INSTITUTE OF INFORMATION TECHNOLOGY

JAHANGIRNAGAR UNIVERSITY

Number of Assignment **:** 01

Submission Date **:** 08/01/2023

Course Tittle **:** Simulation and Modeling

Course Code **:** ICT - 3209

Submitted To

Md. Fazlul Karim Patwary

Professor

IIT – JU

Submitted By

Md. Shakil Hossain

Roll – 2023

3^{nd} year 2^{nd} Semester

IIT – JU

Question 1: The lifetime of light bulbs is known to be normally distributed with $\mu = 100h$ and $\sigma = 8h$. What is the Probability that a bulb Picked at random will have a lifetime between 110 and 120 burning hours?

Solution:

We can use the z-score formula to Convert the 110 and 120 burning hours.

The z-score formula is

$$z = (x - \mu)/\sigma$$

For the lower bound

$$z = \frac{(110 - 100)}{8}$$

$$= 1.25$$

For the upper bound

$$z = \frac{120 - 100}{8}$$

$$= 2.5$$

using calculator we can find the Probability of a random value being between 1.25 and 2.5 Standard deviations above the mean is about 0.216.

Therefore, the Probability that a bulb Picked at random will have a lifetime between 110 and 120 burning hours is about 0.216.

Question: 2 The average number of calls for service received by a machine repir department per 8-hr shift is 10.0. What is the Probability that more than 15 calls will be received during a randomly selected 8-hr shift?

Solution:

Using Poisson distribution the exact Probability is:

$$P(X > 15/\lambda = 10.0) = P(X=16) + P(X=17) + \cdots$$

$$= 0.0217 + 0.0128 + 0.0071 + 0.0037 + 0.0019 + 0.0009 + 0.0004 + 0.0002 + 0.0001$$

$$= 0.04869345$$

Using Normal distribution the Probability is

$$Z = (x - mean)/sd = (15.5 - 10)/sqrt(10) = 1.74$$

$$P(X > 15.5) = P(Z > 1.74)$$

$$= 0.0409$$

Question 3: Suppose a neutrophil count is defined as abnormally high if the number of neutrophils is ≥ 76 and abnormally low if the number of neutrophils is ≤ 49. Calculate the Proportion of People whose neutrophil Counts are abnormally high or low.

Solution:

Binomial Probability function is

$$\binom{100}{k}(0.6)^k(0.4)^{100-k}$$

We get

$$\sum_{k=76}^{100}\binom{100}{k}(0.6)^k(0.4)^{100-k} + \sum_{k=0}^{49}\binom{100}{k}(.6)^k(.4)^{100-k}$$

Normal approximation

$$\text{Mean} = np = 100(0.6) = 60$$

$$\text{Variance} = np(1-p) = 100(0.6)(0.4) = 24$$

Thus 1.7% of the People will be normal.

Question 4: If the time of failure for an electrical component follows an exponential distribution with a mean time to failure of 1000 hours what is the Probabily that a ~~random~~ randomly chosen component will fail before 750 hours?

Solution:

The Probability density function of the exponential distribution is

$$f(x) = \lambda * e^{-\lambda x}$$

where

$$\lambda = 1/1000 = 0.001$$

we can use the following formula to find the area under the curve

$$P(0 < x < 750) = \int f(x)\, dx$$

$$= \int (0.001 * e^{-0.001x})\, dx$$

$$= -e^{-0.001x} \Big|_0^{750}$$

$$= 1 - e^{-0.75}$$

$$= 0.774$$

What do you mean by system, state of a system and simulation? Give examples to explain these terms.

A system is a set of components that work together to achieve a common goal or function. A system can be a physical system, such as a car or a computer, or a abstract system, such as a business process or a social network.

The state of a system refers to the current values of all the variables that define the system at a given time. For example, the state of a car might include variables such as the speed, fuel level, and engine temperature. The state of a social network might include variables such as the number of users, the number of connections between users, and the average age of users.

Simulation is the process of creating a model of a system and using it to study the behavior of the system over time. Simulation can be used to predict the future behavior of the system, identify patterns and trends, and test different scenarios.

Examples:

1. A computer system is a collection of hardware and software components that work together to perform tasks such as data processing, storage, and communication. The state of the system might include variables such as the amount of available memory, the number of active processes, and the CPU utilization.
2. A traffic simulation is a model of a road network and the vehicles that travel on it. The simulation can be used to study the flow of traffic, predict congestion, and evaluate the impact of different traffic management strategies.
3. A business process simulation is a model of a series of activities and decisions that are carried out to achieve a specific business goal. The simulation can be used to study the efficiency of the process, identify bottlenecks, and optimize resource allocation.

Why does Single-server queue is called Discrete-Event Simulation?

A single-server queue is a model of a system where customers arrive at a service point and wait in line to be served by a single server. This type of system is often used to model a variety of real-world scenarios, such as a bank teller, a grocery store checkout, or a customer service call center.

Single-server queues are often studied using discrete-event simulation because they involve a series of discrete events, such as customer arrivals and departures that occur at specific times. In a discrete-event simulation, the model advances from one event to the next by updating the state of the system based on the events that have occurred.

For example, in a single-server queue simulation, the model might start with an empty queue and a server that is idle. When a customer arrives, the model updates the state of the system by adding the customer to the queue and checking whether the server is available to serve them. If the server is available, the model updates the state again by marking the server as busy and removing the customer from the queue. If the server is not available, the customer remains in the queue until the server becomes available.

Discrete-event simulation is a powerful tool for studying the behavior of single-server queues and other systems because it allows researchers to examine how different factors, such as the arrival rate of customers and the service time of the server, impact the performance of the system.

What are the factors that are considered in a Single-server queue?

There are several factors that are commonly considered in the analysis of a single-server queue:

1. Arrival rate: This is the rate at which customers arrive at the service point. It is usually measured in customers per unit of time, such as customers per minute or customers per hour.
2. Service rate: This is the rate at which the server is able to serve customers. It is also usually measured in customers per unit of time.
3. Service time: This is the time it takes for the server to serve a customer. It is usually measured in the same unit of time as the arrival rate and service rate.
4. Queue capacity: This is the maximum number of customers that can be waiting in the queue at any given time. If the queue reaches capacity, new customers may be turned away or may choose to leave the system.
5. Population size: This is the total number of customers in the system, including those in the queue and those being served by the server.
6. Utilization: This is the percentage of time that the server is busy serving customers. It is calculated as the service rate divided by the arrival rate.
7. Mean waiting time: This is the average amount of time that a customer spends waiting in the queue before being served.
8. Mean service time: This is the average amount of time that a customer spends being served by the server.
9. Mean time in system: This is the average amount of time that a customer spends in the system, including both the time spent waiting in the queue and the time spent being served.

These factors can be used to analyze the performance of the single-server queue and to optimize the system to improve customer satisfaction and efficiency.

Explain the mechanisms of Next-event time advance and Fixed-increment time advance in simulation.

In simulation, time advance refers to the process of advancing the simulated time of the model from one time step to the next. There are two main mechanisms for time advance in simulation: next-event time advance and fixed-increment time advance.

Next-event time advance is a mechanism for advancing the simulated time to the time of the next scheduled event in the model. This means that the simulated time only advances when there is an event scheduled to occur, such as a customer arrival or a server completing a service.

For example, consider a single-server queue simulation where customers arrive at random intervals. The simulated time in the model would advance to the time of the next customer arrival each time a customer arrived, and would remain unchanged until the next customer arrived. This approach is useful when events occur at irregular intervals and there is a need to accurately capture the timing of individual events.

Fixed-increment time advance is a mechanism for advancing the simulated time by a fixed increment at each time step. This means that the simulated time advances by a fixed amount, such as one minute or one hour, regardless of whether there are any scheduled events.

For example, consider a simulation of a manufacturing process where the state of the system is updated every hour. The simulated time in the model would advance by one hour at each time step, even if there are no events scheduled to occur during that hour. This approach is useful when events occur at regular intervals and there is less need to capture the timing of individual events.

Both next-event time advance and fixed-increment time advance have their own advantages and disadvantages, and the appropriate mechanism for a given simulation depends on the needs of the model and the goals of the simulation.

Why does probability distribution associated with simulation? Give examples of the use of discrete and continuous probability distribution in simulation.

Probability distributions are often associated with simulation because they are used to model the uncertainty and variability of real-world systems. In a simulation, probability distributions can be used to generate random numbers that represent the occurrence of events, such as the arrival of customers or the completion of a task.

There are two main types of probability distributions: discrete and continuous.

Discrete probability distributions are used to model variables that can take on a finite or countably infinite set of values. Examples of discrete probability distributions include the binomial distribution, the Poisson distribution, and the geometric distribution.

Continuous probability distributions are used to model variables that can take on any value within a given range. Examples of continuous probability distributions include the normal distribution, the uniform distribution, and the exponential distribution.

Examples of the use of discrete and continuous probability distributions in simulation include:

1. A single-server queue simulation might use a Poisson distribution to model the arrival rate of customers. The Poisson distribution is a discrete distribution that describes the probability of a given number of events occurring in a fixed time period.
2. A simulation of a manufacturing process might use a normal distribution to model the processing time of a machine. The normal distribution is a continuous distribution that describes the probability of a given value occurring within a certain range.
3. A simulation of a financial system might use a binomial distribution to model the probability of a stock price rising or falling. The binomial distribution is a discrete distribution that describes the probability of a given number of successes occurring in a fixed number of trials.

Define Negative Binomial Distribution and geometric distribution. Write the application of these distribution in simulation.

The negative binomial distribution is a discrete probability distribution that describes the probability of a given number of failures occurring before a specified number of successes. It is often used to model the number of failures that occur before a given event or threshold is reached.

The geometric distribution is a discrete probability distribution that describes the probability of a given number of failures occurring before the first success. It is often used to model the number of trials required to achieve a given event or outcome.

Both the negative binomial distribution and the geometric distribution are closely related to the binomial distribution, which describes the probability of a given number of successes occurring in a fixed number of trials.

Applications of the negative binomial distribution and the geometric distribution in simulation include:

1. A simulation of a manufacturing process might use a negative binomial distribution to model the number of defects that occur before a given number of units are produced to specification.
2. A simulation of a quality control process might use a geometric distribution to model the number of inspections required to detect a given number of defects.
3. A simulation of a supply chain might use a negative binomial distribution to model the number of orders that are lost or delayed before a given number of orders are fulfilled on time.

4. A simulation of a service system might use a geometric distribution to model the number of calls that are abandoned before a given number of calls are answered by a customer service representative.

What are the application of Q-Q and P-P plot? Construct a Q-Q plots for the inter arrival time of customers 10, 12, 18, 22 in seconds.

Q-Q plots (quantile-quantile plots) and P-P plots (probability-probability plots) are graphical tools used to compare two sets of data to determine whether they come from the same distribution.

Q-Q plots are used to compare the quantiles (fractions) of two datasets. A Q-Q plot is created by plotting the quantiles of one dataset on the x-axis against the quantiles of the other dataset on the y-axis. If the two datasets come from the same distribution, the points in the Q-Q plot should fall approximately on a straight line.

P-P plots are similar to Q-Q plots, but they plot the probability (fraction) of each data point on the x-axis against the probability of the corresponding data point on the y-axis. If the two datasets come from the same distribution, the points in the P-P plot should also fall approximately on a straight line.

Applications of Q-Q and P-P plots include:

1. Testing the goodness of fit of a statistical model to a dataset
2. Comparing the distributions of two datasets to determine whether they are significantly different
3. Evaluating the performance of a simulation model by comparing the output of the model to real-world data

To construct a Q-Q plot for the inter-arrival times of customers in seconds, you would first need to calculate the quantiles of the inter-arrival times. For example, if you have a sample of four inter-arrival times (10, 12, 18, 22) you would calculate the quantiles as follows:

Quantile 1: 10 seconds (0th quantile) Quantile 2: 11 seconds (0.25 quantile) Quantile 3: 15 seconds (0.5 quantile) Quantile 4: 20 seconds (0.75 quantile) Quantile 5: 22 seconds (1.0 quantile)

You would then plot the quantiles of the inter-arrival times on the y-axis against their corresponding quantiles on the x-axis. If the inter-arrival times follow a certain distribution, such as the exponential distribution, you would expect the points in the Q-Q plot to fall approximately on a straight line.

What is regression analysis and how do you should select dependent and independent variables?

Regression analysis is a statistical method used to model the relationship between a dependent variable and one or more independent variables. It is used to predict the value of the dependent variable based on the values of the independent variables.

The dependent variable is the variable that is being predicted or explained by the independent variables. It is also known as the outcome variable or the response variable.

The independent variables, also known as the predictor variables or the explanatory variables, are the variables that are used to predict or explain the dependent variable.

In regression analysis, it is important to carefully select the dependent and independent variables to ensure that the model accurately reflects the relationship between the variables. Here are some guidelines for selecting dependent and independent variables:

1. The dependent variable should be a continuous variable, such as a numerical measurement or a score on a scale.
2. The independent variables should be relevant to the dependent variable and should be able to explain or predict the variation in the dependent variable.
3. The independent variables should be independent of each other, meaning that the value of one independent variable should not depend on the value of any other independent variable.
4. The independent variables should be measured at the same level of precision as the dependent variable.
5. The relationship between the dependent and independent variables should be linear, meaning that the change in the dependent variable should be directly proportional to the change in the independent variable.

Chi-square Goodness-of-fit Test is widely use after a simulation performed, what are the process of conduction this test and why do researchers use this test?

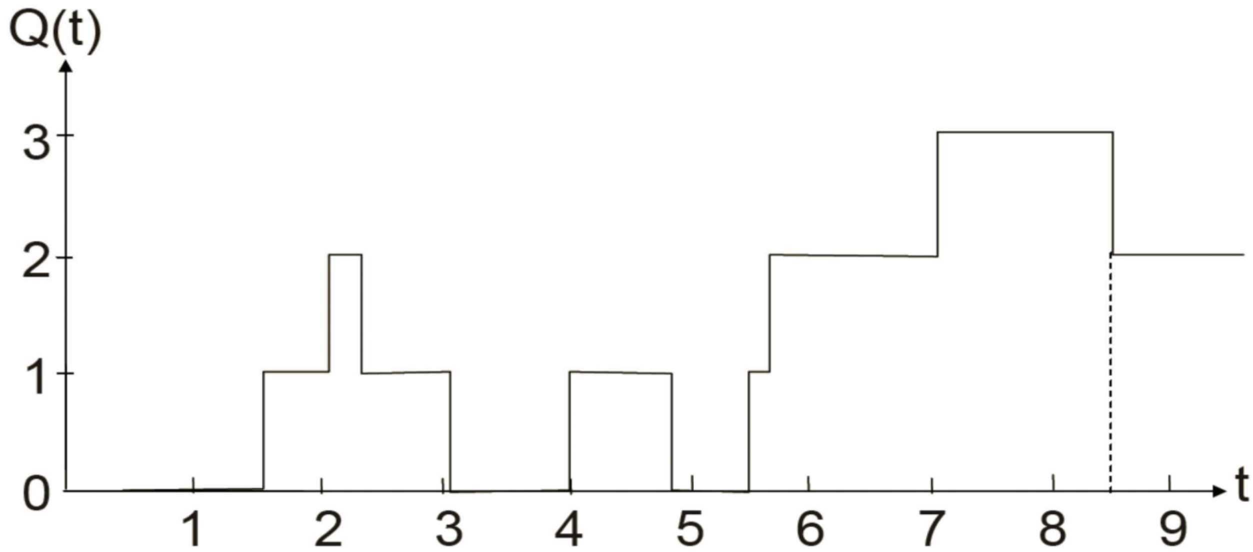
The chi-square goodness-of-fit test is a statistical test used to determine whether a sample data set is consistent with a specified probability distribution. It is often used after a simulation has been performed to assess the quality of the simulation model.

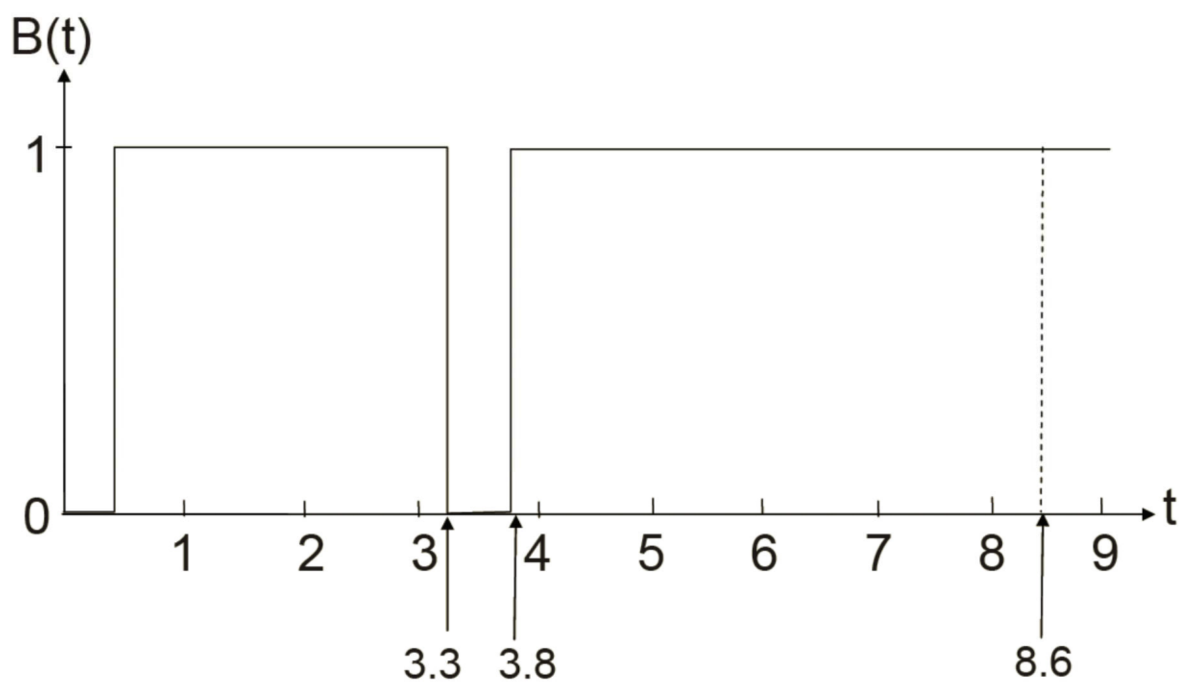
To conduct a chi-square goodness-of-fit test, the following steps should be followed:

1. Specify the probability distribution that you want to test. This could be a theoretical distribution, such as the normal distribution, or a distribution estimated from a sample of data.

2. Divide the range of possible values of the variable into bins or categories. For example, if the variable is a continuous variable, you might divide the range into 10 or 20 bins of equal width.
3. Calculate the expected number of observations in each bin under the specified probability distribution. This is done by evaluating the probability distribution function at the midpoint of each bin and multiplying the result by the sample size.
4. Count the number of observations in each bin in the sample data set.
5. Calculate the chi-square statistic as the sum of the squared differences between the observed and expected counts, divided by the expected counts.
6. Compare the calculated chi-square statistic to the critical value of the chi-square distribution with a degree of freedom equal to the number of bins minus one. If the calculated chi-square statistic is greater than the critical value, the null hypothesis that the sample data come from the specified distribution can be rejected

Event No.	time	Description	Server status	No. in queue	Times of arrival	Time of last event	No. delayed	Total delay	Area under Q(t)	Area under B(t)
1	0.4	Arrival 1	0	0	---	0.4	0	0	0	0
2	1.6	Arrival 2	1	1	1.6	1.6	1	0	0	1.2
3	2.1	Arrival 3	1	2	1.6, 2.1	2.1	1	0	0.5	1.7
4	2.4	Departure 1	1	1	2.1	2.4	2	0.8	1.1	2.0
5	3.1	Departure 2	1	0	---	3.1	3	1.8	1.8	2.7
6	3.3	Departure 3	0	0	---	3.3	3	1.8	1.8	2.9
7	3.8	Arrival 4	1	0	---	3.8	4	1.8	1.8	2.9
8	4.0	Arrival 5	1	1	4.0	4.0	4	1.8	1.8	3.1
9	4.9	Departure 4	1	0	---	4.9	5	2.7	2.7	4.0
10	5.6	Arrival 6	1	1	5.6	5.6	5	2.7	2.7	4.7
11	5.8	Arrival 7	1	2	5.6, 5.8	5.8	5	2.7	2.9	4.9
12	7.2	Arrival 8	1	3	5.6, 5.8, 7.2	7.2	5	2.7	5.7	6.3
13	8.6	Departure 5	1	2	5.8, 7.2	8.6	6	5.7	9.9	7.7


Q(t)
3
2
1
0
1 2 3 4 5 6 7 8 9
t


B(t)
1
0
1 2 3 4 5 6 7 8 9
t
3.3
3.8
8.6