



University of Vavuniya

First Examination in Information Technology - 2022

First Semester - January/February 2024

IT1152 Essentials of Statistics

Answer Four Questions Only

(Non-programmable calculator is allowed to be used)

Allowed Time : Two hours

1. (a) Distinguish between discrete and continuous data with examples. [20%]

- (b) The following data shows the annual salaries of IT professionals in a company. The data is grouped into salary ranges, and the frequency represents the number of IT professionals falling into each salary range.

Salary Range (in US dollars)	Number of IT Professionals
50,000 - 60,000	15
60,000 - 70,000	20
70,000 - 80,000	25
80,000 - 90,000	18
90,000 - 100,000	12

- i. Construct a histogram for this data, and comment on it. [15%]
- ii. Calculate the following measures: mean, median, mode, all quartile values, and standard deviation. [55%]
- iii. Find the Coefficient of Skewness of the data, and check your opinion about the histogram you provided in part (i). [10%]

2. (a) What do you mean by a random event in probability? [10%]
- (b) Describe mutually exclusive events with examples. [15%]
- (c) A new computer virus can enter the system through e-mail or through the internet. There is a 30% chance of receiving this virus through e-mail. There is a 40% chance of receiving it through the internet. Also, the virus enters the system simultaneously through e-mail and the internet with probability 0.15. What is the probability that
- the virus enters the system through e-mail or internet? [15%]
 - the virus enters the system only through the internet? [10%]
 - the virus does not enter a system simultaneously via e-mail and the internet? [10%]
 - the virus does not enter the system at all? [10%]
- (d) Three computer viruses arrived as an e-mail attachment. Virus A damages the system with probability 0.4. Independently of it, virus B damages the system with probability 0.5. Independently of A and B, virus C damages the system with probability 0.2. What is the probability that
- the system gets damaged by all three viruses? [10%]
 - the system gets damaged by only virus A and B? [10%]
 - the system gets damaged by either virus A or B but not by C? [10%]
3. (a) i. Define conditional probability. [10%]
- ii. Suppose you are a system administrator responsible for managing a network of servers. Two critical events are of interest: A , the event that a server experiences a hardware failure, and B , the event that a software update is applied to the server. You have historical data and estimates for the following probabilities:
- Probability of a hardware failure without a software update is 0.03.
 - Probability of a software update being applied is 0.15.
 - Probability of a hardware failure and a software update is 0.01.

[This question is continued on the next page]

Calculate the probability of a hardware failure given that a software update has been applied. [30%]

(b) i. State Bayes' theorem. [20%]

ii. In a large corporate network, there are two data centers, Data Center A and Data Center B. 60% of events originate from Data Center A, and 40% originate from Data Center B. Network administrators are concerned about a specific type of cyber threat (T). They have deployed security measures independently in each data center to detect and prevent this threat. The probability of detecting the threat in Data Center A is 0.95, and the probability of detecting the threat in Data Center B is 0.90. Suppose network administrators detect a threat in a particular event, what is the probability that the event originates from Data Center B? [40%]

4. (a) i. Distinguish between discrete and continuous random variables. [20%]

ii. Suppose an IT system undergoes regular security audits to identify vulnerabilities. The number of vulnerabilities discovered per audit follows a discrete probability distribution. The table below represents the probability distribution of the number of vulnerabilities (X):

Number of Vulnerabilities (X)	Probability (P(X))
0	0.15
1	0.30
2	0.25
3	0.20
4	0.10

Check whether the above probability distribution is correct. If so, calculate

A. the probability that the IT system detects at least two vulnerabilities. [15%]

B. the probability that the IT system does not detect any vulnerabilities. [05%]

C. the average number of identified vulnerabilities along with standard deviation. [15%]

[This question is continued on the next page]

- (b) i. List the conditions for Binomial Distribution. [10%]
- ii. A quality control engineer tests the quality of produced computers. Suppose that 5% of computers have defects, and defects occur independently of each other. In a shipment of twenty computers, find the probability that
- A. exactly three computers are defective. [10%]
- B. at least three computers are defective. [25%]
5. (a) i. List the conditions for Poisson distribution. [10%]
- ii. Messages arrive at an electronic message center at random times, with an average of 9 messages per hour.
- A. What is the probability of receiving at least five messages during the next hour? [25%]
- B. What is the probability of receiving exactly five messages during the next hour? [10%]
- (b) i. Define normal distribution. [10%]
- ii. Suppose you have collected response time data from 1000 requests to a specific API endpoint, and you find that the response times follow a normal distribution with a mean (μ) of 50 milliseconds and a standard deviation (σ) of 10 milliseconds. What is the probability that a randomly selected request will have a response time
- A. less than 60 milliseconds. [15%]
- B. between 45 and 65 milliseconds. [15%]
- C. more than 55 milliseconds. [15%]