical value $\chi^2_{0.05,9}$ = 16.9.Since χ_0^2 is much smaller than the tabulated value of $\chi^2_{0.05,9}$, the null hypothesis of a uniform distribution is not rejected.

5.4.2 Tests for Auto-correlation

The tests for auto-correlation are concerned with the dependence between numbers in a sequence. The list of the 30 numbers appears to have the effect that every 5th number has a very large value. If this is a regular pattern, we can't really say the sequence is random.

0.12	0.01	0.23	0.28	0.89	0.31	0.64	0.28	0.83	0.93
0.99	0.15	0.33	0.35	0.91	0.41	0.60	0.27	0.75	0.88
0.68	0.49	0.05	0.43	0.95	0.58	0.19	0.36	0.69	0.87

The test computes the auto-correlation between every m numbers (m is also known as the lag) starting with the ith number. Thus the autocorrelation ρ_{im} between the following numbers would be of interest.

$$R_i, R_{i+m}, R_{i+2m}, \ldots, R_{i+(M+1)m}$$

Form the test statistic $Z_0 = \frac{\hat{\rho}_{im}}{\sigma_{\hat{\rho}_{im}}}$ which is distributed normally with a mean of zero and a variance of one.

The actual formula for $\hat{\rho}_{im}$ and the standard deviation is $\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} R_{(k+1)m}\right] - 0.25$ and

$$\sigma_{\hat{\rho}_{im}} = \frac{\sqrt{13M+7}}{12(M+1)}$$

After computing Z_0, do not reject the null hypothesis of independence if

$$-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$$

where α is the level of significance.

EXAMPLE : Test whether the 3rd, 8th, 13th, and so on, numbers in the sequence at the beginning of this section are auto correlated. (Use a = 0.05.) Here, i = 3 (beginning with the third number), m = 5 (every five numbers), N = 30 (30 numbers in the sequence), and M = 4 (largest integer such that 3 + (M +1)5 < 30).

0.12	0.01	0.23	0.28	0.89	0.31	0.64	0.28	0.83	0.93
0.99	0.15	0.33	0.35	0.91	0.41	0.60	0.27	0.75	0.88
0.68	0.49	0.05	0.43	0.95	0.58	0.19	0.36	0.69	0.87

Solution:

$$\rho_{im} = \frac{1}{4+1}\left[(0.23)(0.28)+(0.28)(0.33)+(0.33)(0.27)+(0.27)(0.05)+(0.05)(0.36)\right]-0.25$$
$$= -0.1945$$

And

$$\sigma_{\rho_{im}} = \frac{\sqrt{13(4)+7}}{12(4+1)} = 0.1280$$

Then, test for statistic assumes the value

$$Z_0 = -\frac{0.1945}{0.1280} = -1.516$$

Now the critical value from Table A.3 is $Z_{0.025}$=1.96

Therefore, the hypothesis of independence can't be rejected on the basis of this test.

2.Random Variate Generation TECHNIQUES:

• INVERSE TRANSFORMATION TECHNIQUE

• ACCEPTANCE-REJECTION TECHNIQUE

All these techniques assume that a source of uniform (0,1) random numbers is available R1,R2….. where each R1 has probability density function and cumulative distribution function.

Note: The random variable may be either discrete or continuous.

2.1 Inverse Transform Technique

The inverse transform technique can be used to sample from exponential, the uniform, the Weibull and the triangle distributions.

2.1.1 Exponential Distribution

The exponential distribution, has probability density function (pdf) given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & 0 \le x \\ 0, & x < 0 \end{cases}$$

and cumulative distribution function (cdf) given by

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$= \begin{cases} 1 - e^{-\lambda x}, & 0 \le x \\ 0, & x < 0 \end{cases}$$

The parameter λ can be interpreted as the mean number of occurrences per time unit. For example, if interarrival times X_1, X_2, X_3 . . . had an exponential distribution with rate, and then could be interpreted as the mean number of arrivals per time unit, or the arrival rate. For any i,

$$E(X_i) = 1/\lambda$$

And so 1/λ is mean inter arrival time. The goal here is to develop a procedure for generating values X1, X2, X3 . . . which have an exponential distribution.

The inverse transform technique can be utilized, at least in principle, for any distribution. But it is most useful when the cdf. F(x), is of such simple form that its inverse, F^{-1}, can be easily computed.

A step-by-step procedure for the inverse transform technique illustrated by me exponential distribution, is as follows:

Step 1: Compute the cdf of the desired random variable X. For the exponential distribution, the cdf is

$$F(x) = 1-e^{-\lambda x}, \; x \ge 0.$$

Step 2: Set F(X) = R on the range of X. For the exponential distribution, it becomes

$$1 - e^{-\lambda X} = R \text{ on the range } x \ge 0.$$

Since X is a random variable (with the exponential distribution in this case), so $1-e^{-\lambda x}$ is also a random variable, here called R. As will be shown later, R has a uniform distribution over the interval (0,1).,

Step 3: Solve the equation F(X) = R for X in terms of R. For the exponential distribution, the solution proceeds as follows:

$$1 - e^{-\lambda x} = R$$

$$e^{-\lambda x} = 1 - R$$

$$-\lambda X = \ln(1 - R)$$

$$x = -1/\lambda \ln(1 - R) \qquad \ldots(5.1)$$

Equation (5.1) is called a random-variate generator for the exponential distribution. In general, Equation (5.1) is written as $X=F^{-1}(R)$. Generating a sequence of values is accomplished through steps 4.

Step 4: Generate (as needed) uniform random numbers R1, R2, R3,... and compute the desired random variates by

$$\mathbf{Xi = {}_{F}^{-1} (Ri)}$$

For the exponential case, $\mathbf{F^{-1} (R) = (-1/\lambda)ln(1- R)}$ by Equation (5.1),

so that **Xi = -1/λ ln (1 – Ri) …(5.2)** for i = 1,2,3,.... One simplification that is usually employed in Equation (5.2) is to replace 1 – Ri by Ri to yield **Xi = -1/λ ln Ri …(5.3)** which is justified since both Ri and 1- Ri are uniformly distributed on (0,1).

Example: consider the random number As fellows, where λ=1

R_i	0.1306	0.0422	0.6597	0.7965	0.7696

Solution:

Using equation compute Xi

$$x = -1/\lambda \ln(1 - R)$$

I	1	2	3	4	5
R_i	0.1306	0.0422	0.6597	0.7965	0.7696
X_i	0.1400	0.0431	1.078	1.592	1.468

Uniform Distribution :

Consider a random variable X that is uniformly distributed on the interval [a, b]. A reasonable guess for generating X is given by

$$X = a + (b - a)R \quad \ldots\ldots\ldots 5.5$$

[Recall that R is always a random number on (0,1).

The pdf of X is given by

$$f(x) = \begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

The derivation of Equation (5..5) follows steps 1 through 3 of Section 5.1.1:

Step 1. The cdf is given by

$$F(x) = \begin{cases} 0, & x < a \\ (x-a)/(b-a), & a \le x \le b \\ 1, & x > b \end{cases}$$

Step 2. Set F(X) = (X - a)/(b -a) = R

Step 3. Solving for X in terms of R yields

$$X = a + (b - a)R,$$

which agrees with Equation (5.5).

Weibull Distribution:

The weibull distribution was introduce for test the time to failure of the machine or electronic components. The location of the parameters V is set to 0.

$$f(x) = \begin{cases} \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/\alpha)^\beta}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

where α>0 and β>0 are the scale and shape of parameters.

Steps for Weibull distribution are as follows:

step 1: The cdf is given by

$$F(X) = 1 - e^{-(x/\alpha)^\beta}, x \geq 0.$$

step2 :set f(x)=R

$$1 - e^{-(X/\alpha)^\beta} = R.$$

step 3:Solving for X in terms of R yields.

$$X = \alpha[-\ln(1-R)]^{1/\beta}$$

<u>Empirical continuous distribution:</u>

Respampling of data from the sample data in systamtic manner is called empirical continuos distribution.

Step1:Arrange data for smallest to largest order of interval

x(i-1)<x<X(i) i=0,1,2,3….n

Step2:Compute probability 1/n

Step3:Compute cumulative probability i.e i/n where n is interval

step4:calculate a slope i.e

without frequency ai=x(i)-x(i-1)/(1/n)

with frequency ai= x(i)-x(i-1)/(c(i)-c(i-1) where c(i) is cumulative probability

2.1 Acceptance-Rejection technique

- Useful particularly when inverse cdf does not exist in closed form
- Illustration: To generate random variants, *X* ~ U(*1/4*, *1*)
- Procedures:

Step 1: Generate a random number R ~ U [0, 1]

Step 2a: If R ≥ ¼, accept X=R.

Step 2b: If R < ¼, reject R, return to Step 1

- R does not have the desired distribution, but R conditioned (R') on the event {R ³ ¼} does.
- Efficiency: Depends heavily on the ability to minimize the number of rejections.

2.1.1 Poisson Distribution *A Poisson random variable, N, with mean a > 0 has pmf*

$$p(n) = P(N = n) = \frac{e^{-\alpha}\alpha^{n}}{n!}, \; n = 0, 1, 2, \ldots$$

- N can be interpreted as number of arrivals from a Poisson arrival process during one unit of time
- Then time between the arrivals in the process are exponentially distributed with rate α.
- Thus there is a relationship between the (discrete) Poisson distribution and the (continuous) exponential distribution, namely

$$N = n \iff \sum_{i=1}^{n} A_i \le 1 < \sum_{i=1}^{n+1} A_i$$

$$\sum_{i=1}^{n} A_i \le 1 < \sum_{i=1}^{n+1} A_i \iff \sum_{i=1}^{n} -\frac{1}{\alpha}\ln R_i \le 1 < \sum_{i=1}^{n+1} -\frac{1}{\alpha}\ln R_i$$

$$\iff \prod_{i=1}^{n} R_i \ge e^{-\alpha} > \prod_{i=1}^{n+1} R_i$$

The procedure for generating a Poisson random variate, N, is given by the following steps:

Step 1: *Set n = 0, and P = 1*

Step 2: *Generate a random number Rn+1 and let P = P. Rn+1*

Step 3: *If $P < e^{-\alpha}$, then accept N = n. Otherwise, reject current n, increase n by one, and return to step 2*

Example: *Generate three Poisson variants with mean a =0.2 for the given Random number*

0.4357,0.4146,0.8353,0.9952,0.8004

Solution:

Step 1.Set n = 0, P = 1.

tep 2.R1 = 0.4357, P = 1 • R1 = 0.4357.

Step 3. Since P = 0.4357 < e^{-b} = 0.8187, accept N = 0. ***Repeat Above procedure***

n	R_{n+1}	p	accept/reject	Result
0	0.4357	0.4357	$P<e^{-\alpha}$ (accept)	N=0
0	0.4146	0.4146	$P<e^{-\alpha}$ (accept)	N=0
0	0.8353	0.8353	$P\geq e^{-\alpha}$ (reject)	
1	0.9952	0.8313	$P\geq e^{-\alpha}$ (reject)	
2	0.8004	0.6654	$p < e^{-\alpha}$ (accept)	N=2

Gamma distribution:

Is to check the random variants are accepted or rejected based on dependent sample data.

Steps 1: Refer the steps which given in problems.

Unit-5

Random Number Generation & Random Variate Generation

Problems :

① Generate a sequence of 5 integer Random number with $a=19$, $m=100$, $x_0=63$, $\boxed{c=1}$ (mixed Linear Congruential Method)

Solution : $\boxed{X_{i+1} = (a x_i + c) \bmod m,}$ $i = 0 \ldots m-1$

→ $i=0$ $X_1 = (19 \times 63 + 1) \bmod 100$

$= 1198 \bmod 100 = 98$

$i=1$ $\boxed{R_i = \frac{x_i}{m}, \quad i = 1 \ldots m}$

$R_1 = \frac{98}{100} = 0.98$

→ $i=1$ $X_2 = (19 \times 98 + 1) \bmod 100 = 63$

$i=2$ $R_2 = \frac{63}{100} = 0.63$

→ $i=2$ $X_3 = (19 \times 63 + 1) \bmod 100 = 98$

$i=3$ $R_3 = 98/100 = 0.98$

→ $i=3$ $X_4 = (19 \times 98 + 1) \bmod 100 = 63$

$i=4$ $R_4 = 63/100 = 0.63$

→ $i=4$ $X_5 = (19 \times 63 + 1) \bmod 100 = 98$

$i=5$ $R_5 = 98/100 = 0.98$

∴ Random numbers are 0.98, 0.63, 0.98, 0.63, 0.98

(2) Using Multiplicative Congruential Method, Generate sequence of 5 integer number where $a = 24$, $m = 100$, seed value $= 64$.

Solution:

Given
$a = 24$
$m = 100$
$x_0 = 64$
$c = 0$

$X_{i+1} = (a X_i + c) \bmod m, \quad i = 0 \dots m-1$

$R_i = \frac{X_i}{m}, \quad i = 1 \dots m$

$X_1 = (24 \times 64 + 0) \bmod 100 = 36$

$R_1 = 36/100 = \underline{\underline{0.36}}$

$X_2 = (24 \times 36 + 0) \bmod 100 = 64$

$R_2 = 64/100 = \underline{\underline{0.64}}$

$X_3 = (24 \times 64 + 0) \bmod 100 = 36$

$R_3 = 36/100 = \underline{\underline{0.36}}$

$X_4 = (24 \times 36 + 0) \bmod 100 = 64$

$R_4 = 64/100 = \underline{\underline{0.64}}$

$X_5 = (24 \times 64 + 0) \bmod 100 = 36$

$R_5 = 36/100 = \underline{\underline{0.36}}$

$\therefore$ Random numbers are 0.36, 0.64, 0.36, 0.64, 0.36

(3) Using **K.S Test** with $\alpha = 0.05$, to Test wether the data shown below are uniformly distributed 0.44, 0.81, 0.14, 0.05, 0.93 & also plot graph for $f(x)$.

Solution:

given, $\alpha = 0.05$
$N = 5$ (No. of random number given)

i	$R_{(i)}$	i/N	$\frac{i-1}{N}$	$D^+ = \max_{1 \le i \le N} \{ \frac{i}{N} - R_i \}$	$D^- = \max_{1 \le i \le N} \{ R_i - \frac{i-1}{N} \}$
1	0.05	0.20	0	0.15	0.05
2	0.14	0.40	0.20	$\boxed{0.26}$ max.	~
3	0.44	0.60	0.40	0.16	0.04
4	0.81	0.80	0.60	~	$\boxed{0.21}$ max.
5	0.93	1.00	0.80	0.07	0.13

$$D = \max(D^+, D^-)$$
$$= \max(0.26, 0.21) = \underline{0.26}$$

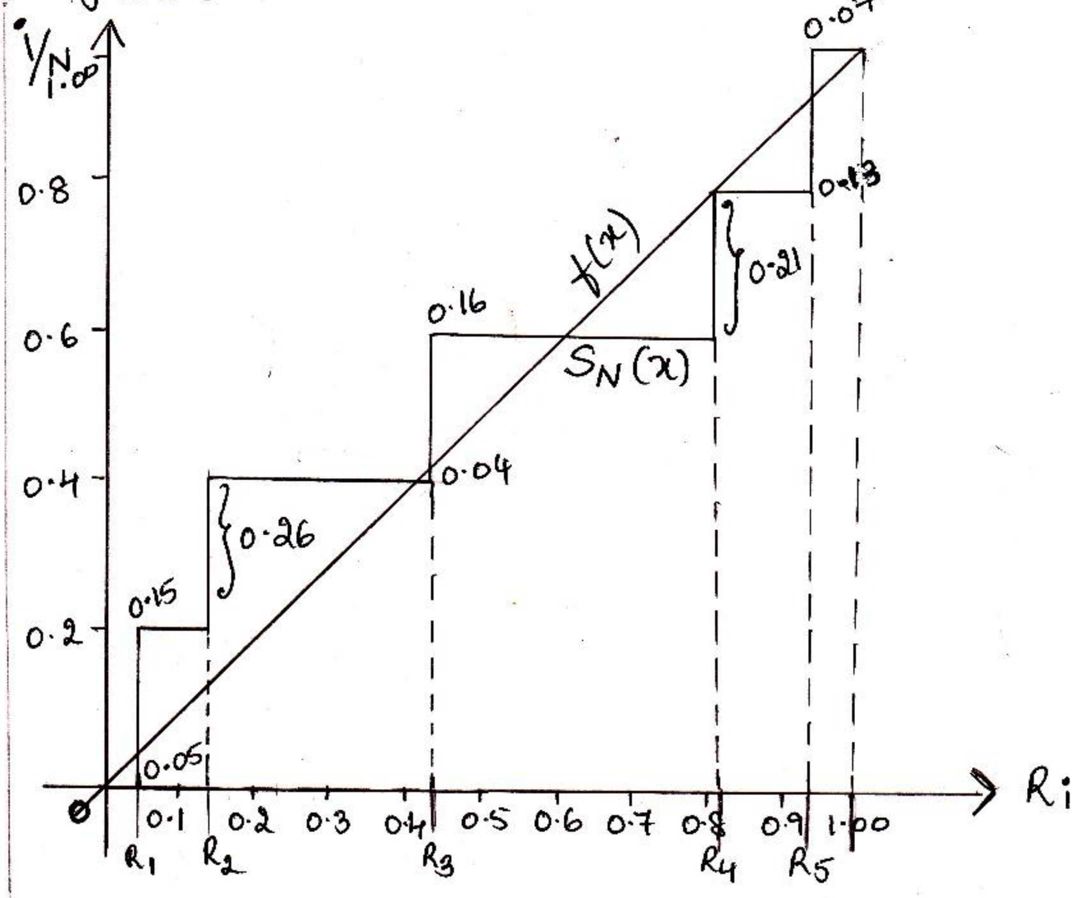
$D_{\alpha, N} \Rightarrow D_{0.05, 5} = 0.565$, from table A8

$\therefore D \le D_\alpha \Rightarrow 0.26 \le 0.565$

$\therefore$ Accepted Null hypothesis

hence Random Numbers are uniformly distributed.

graph :



(4) Using K.S Test with $\alpha = 0.05$, & sample data are given below also plot graph for $f(x)$.

0.46, 0.18, 0.23, 0.64, 0.36, 0.44

Solution: given $\alpha = 0.05$, $N = 6$.

i	R_i	i/N	$\frac{i-1}{N}$	$D^+ = \max_{1 \le i \le N} \{ i/N - R_i \}$	$D^- = \max_{1 \le i \le N} \{ R_i - \frac{i-1}{N} \}$
1	0.18	0.16	0	~	$\boxed{0.18}$ max.
2	0.23	0.33	0.16	0.1	0.07
3	0.36	0.5	0.33	0.14	0.03
4	0.44	0.66	0.5	0.22	~
5	0.46	0.83	0.66	$\boxed{0.37}$ max.	~
6	0.64	1.00	0.83	0.36	~

$\therefore D = \max(D^+, D^-) \Rightarrow \max(0.37, 0.18) = 0.37$

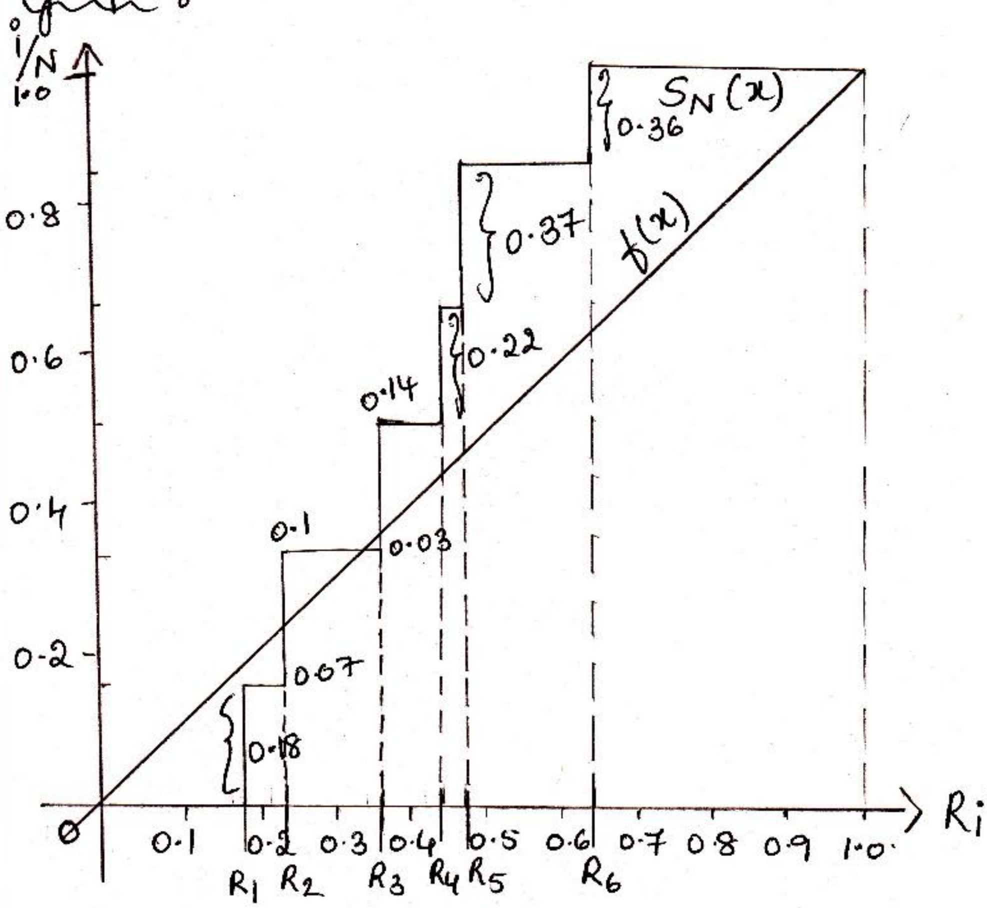
from table A8, $D_{\alpha, N} = D_{0.05, 6} = 0.521$

$\therefore D \leq D_\alpha = 0.37 \leq 0.521$

$\therefore$ Accepted Null hypothesis

hence random numbers are uniformly distributed.

Graph :



---

(5) Using chi-square Test with level of significance $\alpha = 0.05$, Test wether given data are uniformly distributed. with test uses 10 interval of equal length. total number of sample data are 100.

VVI

Interval	1	2	3	4	5	6	7	8	9	10
Observation	8	8	10	9	12	8	10	14	10	11

<u>Solution :</u>

given, $\alpha = 0.05$, $n = 10$, $N = 100$

$$E_i = \frac{N}{n} = \frac{100}{10} = 10$$

interval	O_i	$O_i - E_i$	$(O_i - E_i)^2$	E_i	$X_0^2 = \sum \frac{(O_i - E_i)^2}{E_i}$
1	8	-2	4	10	0.4
2	8	-2	4	10	0.4
3	10	0	0	10	0
4	9	-1	1	10	0.1
5	12	2	4	10	0.4
6	8	-2	4	10	0.4
7	10	0	0	10	0
8	14	4	16	10	1.6
9	10	0	0	10	0
10	11	1	1	10	0.1
Sum	100				3.4

$\therefore X_0^2 = 3.4$

from table A6, $X^2_{\alpha, n-1} = X^2_{0.05, 9} = 16.9$

$\therefore X_0^2 \leq X^2_{\alpha, n-1} = 3.4 \leq 16.9$

$\therefore$ Accepted null hypothesis

(6) use chi-square Test with $\alpha = 0.05$ where $n = 10$, intervals of equal length. sample data are given below:

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, 0.4
0.10, 0.17, 0.10, 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,
0.72, 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.

Solution: given, $\alpha = 0.05$, $n = 10$, $N = 100$

$E_i = N/n = 100/10 = 10$

interval	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum \frac{(O_i - E_i)^2}{E_i}$
0.01-0.10	10	10	0	0	0
0.11-20	8	10	-2	4	0.4
0.21-30	10	10	0	0	0
0.31-40	9	10	-1	1	0.1
0.41-50	12	10	2	4	0.4
0.51-60	8	10	-2	4	0.4
0.61-70	10	10	0	0	0
0.71-80	14	10	4	16	1.6
0.81-90	10	10	0	0	0
0.91-1.00	9	10	-1	1	0.1

$X_0^2 = 3$

from Table A6, $X^2_{\alpha, n-1} = X^2_{0.05, 9} = 16.9$

$\therefore X_0^2 \leq X^2_{\alpha, n-1}$

$3 \leq 16.9$

$\therefore$ Accepted Null hypothesis

(7) Using Auto correlation Test to test wether numbers are uniformly distributed with starting period 3rd, 8th, 13th & so on and largest integer number is 4. $Z_{\alpha/2} = 1.96$. the sample data are given below:

0.12, 0.01, 0.23, 0.28, 0.89, 0.31, 0.64, 0.28, 0.83, 0.93, 0.99, 0.15, 0.33, 0.35, 0.91, 0.41, 0.60, 0.27, 0.75, 0.88, 0.68, 0.49, 0.05, 0.43, 0.95, 0.58, 0.19, 0.36, 0.69, 0.87.

Solution: given, $i = 3$ (period starts from 3rd)

$m = 5$ (difference b/w periods i.e 8-3, 13-8 ...)

$M = 4$ (largest number)

$Z_{\alpha/2} = 1.96$

$$\boxed{\hat{P}_{im} = \frac{1}{M+1}\left[\sum_{K=0}^{M} R_{i+Km} R_{i+(K+1)m}\right] - 0.25}$$

$$= \frac{1}{4+1}\left[(0.23)(0.28) + (0.28)(0.33) + (0.33)(0.27) + (0.27)(0.05) + (0.05)(0.36)\right] - 0.25$$

$$= -0.1945$$

$$\boxed{\sigma_{\hat{P}_{iM}} = \frac{\sqrt{13M+7}}{12(M+1)}}$$

$$= \frac{\sqrt{13 \times 4 + 7}}{12(4+1)} = 0.1280$$

$$\boxed{Z_0 = \frac{\hat{P}_{iM}}{\sigma_{\hat{P}_{iM}}}} = \frac{-0.1945}{0.1280} = -1.5196$$

$\therefore$ $-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$

$-1.96 \leq -1.5196 \leq 1.96$

$\therefore$ Accepted Null Hypothesis

(8). Generate 5 Random Variation from sample data 0.30, 0.20, 0.10, 0.50, 0.60 using Exponential distribution with mean = 1

Imp.

Solution: given $\lambda = 1$

i	1	2	3	4	5
R_i	0.30	0.20	0.10	0.50	0.60
X_i	0.3566	0.2231	0.1053	0.6931	0.9162

$$\boxed{X_i = -\frac{1}{\lambda} \ln(1-R_i)}$$

$$X_1 = -\frac{1}{1} \ln(1-0.30) = 0.3566$$

$$X_2 = -\frac{1}{1} \ln(1-0.20) = 0.2231$$

$$X_3 = \frac{-1}{1} \ln(1-0.10) = \underline{\underline{0.1053}}$$

$$X_4 = \frac{-1}{1} \ln(1-0.50) = \underline{\underline{0.6931}}$$

$$X_5 = \frac{-1}{1} \ln(1-0.60) = \underline{\underline{0.9162}}$$

---

(9) Generate 5 Random Variation using Uniform Distribution technique with interval $0.3 \leq x \leq 2$.
consider sample data, 0.30, 0.25, 0.80, 0.75, 2.5

Solution: given, $a = 0.3 \quad b = 2$

$$\boxed{X_i = a + (b-a) R_i \quad , i = 1 \ldots n}$$

i	1	2	3	4	5
R_i	0.30	0.25	0.80	0.75	2.5
X_i	0.81	0	1.66	1.575	1

$X_1 = 0.3 + (2-0.3)\,0.3 = \underline{\underline{0.81}}$

$X_2 = 0.25 < 0.3 = \underline{\underline{0}}$

$X_3 = 0.3 \leq 0.8 \leq 2 \Rightarrow 0.3 + (2-0.3)\,0.80 = \underline{\underline{1.66}}$

$X_4 = 0.3 \leq 0.75 \leq 2 \Rightarrow 0.3 + (2-0.3)\,0.75 = \underline{\underline{1.575}}$

$X_5 = 2.5 > 2 \Rightarrow \underline{\underline{1}}$

(10) Generate 6 random variation using Weibull Distribution with slope = 10 & shape = 2. the sample data are 0.10, 0.60, 0.50, 0.80, 0.20, 0.45

Solution: given $\alpha = 10$ (slope)
$\beta = 2$ (shape)

$$\boxed{X_i = \alpha\left[-\ln(1-R_i)\right]^{1/\beta}}$$

i	1	2	3	4	5	6.
R_i	0.10	0.60	0.50	0.80	0.20	0.45
X_i	3.24	9.57	8.32	12.68	4.72	7.73

$$X_1 = 10 \times \left[-\ln(1-0.10)\right]^{1/2} = \underline{\underline{3.24}}$$

$$X_2 = 10\left[-\ln(1-0.60)\right]^{1/2} = \underline{\underline{9.57}}$$

$$X_3 = 10\left[-\ln(1-0.50)\right]^{1/2} = \underline{\underline{8.32}}$$

$$X_4 = 10\left[-\ln(1-0.80)\right]^{1/2} = \underline{\underline{12.68}}$$

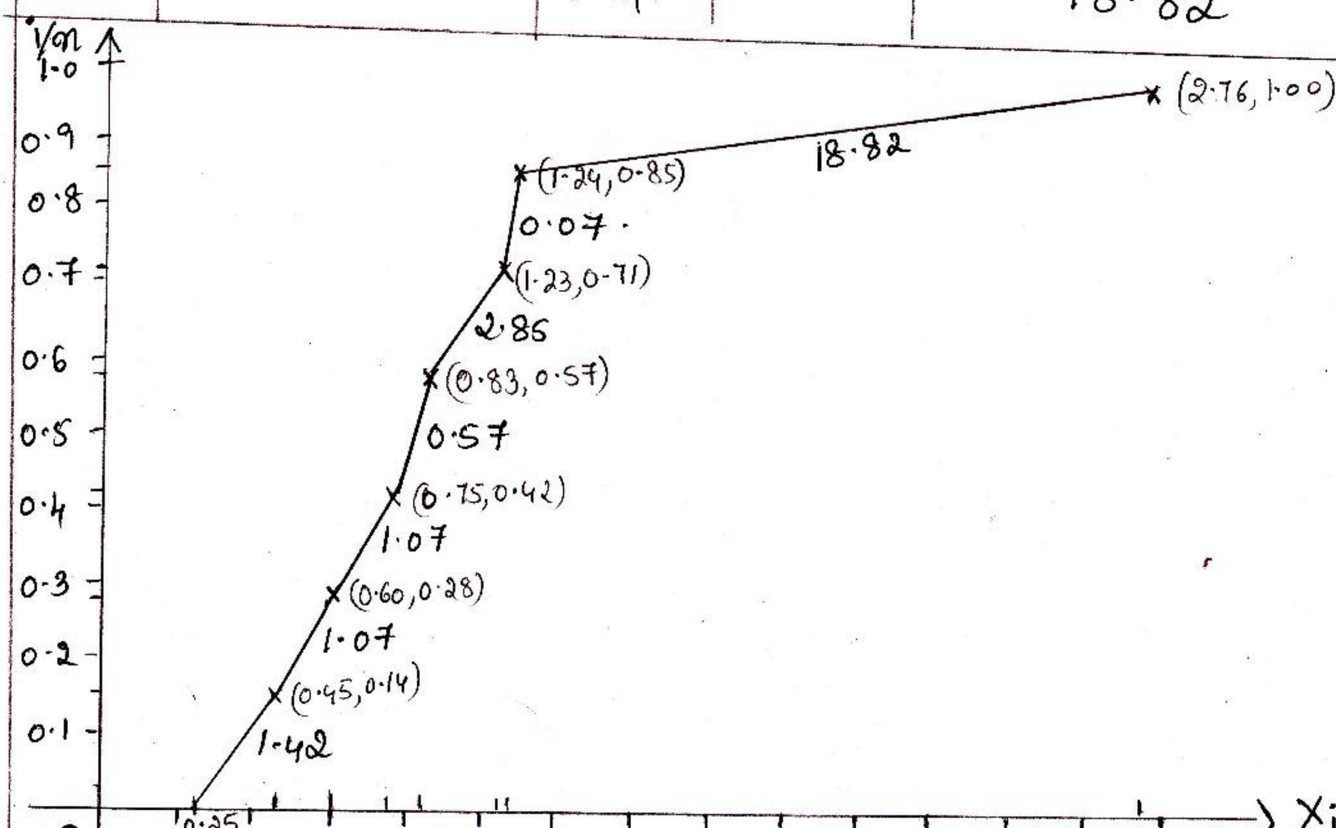
$$X_5 = 10\left[-\ln(1-0.20)\right]^{1/2} = \underline{\underline{4.72}}$$

$$X_6 = 10\left[-\ln(1-0.45)\right]^{1/2} = \underline{\underline{7.73}}$$

(11) consider the data : 0.83, 0.45, 1.23, 0.60, 0.75, 2.76, 1.24 with the probability = $1/n$ where $x_0 = 0.25$. Find slope of ith line segment using Emphirical Continuous distribution method. (without frequency)

Solution:

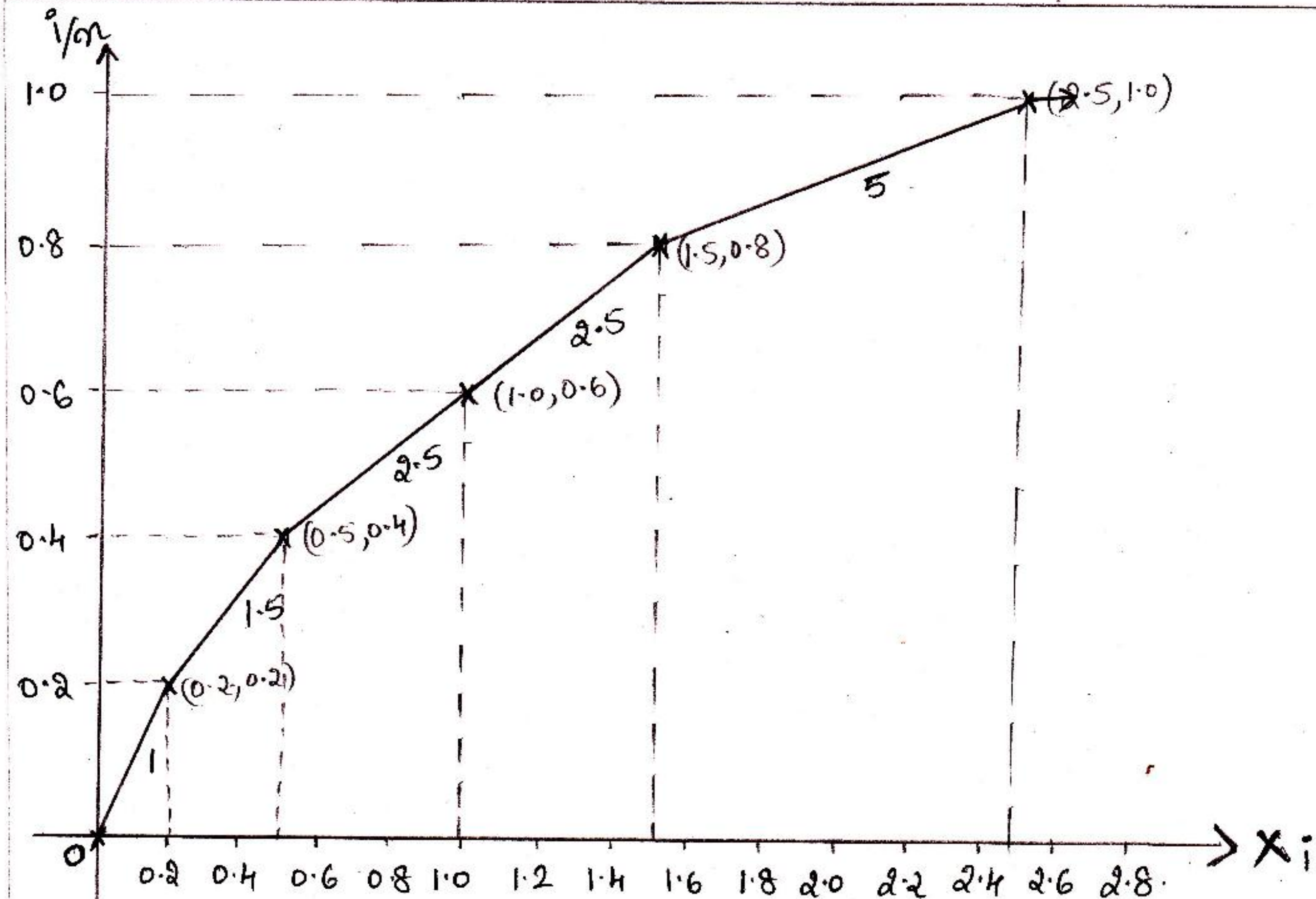
i	$x_{i-1} \leq x \leq x_i$	$1/n$	i/n	$a_i = \frac{x_i - x_{i-1}}{1/n}$
1	$0.25 \leq x \leq 0.45$	0.14	0.14	$a_i = \frac{0.45-0.25}{0.14} = 1.42$
2	$0.45 \leq x \leq 0.60$	0.14	0.28	1.07
3	$0.60 \leq x \leq 0.75$	0.14	0.42	1.07
4	$0.75 \leq x \leq 0.83$	0.14	0.57	0.57
5	$0.83 \leq x \leq 1.23$	0.14	0.71	2.85
6	$1.23 \leq x \leq 1.24$	0.14	0.85	0.07
7	$1.24 \leq x \leq 2.76$	0.14	1.00	18.82



(12) Consider the data 1·0, 0·5, 0·20, 1·5, 2·5 & Frequency are 31, 10, 25, 24, 30. Find the slope of ith line segment using Empirical Continuous Distribution (with frequency)

Solution: Consider $x_0 = 0.0$ & $c_0 = 0.0$

i	$x_{i-1} \leq x \leq x_i$	Frequency	Relative Frequency	Cumulative Probability (i/n)	$a_i = \frac{x_i - x_{i-1}}{c_i - c_{i-1}}$
1	$0.0 \leq x \leq 0.2$	25	0·25	0·2	$a_i = \frac{0.2-0.0}{0.2-0.0} = 1$
2	$0.2 \leq x \leq 0.5$	10	0·10	0·4	$\frac{0.5-0.2}{0.4-0.2} = 1.5$
3	$0.5 \leq x \leq 1.0$	31	0·31	0·6	$\frac{1.0-0.5}{0.6-0.4} = 2.5$
4	$1.0 \leq x \leq 1.5$	24	0·24	0·8	$\frac{1.5-1.0}{0.8-0.6} = 2.5$
5	$1.5 \leq x \leq 2.5$	30	0·30	1·00	$\frac{2.5-1.5}{1.0-0.8} = 5$



(13) 10m * Generate 3 poisson variate with mean 0.2. consider random numbers 0.4357, 0.4146, 0.8353, 0.9952, 0.8004.
vvimp

Solution : Steps : (2marks)

Step 0: Set $n=0$, $p=1$

2: $R_1 = 0.4357$ $\quad P = P \cdot R_1 = 1 \times 0.4357 = 0.4357$

$e^{-\alpha} = e^{-0.2} = 0.8187$

3: $P < e^{-\alpha} = 0.4357 < 0.8187$ $\boxed{\therefore \text{ Accept } N=0}$

Step 1: $n=0$, $p=1$

2: $R_2 = 0.4146$ $\quad P = P \cdot R_2 = 1 \times 0.4146 = 0.4146$

3: $0.4146 < 0.8187$ $\boxed{\therefore \text{ Accept } N=0}$

Step 1: $n=0$, $p=1$

2: $R_3 = 0.8353$ $\quad p = 1 \times 0.8353 = 0.8353$

3: $0.8353 < 0.8187$ $\boxed{\therefore \text{ Reject } n=1}$ $\quad p = 0.8353.$

$n=1$, $p = 0.8353$

Step 2: $R_4 = 0.9952$ $\quad p = P \cdot R_4 = 0.8353 \times 0.9952 = 0.8312$

3: $0.8312 < 0.8187$ $\boxed{\therefore \text{ Reject } n=2}$

$n=2$, $p = 0.8312$

Step 2: $R_5 = 0.8004$ $\quad p = 0.8312 \times 0.8004 = 0.6652$

3: $0.6652 < 0.8187$ $\boxed{\therefore \text{ Accept } N=2}$

* table : (8 marks)

n	P	R_{n+1}	$P = P \cdot R_{n+1}$	$P < e^{-\alpha}$ A/R	$N = n$.
0	1	0.4357	0.4357	0.4357 < 0.8187 Accept	N=0
0	1	0.4146	0.4146	0.4146 < 0.8187 Accept	N=0
0	1	0.8353	0.8353	0.8353 < 0.8187 Reject	–
1	0.8353	0.9952	0.8312	0.8312 < 0.8187 Reject	–
2	0.8312	0.8004	0.6652	0.6652 < 0.8187 Accept	N=2

∴ the 3 poisson variates are N=0, N=0, N=2.

(14) Generate 5 poisson variate with mean = 0.2

Solution : Steps :

$e^{-\alpha} = e^{-0.2} = 0.8187$.

$n = 5$

*) If random numbers are not given, then the random numbers are selected so that it satisfies following conditions : If x is random number then,

$\frac{1}{4} \leq x \leq 1$ & $x \leq e^{-\alpha}$ then accepted, else reject.

Step 1: $n=0$, $p=1$

2: $R_1 = 0.3568$ $\quad p = p \times R_1 = 1 \times 0.3568 = 0.3568$

3: $0.3568 \leq 0.8187$ & $0.25 \leq 0.3568 \leq 1$ $\boxed{\therefore \text{Accept } N=0}$

Step 1: $n=0$ $\quad p=1$

2: $R_2 = 0.4123$ $\quad p = 1 \times 0.4123 = 0.4123$

3: $0.4123 \leq 0.8187$ & $0.25 \leq 0.4123 \leq 1$ $\boxed{\therefore \text{Accept } N=0}$

Step 1: $n=0$, $p=1$

2: $R_3 = 0.5067$ $p = 1 \times 0.5067 = 0.5067$.

3: $0.5067 \leq 0.8187$ & $0.25 \leq 0.5067 \leq 1$ $\boxed{\therefore \text{Accept } N=0}$

$R_4 = 0.6818$ $p = 1 \times 0.6818 = 0.6818$

$0.6818 \leq 0.8187$ & $0.25 \leq 0.6818 \leq 1$ $\boxed{\therefore \text{Accept } N=0}$

Step 1: $n=0$, $p=1$

2: $R_5 = 0.7293$ $p = 1 \times 0.7293 = 0.7293$

3: $0.7293 \leq 0.8187$ & $0.25 \leq 0.7293 \leq 1$ $\boxed{\therefore \text{Accept } N=0}$

Table:

n	P	R_{n+1}	$P = P \cdot R_{n+1}$	$P \leq e^{-\alpha}$ A/R	$N = n$.
0	1	0.3568	0.3568	$0.3568 \leq 0.8187$ Accept	N=0
0	1	0.4123	0.4123	$0.4123 \leq 0.8187$ Accept	N=0
0	1	0.5067	0.5067	$0.5067 \leq 0.8187$ Accept	N=0
0	1	0.6818	0.6818	$0.6818 \leq 0.8187$ Accept	N=0
0	1	0.7293	0.7293	$0.7293 \leq 0.8187$ Accept	N=0

$\therefore$ The 5 poisson variates are N=0, N=0, N=0, N=0, N=0

(15) Generate 5 gamma variate using Gamma distribution with slope, $\beta = 2.30$ & mean, $\theta = 0.4545$.
Consider the random numbers : 0.4357, 0.1806, 0.1508, 0.8353, 0.1202, 0.8004, 0.9550, 0.1460, 0.196, 0.234.

Solution :

given $\beta = 2.30$, $\theta = 0.4545$

Formula :

$$a = (2\beta - 1)^{1/2}$$

$$b = 2\beta - \ln 4 + 1/a$$

$$X = \beta\left[R_1/(1-R_1)\right]^a$$

$X \le b - \ln(R_1^2 \cdot R_2)$ then Accept & $X = X/\beta\theta$ else Reject.

Steps :

1: $a = (2 \times 2.3 - 1)^{1/2} = 1.8973$

$b = 2 \times 2.3 - \ln 4 + \frac{1}{1.8973} = 3.7407$

2: $R_1 = 0.4357$ $R_2 = 0.1806$

3: $X = 2.30\left[0.4357/(1-0.4357)\right]^{1.8973} = 1.4080$

4: $1.4080 \le 3.7407 - \ln(0.4357^2 \times 0.1806) = 7.1137$

$\boxed{\therefore \text{Accept}}$ hence $X = 1.4080/2.30 \times 0.4545 = \boxed{1.3469}$

1: $a = 1.8973$ $b = 3.7407$

2: $R_1 = 0.1508$ $R_2 = 0.8353$

3: $X = 2.30\left[0.1508/(1-0.1508)\right]^{1.8973} = 0.0866$

4: $0.0866 \le 3.7407 - \ln(0.1508^2 \times 0.8353)$

$0.0866 \le 7.7042$ $\boxed{\therefore \text{Accept}}$

$X = 0.0866/2.3 \times 0.4545 = \boxed{0.0828}$

1: $a = 1.8973$, $b = 3.7407$

2: $R_1 = 0.1202$ $R_2 = 0.8004$

3: $X = 2.3\left[0.1202/(1-0.1202)\right]^{1.8973} = 0.0526$

4: $0.0526 \leq 8.2005$ $\boxed{\therefore \text{Accept}}$

$X = 0.0526/2.3 \times 0.4545 = \boxed{0.0503}$

Step 1: $a = 1.8973$ $b = 3.7407$.

2: $R_1 = 0.9550$ $R_2 = 0.1460$

3: $X = 756.9167$

4: $756.9167 \nleq 5.7569$ $\boxed{\therefore \text{Reject}}$

Step 2: $R_1 = 0.1960$ $R_2 = 0.2340$

3: $X = 0.1580$

4: $0.1580 \leq 8.4524$ $\boxed{\therefore \text{Accept}}$

$X = 0.1580/2.3 \times 0.4545 = \boxed{0.1511}$

table:

a	b	R_1 & R_2	X	$X < b - \ln(R_1^2 R_2)$ A / R	$X = X/\beta\theta$
1.8973	3.7407	$0.4357 = R_1$ $0.1806 = R_2$	1.4080	$1.4080 \leq 7.1137$ Accept	1.3469.
1.8973	3.7407	$R_1 = 0.1508$ $R_2 = 0.8353$	0.0866	$0.0866 \leq 7.7042$ Accept	0.0828
1.8973	3.7407	$R_1 = 0.1202$ $R_2 = 0.8004$	0.0526	$0.0526 \leq 8.2005$ Accept	0.0503
1.8973	3.7407	$R_1 = 0.9550$ $R_2 = 0.1460$	756.916	$756.916 \leq 5.7569$ Reject	-
-	-	$R_1 = 0.196$ $R_2 = 0.234$	0.1580	$0.1580 \leq 8.4524$ Accept	0.1511

$\therefore$ The Gamma Variates are 1.34, 0.082, 0.0503, 0.1511

6. INPUT MODELING

- Input data provide the driving force for a simulation model. In the simulation of a queuing system, typical input data are the distributions of time between arrivals and service times.
- For the simulation of a reliability system, the distribution of time-to=failure of a component is an example of input data.

There are four steps in the development of a useful model of input data:

- Collect data from the real system of interest. This often requires a substantial time and resource commitment. Unfortunately, in some situations it is not possible to collect data
- Identify a probability distribution to represent the input process. When data are available, this step typically begins by developing a frequency distribution, or histogram, of the data.
- Choose parameters that determine a specific instance of the distribution family. When data are available, these parameters may be estimated from the data.
- Evaluate the chosen distribution and the associated parameters for good-of- fit. Goodness-of-fit may be evaluated informally via graphical methods, or formally via statistical tests. The chisquare and the Kolmo-gorov-Smirnov tests are standard goodness-of-fit tests. If not satisfied that the chosen distribution is a good approximation of the data, then the analyst returns to the second step, chooses a different family of distributions, and repeats the procedure. If several iterations of this procedure fail to yield a fit between an assumed distributional form and the collected data

6.1 Data Collection

- Data collection is one of the biggest tasks in solving real problem. It is one of the most important and difficult problems in simulation. And even if when data are available, they have rarely been recorded in a form that is directly useful for simulation input modeling.

The following suggestions may enhance and facilitate data collection, although they are not all – inclusive.

1. A useful expenditure of time is in planning. This could begin by a practice or pre observing session. Try to collect data while pre-observing.
2. Try to analyze the data as they are being collected. Determine if any data being collected are useless to the simulation. There is no need to collect superfluous data.
3. Try to combine homogeneous data sets. Check data for homogeneity in successive time periods and during the same time period on successive days.
4. Be aware of the possibility of data censoring, in which a quantity of interest is not observed in its entirety. This problem most often occurs when the analyst is interested in the time required to complete some process (for example, produce a part, treat a patient, or have a component fail), but the process begins prior to, or finishes after the completion of, the observation period.
5. To determine whether there is a relationship between two variables, build a scatter diagram.
6. Consider the possibility that a sequence of observations which appear to be independent may possess autocorrelation. Autocorrelation may exist in successive time periods or for successive customers.
7. Keep in mind the difference between input data and output or performance data, and be sure to collect input data. Input data typically represent the uncertain quantities that are largely beyond the control of the system and will not be altered by changes made to improve the system.

6.2 Identifying the Distribution with Data.

- In this section we discuss methods for selecting families of input distributions when data are available.

6.2.1 Histogram

- A frequency distribution or histogram is useful in identifying the shape of a distribution. A histogram is constructed as follows:

1. Divide the range of the data into intervals (intervals are usually of equal width;

however, unequal widths however, unequal width may be used if the heights of the frequencies are adjusted).

2. Label the horizontal axis to conform to the intervals selected.
3. Determine the frequency of occurrences within each interval.
4. Label the vertical axis so that the total occurrences can be plotted for each interval.
5. Plot the frequencies on the vertical axis.

- If the intervals are too wide, the histogram will be coarse, or blocky, and its shape and other details will not show well. If the intervals are too narrow, the histogram will be ragged and will not smooth the data.
- The histogram for continuous data corresponds to the probability density function of a theoretical distribution.

Example 6.2 : The number of vehicles arriving at the northwest corner of an intersection in a 5 min period between 7 A.M. and 7:05 A.M. was monitored for five workdays over a 20-week period. Table shows the resulting data. The first entry in the table indicates that there were 12:5 min periods during which zero vehicles arrived, 10 periods during which one vehicles arrived, and so on,

Table 6:1 Number of Arrivals in a 5 Minute period

Arrivals Per period	Frequency	Arrivals Per Period	Frequency
0	12	6	7
1	10	7	5
2	19	8	5
3	17	9	3
4	10	10	3
5	8	11	1

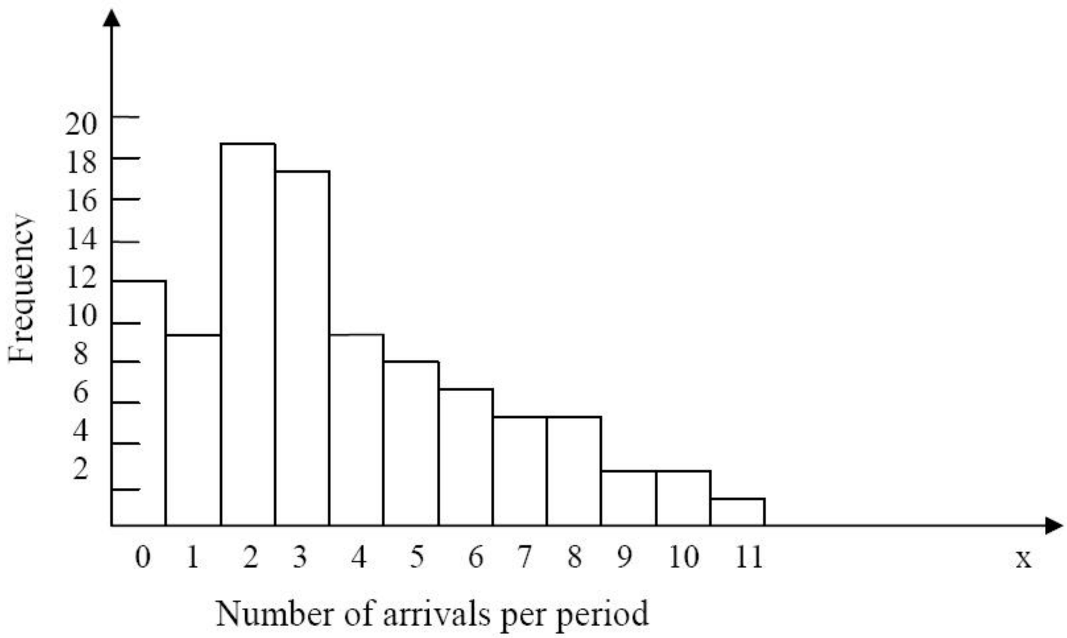


Fig 6.2 Histogram of number of arrivals per period.

6.2.2 Selecting the Family of Distributions

- Additionally, the shapes of these distributions were displayed. The purpose of preparing histogram is to infer a known pdf or pmf. A family of distributions is selected on the basis of what might arise in the context being investigated along with the shape of the histogram.

- Thus, if interarrival-time data have been collected, and the histogram has a shape similar to the pdf in Figure 5.9.the assumption of an exponential distribution would be warranted.

- Similarly, if measurements of weights of pallets of freight are being made, and the histogram appears symmetric about the mean with a shape like that shown in Fig 5.12, the assumption of a normal distribution would be warranted.

- The exponential, normal, and Poisson distributions are frequently encountered and are not difficult to analyze from a computational standpoint. Although more difficult to analyze, the gamma and Weibull distributions provide array of shapes, and should not be overlooked when modeling an underlying probabilistic process. Perhaps an exponential distribution was assumed, but it was found not to fit the data. The next step would be to examine where the lack of fit occurred.

- If the lack of fit was in one of the tails of the distribution, perhaps a gamma or Weibull distribution would more adequately fit the data.

- Literally hundreds of probability distributions have been created, many with some specific physical process in mind. One aid to selecting distributions is to use the physical basis of the distributions as a guide. Here are some examples:

6.2.3 Quantile-Quantile Plots

- Further, our perception of the fit depends on widths of the histogram intervals. But even if the intervals are well chosen, grouping of data into cells makes it difficult to compare a histogram to a continues probability density function

- If X is a random variable with cdf F, then the q-quintile of X is that y such that F(y) = P(X < y) = q, for 0 < q < 1. When F has an invererse, we write y = F-1(q).

- Now let {Xi, i = 1, 2,...,n} be a sample of data from X. Order the observations from the smallest to the largest, and denote these as {yj, j =1,2 ,,,n}, where y1 < y2 < ….. < y_n- Let j denote the ranking or order number. Therefore, j = 1 for the smallest and j = n for the largest. The q-q plot is based on the fact that y1 is an estimate of the (j — 1/2)/n quantile of X other words,

$$Yj \text{ is approximately } F^{-1} \left[\frac{J - \frac{1}{2}}{n} \right]$$

- Now suppose that we have chosen a distribution with cdf F as a possible representation of the distribution of X. If F is a member of an appropriate family of distributions, then a plot of y_j versus F^{-1}((j —1/2)/n) will be approximately a straight line.

6.3 Parameter Estimation

- After a family of distributions has been selected, the next step is to estimate the parameters of the distribution. Estimators for many useful distributions are described in this section. In addition, many software packages—some of them integrated into simulation languages—are now available to compute these estimates.

6.3.1 Preliminary Statistics: Sample Mean and Sample Variance

- In a number of instances the sample mean, or the sample mean and sample variance, are used to estimate of the parameters of hypothesized distribution;

- If the observations in a sample of size n are X1, X2,..., Xn, the sample mean (X) is defined by

$$\overline{X} = \frac{\sum^{n}_{i=1} Xi}{n} \qquad 9.1$$

and the sample variance, s^2 is defined by

$$S^2 = \frac{\sum^{n}_{i=1} Xi^2 - n\overline{X}^2}{n - 1} \qquad 9.2$$

If the data are discrete and grouped in frequency distribution, Equation (9.1) and (.2) can be modified to provide for much greater computational efficiency, The sample mean can be computed by

$$\overline{X}^2 = \frac{\sum^{n}_{j=1} fjXj}{n} \qquad 9.3$$

And the sample variance by

$$X^2 = \frac{\sum_{j=1}^{k} fjXj^2 - n\overline{X}^2}{n - 1} \qquad 9\,4$$

where k is the number of distinct values of X and fj is the observed frequency of the value Xj, of X.

6.3.2 Suggested Estimators

- Numerical estimates of the distribution parameters are needed to reduce the family of distributions to a specific distribution and to test the resulting hypothesis.
- These estimators are the maximum-likelihood estimators based on the raw data. (If the data are in class intervals, these estimators must be modified.)
- The triangular distribution is usually employed when no data are available, with the parameters obtained from educated guesses for the minimum, most likely, and maximum possible value's; the uniform distribution may also be used in this way if only minimum and maximum values are available.

Distribution	Parameter	Estimator
Poisson	α	$\hat{\alpha} = \overline{X}$
Exponential	λ	$\hat{\lambda} = \frac{1}{\overline{X}}$
Gamma	β, θ	$\hat{\theta} = \frac{1}{\overline{X}}$
Normal	μ, σ^2	$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = S^2$
Lognormal	μ, σ^2	$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = S^2$

6.4 Goodness-of-Fit Tests

- These two tests are applied in this section to hypotheses about distributional forms of input data. Goodness-of-fit tests provide help full guidance for evaluating the suitability of a potential input model.
- However, since there is no single correct distribution in a real application, you should not be a slave to the verdict of such tests.
- It is especially important to understand the effect of sample size. If very little data are available, then a goodness-of-fit test is unlikely to reject any candidate distribution; but if a lot of data are available, then a goodness-of-fit test will likely reject all candidate distribution.

6.4.1 Chi-Square Test

- One procedure for testing the hypothesis that a random sample of size n of the random variable X follows a specific distributional form is the chi-square goodness-offit test.
- This test formalizes the intuitive idea of comparing the histogram of the data to the shape of the candidate density or mass function, The test is valid for large sample sizes, for both discrete and continuous distribution assumptions, When parameters are estimated by maximum likelihood.

$$X_0^2 = \sum_{I=1}^{k} \frac{(Oi - Ei)^2}{Ei} \qquad 9.16$$

- where 0, is the observed frequency in the ith class interval and Ei, is the expected frequency in that class interval. The expected frequency for each class interval is computed as Ei=npi, where pf is the theoretical, hypothesized probability associated with the ith class interval.
- It can be shown thatX02 approximately follows the chi-square distribution with k-s-1 degrees of freedom, where s represents the number of parameters of the hypothesized distribution estimated by sample statistics. The hypotheses are :

H0: the random variable, X, conforms to the distributional assumption with the parameter(s) given by the parameter estimate(s)

H1 : the random variable X does not conform

- If the distribution being tested is discrete, each value of the random variable should be a class interval, unless it is necessary to combine adjacent class intervals to meet the minimum expected cell-frequency requirement. For the discrete case, if combining adjacent cells is not required,

$Pi = P(X_I) = P(X\ X_i)$

Otherwise, pi, is determined by summing the probabilities of appropriate adjacent cells.

- If the distribution being tested is continuous, the class intervals are given by $[a_{i-1}, ai)$, , where ai-1 and ai, are the endpoints of the ith class interval. For the continuous case with assumed pdf f(x), or assumed cdf F(x), pi, can be computed By

$Pi = {}_{ai-1}^{\ ai}\ f(x)\ dx = F(a_i) - F(a_{i\ -1})$

6.4.2 Chi-Square Test with Equal Probabilities

- If a continuous distributional assumption is being tested, class intervals that are equal in probability rather than equal in width of interval should be used.
- Unfortunately, there is as yet no method for deter mining the; probability associated with each interval that maximize the; power of a test o f a given size.

Ei = n p i 5

- Substituting for p i yields n/k 5

- and solving for k yields k n/5

6.4.3 Kolmogorov - Smirnov Goodness-of-Fit Test

- The chi-square goodness-of-fit test can accommodate the estimation of parameters from the data with a resultant decrease in the degrees of freedom (one for J each parameter estimated). The chi-square test requires that the data be placed in class intervals, and in the case of continues distributional assumption, this grouping is arbitrary.

- Also, the distribution of the chi-square test statistic is known only approximately, and the power of the test is sometimes rather low. As a result of these considerations, goodness-of-fit tests, other than the chi-square, are desired.

- The Kolmogorov-Smirnov test is particularly useful when sample sizes are small and when no parameters have been estimated from the data.

- (Kolmogoro-Smirnov Test for Exponential Distribution)

Ho : the interarrival times are exponentially distributed
H1: the interarrival times are not exponentially distributed

- The data were collected over the interval 0 to T = 100 min. It can be shown that if the underlying distribution of interarrival times { T1, T2, … } is exponential, the arrival times are uniformly distributed on the interval (0,T).

- The arrival times T1, T1+T2, T1+T2+T3,…..,T1+…..+T50 are obtained by adding interarrival times.
- On a (0,1) interval, the points will be [T1/T, (T1+T2)/T,…..,(T1+….+T50)/T].

6.5 Selecting Input Models without Data

Unfortunately. it is often necessary in practice to develop a simulation model for demonstration purposes or a preliminary study—before any i data are available.) In this case the modeler must be resourceful in choosing input models and must carefully check the sensitivity of results to the models.

Engineering data : Often a product or process has performance ratings pro vided by the manufacturer.

Expert option : Talk to people who are experienced with the procesws or similar processes. Often they can provide optimistic, pessimistic and most likely times.

Physical or conventional limitations : Most real processes have physical limit on performance. Because of company policies, there may be upper limits on how long a process may take. Do not ignore obvious limits or bound: that narrow the range of the input process.

The nature of the process It can be used to justify a particular choice even when no data are available.

6.6 Multivariate and Time-Series Input Models

The random variables presented were considered to be independent of any other variables within the context of the problem. However, variables may be related, and if the variables appear in a simulation model as inputs, the relationship should be determined and taken into consideration.

Step 1. Generate Z_1 and Z_2, independent standard normal random variables.

Step 2. Set $X_1 = \mu_1 + \sigma_1 Z_1$

Step 3. Set $X_2 = \mu_2 + \sigma_2 \left(\rho Z_1 + \sqrt{1-\rho^2} Z_2\right)$

6.7 Time series input model:

If X1,X2..Xn is a sequence of identically distributed,but dependent and convarianc stationary random variables,then there are a number of times series model that can be used to represent the process. The two models that have the characteristics that the autocorrelatrion take the form.

$$\rho_h = \text{corr}(X_t, X_{t+h}) = \rho^h$$

for h=1,2,..n that the log-h autocorrelation decreases geometrically as the lag increases.

AR(1) Model:

consider the time series model

$$X_t = \mu + \phi(X_{t-1} - \mu) + \varepsilon_t$$

for t=2,3,..n where ε2, ε3 are the independent and identically distributed with men 0 and variance σ^2_ε and -1< ϕ<1. If the initial value x1 is chosen appropriately,then x1,x2..are all normal distributed with mean u and variance $\sigma_\varepsilon^2/(1-\phi^2)$,

Step 1. Generate X_1 from the normal distribution with mean μ and variance $\sigma_\varepsilon^2/(1-\phi^2)$. Set $t = 2$.

Step 2. Generate ε_t from the normal distribution with mean 0 and variance σ_ε^2.

Step 3. Set $X_t = \mu + \phi(X_{t-1} - \mu) + \varepsilon_t$.

Step 4. Set $t = t + 1$ and go to Step 2.

EAR(1) Model:

Consider the time series model

$$X_t = \begin{cases} \phi X_{t-1}, & \text{with probability } \phi \\ \phi X_{t-1} + \varepsilon_t, & \text{with probability } 1-\phi \end{cases}$$

for t=2,3,..n where ε2, ε3 are the independent and identically distributed with mean $1/\lambda$ and 0< ϕ<1. If the initial value x1 is chosen appropriately, then x1,x2.. are all exponentially distributed with mean $1/\lambda$ and variance $\sigma_\varepsilon^2/(1-\phi^2)$,

Step 1. Generate X_1 from the exponential distribution with mean $1/\lambda$. Set $t = 2$.

Step 2. Generate U from the uniform distribution on [0, 1]. If $U \le \phi$, then set

$$X_t = \phi X_{t-1}$$

Otherwise, generate ε_t from the exponential distribution with mean $1/\lambda$ and set

$$X_t = \phi X_{t-1} + \varepsilon_t$$

Step 3. Set $t = t + 1$ and go to Step 2.

Goodness-of-Fit Tests

⊙ Chi-Square Test with poisson Assumption

Step 1: Compute poisson distribution using

$$P(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^{x}}{x!} & x = 0, 1, 2, \ldots \end{cases}$$

Step 2: Compute expected frequency

$$E_i = n.P(i)$$ where n is sum of sample data

Reduce interval i.e

$$E_i > 5$$

Step 3: Compute chi-square test i.e

$$X_0^2 = \sum_{i=0}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Step 4: Obtain chi-square test value from table A.6

$$X^2_{\alpha, K-S-1}$$

Step 5: Check hypothesis or Null hypothesis

$$X_0^2 \lessgtr X^2_{\alpha, K-S-1}$$ Accepted hypothesis / Rejected Null hypothesis

*) Chi-Square test for Exponential distribution (Equal probability)

Step 1: Determine the probability

$$\boxed{P = 1/K}$$ where K is interval

Step 2: Determine the mean

$$\lambda = \frac{1}{\bar{X}}$$

$$\bar{X} = \sum_{i=0}^{n} \frac{X_i}{n}$$

Step 3: Compute class interval

$$\boxed{a_i = -\frac{1}{\lambda} \ln(1 - iP)}$$ $i = 0, 1, 2, \ldots K$

Step 4: Compute expected frequency

$$\boxed{E_i = \frac{N}{K}}$$

N - Sum of Sample data

K - interval

Step 5: Compute chi-Square test

$$\boxed{X_0^2 = \sum_{i=0}^{n} \frac{(O_i - E_i)^2}{E_i}}$$

Step 6: Obtain chi-square test Value from table A.6

$$\boxed{X^2_{\alpha, K-s-1}}$$

Step 7: Check hypothesis or Null hypothesis

$$\boxed{X_0^2 < X^2_{\alpha, K-s-1}}$$ accepted

*) Kolmogorov - Smirnov test for exponential distributions

Step 1: Calculate inter arrival points

$$\boxed{R_i = \{ T_1/T,\ (T_1+T_2)/T,\ (T_1+T_2+T_3)/T,\ (T_1+T_2+T_n)/T \}}$$

T is total No's of Sample data

T_i - is the Sample data

Step 2: Compute

$$\boxed{\begin{aligned} D^+ &= \max_{1\le i\le n} \{ i/n - R(i) \} \\ D^- &= \max_{1\le i\le n} \{ R(i) - \tfrac{i-1}{n} \} \end{aligned}}$$

Step 3: Compute

$$\boxed{D = \max (D^+, D^-)}$$

Step 4: Obtain KS-test Value from table A.8

$$\boxed{D_{\alpha, n}}$$

Step 5: Check hypothesis or Null hypothesis

$\boxed{D \le D_{\alpha, n}}$ accepted

(1) UNIT-6 : INPUT MODELLING

I. Chi Square Test using Poisson Assumption

1. Using goodness of fit test, test whether random Nos are uniformly distributed based on poisson assumption with level of significance $\alpha = 0.05$. $\hat{\alpha} = 3.64$. Sample data are :

Interval :	0	1	2	3	4	5	6	7	8	9	10	11
observed Frequency :	12	10	19	17	10	8	7	5	5	3	3	1

⇒ Given : $\alpha = 0.05$

$\hat{\alpha} = 3.64$

$n = 12 + 10 + 19 + 17 + 10 + 8 + 7 + 5 + 5 + 3 + 3 + 1 = 100$

Step 1 : Compute Poisson Distribution

$$P(x) = \frac{e^{-\alpha}\alpha^{x}}{x!} \quad \text{where } x = 0, 1, \ldots, 11$$

$$P(0) = \frac{e^{-3.64} * (3.64)^{0}}{0!} = 0.026$$

$P(1) = 0.096$

$P(2) = 0.174$

$P(3) = 0.211$

$P(4) = 0.192$

$P(5) = 0.140$

$P(6) = 0.085$

$P(7) = 0.044$

$P(8) = 0.020$

$P(9) = 0.008$

$P(10) = 0.003$

$P(11) = 0.001$

Step 2 : Apply Chi Square with poisson assumption.

X_i	O_i	$E_i = n \cdot P_i$	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$
$\begin{matrix}0\\1\end{matrix}$	$\left.\begin{matrix}12\\10\end{matrix}\right\} 22$	$\left.\begin{matrix}2.6\\9.6\end{matrix}\right\} 12.2$	9.8	96.04	7.87
2	19	17.4	1.6	2.56	0.15
3	17	21.1	−4.1	16.81	0.79
4	10	19.2	−9.2	84.64	4.41
5	8	14.0	−6	36	2.57
6	7	8.5	−1.5	2.25	0.27
$\begin{matrix}7\\8\\9\\10\\11\end{matrix}$	$\left.\begin{matrix}5\\5\\3\\3\\1\end{matrix}\right\} 17$	$\left.\begin{matrix}4.4\\2.0\\0.8\\0.3\\0.1\end{matrix}\right\} 7.6$	9.4	88.36	11.63

We have $k = 7$, $s = 1$ $\qquad X_0^2 = \underline{27.69}$

Step 3 : Compute level of Significance from Table A6

$$X^2_{0,\alpha,k-s-1} = X^2_{0\ 0.05,\ 7-1-1}$$

$$= X^2_{0\ 0.05,\ 5} = \underline{11.1}$$

Step 4 : Check whether Random No's are uniformly distributed.

Compare X_0^2 & $X^2_{0\ 0.05,\ 5}$

$\because$ $27.69 > 11.1$ $\Rightarrow$ Random No's are not uniformly distributed.

2. Using goodness of fit test, check whether Random No.s are uniformly distributed over interval [0, 1] using poisson assumption with level of significance = 0.05. Simulation table for critical values is given:

Interval (X_i) :	0	1	2	3	4	5	6	7
Frequency (f_i) :	5	10	5	8	12	10	8	12

⇒ Given : $\alpha = 0.05$

$n = 5+10+5+8+12+10+8+12 = 70$

$\hat{\alpha} = ?$

$$\hat{\alpha} = \bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{n} = \frac{0+10+10+24+48+50+48+84}{70}$$

$$\hat{\alpha} = \frac{274}{70} = 3.91$$

Step 1 : Compute Poisson Distribution

$$P(x) = \frac{e^{-\alpha}\alpha^{x}}{x!}, \quad x = 0, 1, \ldots, 7. \quad \& \quad \alpha = 3.91$$

$P(0) = 0.020$

$P(1) = 0.078$

$P(2) = 0.153$

$P(3) = 0.199$

$P(4) = 0.195$

$P(5) = 0.153$

$P(6) = 0.099$

$P(7) = 0.056$

Step 2 : Apply Chi Square test with poisson assumption

X_i	O_i	$E_i = n \cdot P_i$	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$
0	$\left.\begin{matrix}5\\10\end{matrix}\right\}15$	$\left.\begin{matrix}1.4\\5.46\end{matrix}\right\}6.86$	8.14	66.26	9.66
1					
2	5	10.71	−5.71	32.60	3.04
3	8	13.93	−5.93	35.16	2.52
4	12	13.65	−1.65	2.72	0.19
5	10	10.71	−0.71	0.50	0.05
6	$\left.\begin{matrix}8\\12\end{matrix}\right\}20$	$\left.\begin{matrix}6.93\\3.92\end{matrix}\right\}10.85$	9.15	83.72	7.72
7					

$X_0^2 = 23.18$

Here $k = 6$, $s = 1$

Step 3 : Compute level of Significance from Table A6

$X_{0\ \alpha,\ k-s-1}^2 = X_{0\ 0.05,\ 6-1-1}^2 = 9.49$

Step 4 : Check whether Random No.s are uniformly distributed.

Compare X_0^2 & $X_{0\ 0.05,\ 4}^2$

$\because 23.18 > 9.49 \Rightarrow$ Random No.s are not uniformly distributed

3. Apply goodness of fittest, check whether Random Nos (3) are uniformly distributed over Interval [0, 1] with given size of data 100. Assume $\alpha = 0.01$. Simulation table to check critical values using poisson assumption is given below :

Interval :	1	2	3	4	5	6	7	8	9	10
Frequency :	8	6	10	11	12	8	10	12	12	11

$\Rightarrow$ Given : $\alpha = 0.01$ $\hat{\alpha} = ?$

$n = 100$

$$\alpha = \bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{n} = \frac{586}{100} = 5.86$$

Step 1 : Compute Poisson Distribution

$$P(x) = \frac{e^{-\alpha} \cdot \alpha^{x}}{x!} \quad \text{where } x = 1, 2 \ldots 10 \quad \& \quad \alpha = 5.86$$

$P(1) = 0.017$

$P(2) = 0.049$

$P(3) = 0.096$

$P(4) = 0.140$

$P(5) = 0.164$

$P(6) = 0.160$

$P(7) = 0.134$

$P(8) = 0.098$

$P(9) = 0.064$

$P(10) = 0.038$

Step 2 : Apply Chi Square with poisson assumption

X_i	O_i	$E_i = n \cdot P_i$	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum_{i=1} \frac{(O_i - E_i)^2}{E_i}$
1	8 } 14	1.7 } 6.6	7.4	54.76	8.29
2	6	4.9			
3	10	9.6	0.4	0.16	0.02
4	11	14.0	−3	9	0.64
5	12	16.4	−4.4	19.36	1.18
6	8	16.0	−8	64	4
7	10	13.4	−3.4	11.56	0.86
8	12	9.8	2.2	4.84	0.49
9	12 } 23	6.4 } 10.2	12.8	163.84	16.06
10	11	3.8			

$k = 8, \ s = 1$ $\qquad$ $X_0^2 = 31.54$

Step 3 : Compute level of Significance from Table A6

$X_{0\,\alpha}^2 , k-s-1 = X_0^2 \ 0.01, \ 8-1-1 = 20.1$

Step 4 : Check whether Random No.s are uniformly distributed

Compare X_0^2 & X_0^2 0.05, 6

$31.54 > 20.1 \Rightarrow$ Random No.s are not uniformly distributed

(4)

II. Chi Square Test with Equal Probability (Exponential Dist.)

1. Apply goodness of fit test to check whether random No.s are uniformly distributed over [0, 1] using equal probability. Use $\alpha = 0.05$, interval $k = 8$ to check whether given sample datas are accepted or rejected.

79.918	3.081	0.062	1.961	5.845	3.027	6.505
0.021	0.013	0.123	6.769	59.899	1.192	34.760
5.009	18.387	0.141	43.565	24.420	0.433	144.695
2.663	17.967	0.091	9.003	0.941	0.878	3.371
2.157	7.579	0.624	5.380	3.148	7.078	23.960
0.590	1.928	0.300	0.002	0.543	7.004	31.764
1.005	1.147	0.219	3.217	14.382	1.008	2.336
4.562						

⇒ Given : $k = 8$, $\alpha = 0.05$

Step 1 : Compute mean

$$\bar{\lambda} = \frac{1}{\bar{X}} \quad \text{where} \quad \bar{X} = \frac{\sum X_i}{n}$$

$$\bar{X} = \frac{594.674}{50} = \underline{\underline{11.894}}$$

$$\bar{\lambda} = \frac{1}{11.894}$$

$$\boxed{\bar{\lambda} = 0.084}$$

Step 2 : Compute class intervals

$$P = \frac{1}{k} = \frac{1}{8} = \underline{\underline{0.125}}$$

$$a_i = \frac{-1}{\lambda} \ln[1 - i * p] \quad \text{where } i = 0 \ldots 8$$

$\lambda = 0.084$

$P = 0.125$

$a_0 = 0$

$a_1 = 1.589$ $\qquad$ $a_5 = 11.677$

$a_2 = 3.425$ $\qquad$ $a_6 = 16.504$

$a_3 = 5.595$ $\qquad$ $a_7 = 24.755$

$a_4 = 8.252$ $\qquad$ $a_8 = \infty$

Step 3 : Compute Chi square with equal probability

Class Interval	O_i	$E_i = \frac{N}{k}$ (50/8)	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum_{i=1} \frac{(O_i - E_i)^2}{E_i}$
0 – 1.589	19	6.25	12.75	162.563	26.01
1.589 – 3.425	10	6.25	3.75	14.063	2.25
3.425 – 5.595	3	6.25	−3.25	10.563	1.69
5.595 – 8.252	6	6.25	−0.25	0.0625	0.01
8.252 – 11.677	1	6.25	−5.25	27.563	4.41
11.677 – 16.504	1	6.25	−5.25	27.563	4.41
16.504 – 24.755	4	6.25	−2.25	5.0623	0.81
24.755 – ∞	6	6.25	−0.25	0.063	0.01
	50				$X_0^2 = 39.60$

Step 3 : Compute level of Significance from table A6.

$X_0^2\ \alpha,\ k-s-1 = X_0^2\ 0.05,\ 8-1-1 = 12.6$

Step 4 : Check whether random No.s are uniformly distributed.

$39.60 > 12.6 \Rightarrow$ Random No.s are rejected.

3. Consider goodness of fit test using Chi Square test with equal probability. Given $k = 6$, $\alpha = 0.05$. Sample data:

(5)

0.34	0.90	1.88	1.90	0.74	2.62	2.67	3.53	4.91
5.50	1.10	1.03	1.73	1.00	2.03	1.49	2.16	0.80
0.48	5.60	0.45	0.26	0.24	0.63	0.36	1.28	0.82
2.16	0.05	0.04	0.39	0.21	0.79	0.53	3.53	2.62
0.53	1.50	2.81						

Given: $k = 6$ $\alpha = 0.05$ $N = 39$

Step 1: Compute Mean

$$\lambda = \frac{1}{\bar{X}} \quad \text{where} \quad \bar{X} = \frac{\sum X_i}{N} = \frac{61.61}{39} = 1.579$$

$$\lambda = \frac{1}{1.579} = \underline{\underline{0.63}}$$

Step 2: Compute class intervals

$$P = \frac{1}{k} = \frac{1}{6} = \underline{\underline{0.17}}$$

$$a_i = \frac{-1}{\lambda} \ln\left[1 - i * p\right] \quad \text{where} \quad i = 0, 1 \ldots 6$$

$\lambda = 0.63$

$p = 0.17$

$a_0 = 0$

$a_1 = 0.29$

$a_2 = 0.66$

$a_3 = 1.13$

$a_4 = 1.81$

$a_5 = 3.01$

$a_6 = \infty$

Class Interval	O_i	$E_i = \frac{N}{K}$	$O_i - E_i$	$(O_i - E_i)^2$	$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$
0 - 0.29	5	6.5	-1.5	2.25	0.35
0.29 - 0.66	8	6.5	1.5	2.25	0.35
0.66 - 1.13	8	6.5	1.5	2.25	0.35
1.13 - 1.81	4	6.5	-2.5	6.25	0.96
1.81 - 3.01	9	6.5	2.5	6.25	0.96
3.01 - ∞	5	6.5	-1.5	2.25	0.35

$X_0^2 = 3.32$

Step 3 : Compute level of Significance from table A6

$X_0^2{}_{\alpha,\, k-s-1} = X_0^2{}_{0.05,\, 6-1-1} = 9.49$

Step 4 : Check whether Random No.s are uniformly distributed.

$\because 3.32 < 9.49 \Rightarrow$ Random No.s are accepted

III K-S Test

1. Apply goodness of fit test to check whether Random No.s are uniformly distributed over $(0, T)$ for an interval 100. Take $\alpha = 0.05$

Simulation table for Critical values :

0.44 0.53 2.04 2.74 2.00 0.30 2.54 0.52 2.02 1.89

Given : $\alpha = 0.05$, $n = 10$, $T = 100$

$$R_{(i)} = \left\{ \frac{T_1}{T}, \frac{T_1 + T_2}{T}, \ldots\ldots, \frac{T_1 + T_2 + \ldots T_n}{T} \right\}$$

⑥ Step 1 :

$R_{(i)} = \{0.0044, 0.0097, 0.0301, 0.0575, 0.0775, 0.0805, 0.1059, 0.1111, 0.1313, 0.1502\}$

Step 2 :

i	$R_{(i)}$	i/n	$i-1/n$	$D^+ = \max\{\frac{i}{n} - R_{(i)}\}$	$D^- = \max\{R_{(i)} - \frac{i-1}{n}\}$
1	0.0044	0.1	0	0.0956	0.0044
2	0.0097	0.2	0.1	0.1903	~
3	0.0301	0.3	0.2	0.2699	~
4	0.0575	0.4	0.3	0.3425	~
5	0.0775	0.5	0.4	0.4225	~
6	0.0805	0.6	0.5	0.5195	~
7	0.1059	0.7	0.6	0.5941	~
8	0.1111	0.8	0.7	0.6889	~
9	0.1313	0.9	0.8	0.7687	~
10	0.1502	1.0	0.9	0.8498	~

Step 3 :

$D = \max\{D^+, D^-\} = \max\{0.8498, 0.0044\}$

$D = \underline{0.8498}$

Step 4 :

D_{α}, n from A8 table.

$D_{0.05}, 10 = \underline{0.410}$

Step 5 :

$\therefore 0.8498 > 0.410 \Rightarrow$ Random Nos are rejected

2. Consider Sample data. Perform KS Test.

0.10 1.42 0.46 0.07 1.09 0.76 5.53 3.93 1.07

2.26 2.88 0.67 1.12 0.26

Interval (0-T) = 100 min $\alpha = 0.05$ $n = 14$

= Step 1 :

$R_{(i)} = \{0.0010, 0.0152, 0.0198, 0.0205, 0.0314, 0.039, 0.0943, 0.1336, 0.1443, 0.1669, 0.1957, 0.2024, 0.2136, 0.2162\}$

Step 2 :

i	$R_{(i)}$	i/n	i-1/n	$D^+ = \max\{\frac{i}{n} - R_{(i)}\}$	$D^- = \max\{R_{(i)} - \frac{i-1}{n}\}$
1	0.0010	0.0714	0	0.0704	0.0010
2	0.0152	0.1429	0.0714	0.1277	~
3	0.0198	0.2143	0.1429	0.1945	~
4	0.0205	0.2857	0.2143	0.2652	~
5	0.0314	0.3571	0.2857	0.3257	~
6	0.039	0.4286	0.3571	0.3896	~
7	0.0943	0.5	0.4286	0.4057	~
8	0.1336	0.5714	0.5	0.4378	~
9	0.1443	0.6429	0.5714	0.4986	~
10	0.1669	0.7143	0.6429	0.5474	~
11	0.1957	0.7857	0.7143	0.59	~
12	0.2024	0.8571	0.7857	0.6547	~
13	0.2136	0.9286	0.8571	0.715	~
14	0.2162	1	0.9286	0.7838	~

Step 3 : $D = \max\{0.7838, 0.0010\} = \underline{0.7838}$

Step 4 : $D_{0.05, 14} = 0.349$ (A8 table)

Step 5 : $\because 0.7838 > 0.349 \Rightarrow$ Random No.s are rejected

OUTPUT ANALYSIS FOR A SINGLE MODEL

Estimate system performance via simulation

- If *q* is the system performance, the precision of the estimator can be measured by:
    1. The standard error of .
    2. The width of a confidence interval (CI) for *q*.
- Purpose of statistical analysis:
    1. To estimate the standard error or CI .
    2. To figure out the number of observations required to achieve desired error/CI.
- Potential issues to overcome:
    1. Autocorrelation, e.g. inventory cost for subsequent weeks lack statistical independence.
    2. Initial conditions, e.g. inventory on hand and # of backorders at time 0 would most likely influence the performance of week 1.

7.1 Type of Simulations

- Terminating verses non-terminating simulations
- Terminating simulation:
    1. Runs for some duration of time T_E, where E is a specified event that stops the simulation.
    2. Starts at time *0* under well-specified initial conditions.
    3. Ends at the stopping time T_E.
    4. Bank example: Opens at 8:30 am (time *0*) with no customers present and *8* of the *11* teller working (initial conditions), and closes at 4:30 pm (Time $T_E = 480$ minutes).
    5. The simulation analyst chooses to consider it a terminating system because the object of interest is one day‘s operation.

7.2 Stochastic Nature of Output Data

- Model output consist of one or more random variables (r. v.) because the model is an input-output transformation and the input variables are r.v.‘s.

- M/G/1 queuing example:

    1. Poisson arrival rate = *0.1* per minute;
       service time ~ $N(m = 9.5, s = 1.75)$.

    2. System performance: long-run mean queue length, $L_Q(t)$.

    3. Suppose we run a single simulation for a total of 5,000 minutes

- Divide the time interval [*0, 5000*) into *5* equal subintervals of *1000* minutes.

Average number of customers in queue from time *(j-1)1000* to *j(1000)* is Y_j .

- M/G/1 queueing example (cont.):

- Batched average queue length for 3 independent replications:

Batching Interval (minutes)	Batch, j	Replication 1, Y_{1j}	2, Y_{2j}	3, Y_{3j}
[0, 1000)	1	3.61	2.91	7.67
[1000, 2000)	2	3.21	9.00	19.53
[2000, 3000)	3	2.18	16.15	20.36
[3000, 4000)	4	6.92	24.53	8.11
[4000, 5000)	5	2.82	25.19	12.62
[0, 5000)		3.75	15.56	13.66

- Inherent variability in stochastic simulation both within a single replication and across different replications.
- The average across 3 replications, can be regarded as independent observations, but averages within a replication, $Y_{11}, \ldots, Y_{15}$, are not.

7.3 Measures of performance

- Consider the estimation of a performance parameter, *q* (or *f*), of a simulated system.

  1. Discrete time data: [$Y_1, Y_2, \ldots, Y_n$], with ordinary mean: *q*

  2. Continuous-time data: {$Y(t), 0 \le t \le T_E$} with time-weighted mean: *f*

7.3.1 Point Estimator

- Point estimation for discrete time data[$Y_1, Y_2, \ldots, Y_n$] is defined by.

  The point estimator:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Where $\hat{\theta}$ is a sample mean based on sample of size n The pointer estimator $\hat{\theta}$ is said to be unbiased for θ if its expected value is θ, that is if: Is biased

$$E(\hat{\theta}) = \theta$$

- Point estimation for continuous-time data. The point estimator:

$$\hat{\phi} = \frac{1}{T_E}\int_0^{T_E} Y(t)dt$$

- An unbiased or low-bias estimator is desired.

- Usually, system performance measures can be put into the common framework of *q* or *f:*

  the proportion of days on which sales are lost through an out-of-stock situation, let:

$$Y(t) = \begin{cases} 1, & \text{if out of stock on day } i \\ 0, & \text{otherwise} \end{cases}$$

- Performance measure that does not fit: quantile or percentile:
- Estimating quantiles: the inverse of the problem of estimating a proportion or probability. $\Pr\{Y \le \theta\} = p$
- Consider a histogram of the observed values *Y*:
- Find such that *100p*% of the histogram is to the left of (smaller than) .

7.3.2 Confidence-Interval Estimation

To understand confidence intervals fully, it is important to distinguish between measures of error, and measures of risk, e.g., confidence interval versus prediction interval.

Suppose the model is the normal distribution with mean *q*, variance s^2 (both unknown).

- Let Y_i be the average cycle time for parts produced on the i^{th} replication of the simulation (its mathematical expectation is *q*).
- Average cycle time will vary from day to day, but over the long-run the average of the averages will be close to *q*.
- Sample variance across *R* replications: $S^2 = \frac{1}{R-1}\sum_{i=1}^{R}(Y_{i.} - Y_{..})^2$

7.3.3 Confidence-Interval Estimation

- Confidence Interval (CI):
  - A measure of error.
  - Where $Y_{i.}$ are normally distributed.

$$Y_{..} \pm t_{\alpha/2,R-1}\frac{S}{\sqrt{R}}$$

- We cannot know for certain how far is from q but CI attempts to bound that error.
- A CI, such as 95%, tells us how much we can trust the interval to actually bound the error between and q .
- The more replications we make, the less error there is in (converging to 0 as R goes to infinity).

7.3.4 Confidence-Interval Estimation

- Prediction Interval (PI):
  - A measure of risk.
  - A good guess for the average cycle time on a particular day is our estimator but it is unlikely to be exactly right.
  - PI is designed to be wide enough to contain the *actual* average cycle time on any particular day with high probability.
  - Normal-theory prediction interval:

$$\bar{Y}_{..} \pm t_{\alpha/2,R-1} S \sqrt{1+\frac{1}{R}}$$

  - The length of PI will not go to 0 as R increases because we can never simulate away risk.
  - PI's limit is: $\theta \pm z_{\alpha/2}\sigma$

UNIT 8: Verification and validation modeling

- One of the most important and difficult tasks facing a model developer is the Verification and validation of the simulation model.

- It is the job of the model developer to work closely with the end users Throughout the period (development and validation to reduce this skepticism And to increase the credibility.

<u>The goal of the validation process is twofold:</u>

1: To produce a model that represents true system behavior closely enough for the model to be used as a substitute for the actual system for the purpose of experimenting with system.

2: To increase an acceptable, level the credibility of the model ,so that the model will be used by managers and other decision makers. |

<u>The verification and validation process consists of the following components:-</u>

1:Verification is concerned with building the model right. It is utilized in comparison of the conceptual model to the computer representation that implements that conception. It asks the questions: Is the model implemented correctly in the computer? Are the input parameters and logical structure of the model correctly represented?

2: Validation is concerned with building the right model. It is utilized to determine that a model is an accurate representation of the real system. It is usually achieved through the calibration of the model

7.1 Model Building, Verification, and Validation

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

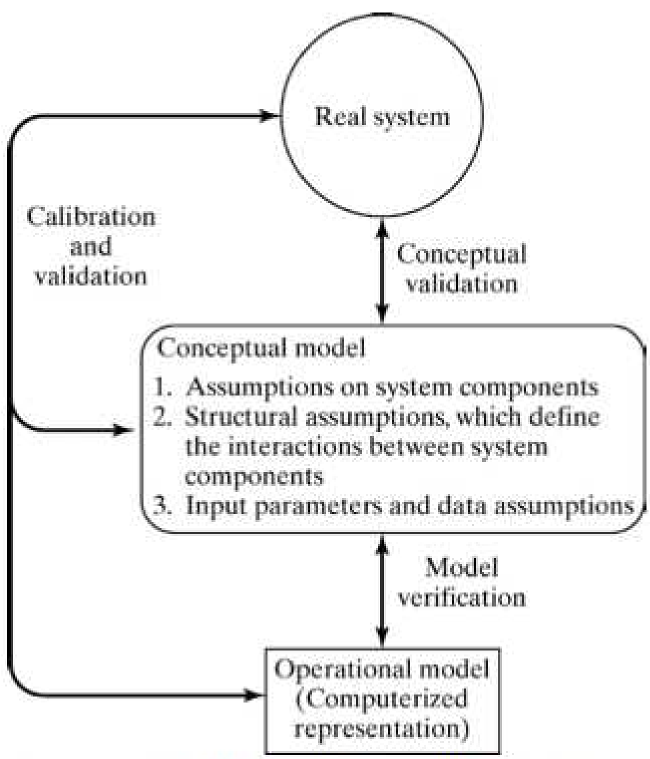


Figure 1 Model building, verification, and validation

7.2 Verification of Simulation Models

- The purpose of model verification is to assure that the conceptual model is reflected accurately in the computerized representation.
- The conceptual model quite often involves some degree of abstraction about system operations, or some amount of simplification of actual operations.

Many common-sense suggestions can be given for use in the verification process:-

- Have the computerized representation checked by someone other than its developer.
- Make a flow diagram which includes each logically possible action a system can take when an event occurs, and follow the model logic for each a for each action for each event type.
- Closely examine the model output for reasonableness under a variety of settings of Input parameters.
- Have the computerized representation print the input parameters at the end of the Simulation to be sure that these parameter values have not been changed inadvertently.
- Make the computerized representation of self-documenting as possible.
- If the computerized representation is animated, verify that what is seen in the animation imitates the actual system.
- The interactive run controller (IRC) or debugger is an essential component of Successful simulation model building. Even the best of simulation analysts makes mistakes or commits logical errors when building a model.

  The IRC assists in finding and correcting those errors in the follow ways:

  (a) The simulation can be monitored as it progresses.

  (b) Attention can be focused on a particular line of logic or multiple lines of logic that constitute a procedure or a particular entity.

  (c) Values of selected model components can be observed. When the simulation has paused, the current value or status of variables, attributes, queues, resources, counters, etc., can be observed

  (d) The simulation can be temporarily suspended, or paused, not only to view information but also to reassign values or redirect entities.

- Graphical interfaces are recommended for accomplishing verification & validation

7.3 Calibration and Validation of Models (As an aid in the validation process or Naylor finger approches):

- Verification and validation although are conceptually distinct, usually are conducted Simultaneously by the modeler.
- Validation is the overall process of comparing the model and its behavior to the real System and its behavior.
- Calibration is the iterative process of comparing the model to the real system, making adjustments to the model, comparing again and so on.
- The following figure 7.2 shows the relationship of the model calibration to the overall validation process.
- The comparison of the model to reality is carried out by variety of test Test are subjective and objective.
    - Subjective test usually involve people, who are knowledgeable about one or more aspects of the system, making judgments about the model and its output.
    - Objective tests always require data on the system's behavior plus the corresponding data produced by the model.

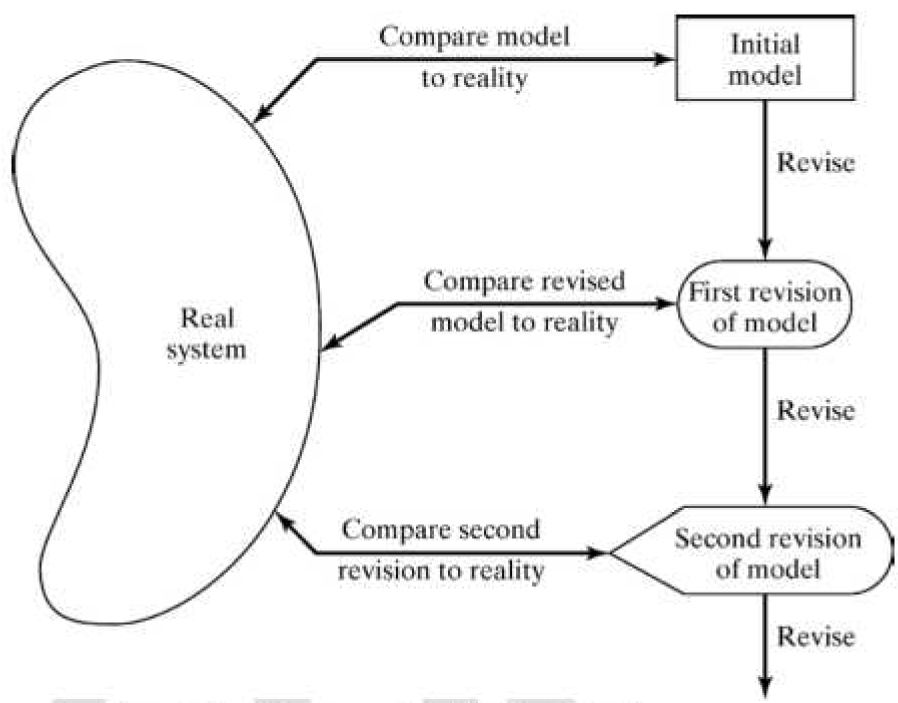


Figure 2 Iterative process of calibration a model

As an aid in the validation process, Naylor finger:

1. Build a model that has high face validity.
2. Validate model assumption.
3. Compare the model input-output transformation to cooresponding input-output transformation for the real system.

7.3.1 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

7.3.2 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.data were collected on:

    1. Inter arrival times of customers during several 2-hour periods of peak loading ("rush-hour" traffic)
    2. Inter arrival times during a slack period
    3. Service times for commercial accounts
    4. Service times for personal accounts

- Validation is not an either/or proposition—no model is ever totally representative of the system under study. In addition, each revision of the model, as in the Figure above involves some cost, time, and effort.
- The procedure for analyzing input data consist of three steps:-

  1: Identifying the appropriate probability distribution.

  2: Estimating the parameters of the hypothesized distribution .

  3: Validating the assumed statistical model by goodness – of – fit test such as the chi square test, KS test and by graphical methods

10.3.3 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.
- 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.

- Thus accurate “ prediction of the past” may replace prediction of the future for purpose of validating the future.
- A necessary condition for input-output transformation is that some version of the system under study exists so that the system data under at least one set of input condition can be collected to compare to model prediction.
- If the system is in planning stage and no system operating data can be collected, complete input-output validation is not possible.
- Validation increases modeler’s confidence that the model of existing system is accurate.
- Changes in the computerized representation of the system, ranging from relatively minor to relatively major include :

  1: Minor changes of single numerical parameters such as speed of the machine, arrival rate of the customer etc.

  2: Minor changes of the form of a statistical distribution such as distribution of service time or a time to failure of a machine.

  3: Major changes in the logical structure of a subsystem such as change in queue discipline for waiting-line model, or a change in the scheduling rule for a job shop model.

  4: Major changes involving a different design for the new system such as computerized inventory control system replacing a non computerized system .

- If the change to the computerized representation of the system is minor such as in items one or two these change can be carefully verified and output from new model can be accepted with considerable confidence.

7.3.4: Input-Output Validation: Using Historical Input Data

- When using artificially generated data as input data the modeler expects the model produce event patterns that are compatible with, but not identical to, the event patterns that occurred in the real system during the period of data collection.
- Thus, in the bank model, artificial input data {X\n, X2n, n = 1,2, , .} for inter arrival and service

times were generated and replicates of the output data Y2 were compared to what was observed in the real system

- An alternative to generating input data is to use the actual historical record, {An, Sn, n = 1,2,...}, to drive simulation model and then to compare model output to system data.
- To implement this technique for the bank model, the data Ai, A2,..., S1 S2 would have to be entered into the model into arrays, or stored on a file to be read as the need arose.
- To conduct a validation test using historical input data, it is important that all input data (An, Sn,...) and all the system response data, such as average delay(Z2), be collected during the same time period.
- Otherwise, comparison of model responses to system responses, such as the comparison of average delay in the model (Y2) to that in the system (Z2), could be misleading.
- responses (Y2 and 22) depend on the inputs (An and Sn) as well as on the structure of the system, or model.
- Implementation of this technique could be difficult for a large system because of the need for simultaneous data collection of all input variables and those response variables of primary interest.

7.3.5: Input-Output Validation: Using a Turing Test

- In addition to statistical tests, or when no statistical test is readily applicable persons knowledgeable about system behavior can be used to compare model output to system output.
- For example, suppose that five reports of system performance over five different days are prepared, and simulation output are used to produce five "fake" reports. The 10 reports should all be in exactly in the same format and should contain information of the type that manager and engineer have previously seen on the system.
- The ten reports are randomly shuffled and given to the engineers, who is asked to decide which report are fake and which are real.
- If engineer identifies substantial number of fake reports the model builder questions the engineer and uses the information gained to improve the model.
- If the engineer cannot distinguish between fake and real reports with any consistency, the modeler will conclude that this test provides no evidence of model inadequacy .

- This type of validation test is called as TURING TEST.

8.4 Optimization via simulation:

- **Optimization via simulation to refer to the problem of maximizing or minimizing the expected performance of a discrete event, stochastic system that is represented by a computer simulation model.**
- **Optimization usually deals with problems with certainty, but in stochastic discrete-event simulation the result of any simulation run is a random variable**
- **let x1,x2,..xm be the m controllable design variable and Y(x1,x2,..xm)be the observed simulation output performance on one run:**
- **To optimize Y(x1,x2,..xm) with respect to x1,x2,..xm is to maximize or minimize the mathematical expectation of performance. E[Y(x1,x2,..xm)]**

- **Optimal for deterministic counterpart.** The idea here is to use an algorithm that would find the optimal solution *if the performance of each design could be evaluated with certainty*. An example might be applying a standard nonlinear programming algorithm to the simulation optimization problem. It is typically up to the analyst to make sure that enough simulation effort is expended (replications or run length) to insure that such an algorithm is not misled by sampling variability. Direct application of an algorithm that assumes deterministic evaluation to a stochastic simulation is not recommended.
- **Robust heuristics.** Many heuristics have been developed for deterministic optimization problems that do not guarantee finding the optimal solution, but nevertheless been shown to be very effective on difficult, practical problems. Some of these heuristics use randomness as part of their search strategy, so one might argue that they are less sensitive to sampling variability than other types of algorithms. Nevertheless, it is still important to make sure that enough simulation effort is expended (replications or run length) to insure that such an algorithm is not misled by sampling variability.

- **Guarantee a prespecified probability of correct selection.** The Two-Stage Bonferroni Procedure in Section 12.2.2 is an example of this approach, which allows the analyst to specify the desired chance of being right. Such algorithms typically require either that every possible design be simulated or that a strong functional relationship among the designs (such as a metamodel) apply. Other algorithms can be found in Goldsman and Nelson [1998].
- **Guarantee asymptotic convergence.** There are many algorithms that guarantee convergence to the global optimal solution as the simulation effort (number of replications, length of replications) becomes infinite. These guarantees are useful because they indicate that the algorithm tends to get to where the analyst wants it to go. However, convergence can be slow, and there is often no guarantee as to how good the reported solution is when the algorithm is terminated in finite time (as it must be in practice). See Andradóttir [1998] for specific algorithms that apply to discrete- or continuous-variable problems.

OutPut analysis of steady state simulation(unit 7 vvimp 10m):

Output Analysis for Steady-State Simulations I

- Consider a single run of a simulation model to estimate a steadystate or long-run characteristics of the system.
- The single run produces observations $Y_1, Y_2, \ldots$ (generally the samples of an autocorrelated time series).
- Performance measure:

$$\theta = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} Y_i, \qquad \text{for discrete measure}$$

$$\phi = \lim_{T_E \to \infty} \frac{1}{T_E} \int_0^{T_E} Y(t)dt, \qquad \text{for continuous measure}$$

independent of initial conditions, both with probability 1

- The sample size is a design choice, with several considerations in mind:

1.Initialization Bias.

2.Error Estimation

3.Replication mathods.

4.Sample size.

5.Batch means.

Initialization Bias I

- Methods to reduce the point-estimator bias caused by using artificial and unrealistic initial conditions:
  - Intelligent initialization.
  - Divide simulation into an initialization phase and data collection phase.
- Intelligent initialization
  - Initialize the simulation in a state that is more representative of long-run conditions.
  - If the system exists, collect data on it and use these data to specify more nearly typical initial conditions.
  - If the system can be simplified enough to make it mathematically solvable, e.g. queueing models, solve the simplified model to find long-run expected or most likely conditions, use that to initialize the simulation.
- Divide each simulation into two phases:
  - An initialization phase, from time 0 to time T_0.

Error Estimation I

- If $\{Y_1, \ldots, Y_n\}$ are not statistically independent, then S^2/n is a biased estimator of the true variance.
- Almost always the case when $\{Y_1, \ldots, Y_n\}$ is a sequence of output observations from within a single replication (autocorrelated sequence, time-series).
- Suppose the point estimator $\widehat{\theta}$ is the sample mean

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Variance of $\overline{Y}$ is very hard to estimate.
- For systems with steady state, produce an output process that is approximately covariance stationary (after passing the transient phase).

Replication Method I

- Use to estimate point-estimator variability and to construct a confidence interval.
- Approach: make R replications, initializing and deleting from each one the same way.
- Important to do a thorough job of investigating the initial-condition bias:
  - Bias is not affected by the number of replications, instead, it is affected only by deleting more data (i.e., increasing T_0) or extending the length of each run (i.e. increasing T_E).
- Basic raw output data $\{Y_{rj}, r = 1, \ldots, R,\ j = 1, \ldots, n\}$ is derived by:
  - Individual observation from within replication r.
  - Batch mean from within replication r of some number of discrete-time observations.

.

Sample Size I

- To estimate a long-run performance measure, θ, within $\pm\varepsilon$ with confidence $100(1-\alpha)\%$.
- M/G/1 queueing example (cont.):
  - We know: $R_0 = 10$, $d = 2$ deleted and $S_0^2 = 25.30$.
  - To estimate the long-run mean queue length, L_Q, within $\varepsilon = 2$ customers with 90% confidence ($\alpha = 10\%$).
  - Initial estimate:

$$R \geq \left(\frac{z_{0.05} S_0}{\varepsilon}\right)^2 = \frac{1.645^2(25.30)}{2^2} = 17.1$$

  - Hence, at least 18 replications are needed, next try $R = 18, 19, \ldots$ using $R \geq (t_{0.05,R-1} S_0/\varepsilon)^2$. We found that

$$R = 19 \geq (t_{0.05,R-1} S_0/\varepsilon)^2 = \left(1.73^2 \cdot 25.3/4\right) = 18.93$$

  - Additional replications needed is $R - R_0 = 19 - 10 = 9$.

Batch Means for Interval Estimation

- Using a single, long replication:
  - Problem: data are dependent so the usual estimator is biased.
  - Solution: batch means.
- Batch means: divide the output data from 1 replication (after appropriate deletion) into a few large batches and then treat the means of these batches as if they were independent.
- A continuous-time process, $\{Y(t), T_0 \le t \le T_0 + T_E\}$:
  - k batches of size $m = T_E/k$, batch means:

$$\overline{Y}_j = \frac{1}{m}\int_{(j-1)m}^{jm} Y(t+T_0)dt,\ j = 1, 2, \ldots, k$$

  - A discrete-time process, $\{Y_i, i = d+1, d+2, \ldots, n\}$:
    - k batches of size $m = (n-d)/k$, batch means:

$$\overline{Y}_j = \frac{1}{m}\sum_{i=(j-1)m+1}^{jm} Y_{i+d}, \qquad j = 1, 2, \ldots, k$$