



# Daffodil International University

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Midterm Examination, Spring 2023

Course Code: STA 221, Course Title: Statistics and Probability

Level: 2 Term: 2 Batch: 60

Time: 01.5 Hrs

Marks: 25

## Answer ALL Questions

[The figures in the right margin indicate the full marks and corresponding course outcomes. All portions of each question must be answered sequentially.]

1.	a)	The course teacher asked randomly 10 of his students about the best measures of central tendency after delivered his lecture. Among them 70% students provide correct answer. Select variable, population, sample and statistic from the above statement.	[2]	CO1
	b)	Find the difference between qualitative and quantitative variables? Name out different graphs and charts for both variables.	[2]	
	c)	Choose the best measures of dispersion and give an explanation.	[1]	
2.	a)	The last investment of randomly selected different software farms was surveyed as following 25, 12, 13, 39, 23, 27, 86, 50, 65, 67, 89, 90, 145, 120, 23, 27, 34, 290, 36, 40		CO2
	i)	Construct Box-plot from the above data.	[4]	
	ii)	Select the outlier if have any.	[2]	
	b)	Develop the shape of the distribution (Skewness) from the above data 2 (a)	[4]	
3.	a)	The mean life time of a sample of 50 light tubes produced by a company is found to be 1100 hours and the population standard deviation is 70 hours. Test the hypothesis at the 5% level of significance that the mean of the tubes produced by the company is 1050 hours. [Tabulated Value=1.96]	[5]	CO3
	b)	In a particular department 21, 24, 32, 39, 27, 13, 35, 22, 23 are the number of students in different sections are expert at C++ and 13, 11, 39, 21, 24, 16, 25 are the expert at python. Compare the two programming language according to consistency of the students.	[5]	

## Formula

### Measures of Central Tendency

#### Mean

$$A.M = \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$W.M = \bar{x} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}$$

$$G.M = \bar{x} = ((x_1 \cdot x_2 \cdot x_3 \dots x_n))^{1/n}$$

$$H.M = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

#### Median

If "n" is odd,  $M_e = X_{1/2(n+1)}$

If "n" is even,  $M_e = \frac{1}{2} (X_{n/2} + X_{\frac{n}{2}+1})$

### Measure of Dispersion

$$\text{Range} = X_{\max} - X_{\min}$$

$$\text{Mean Deviation, M.D} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Population variance

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

Population standard deviation,

$$\sqrt{\sigma^2}$$

Sample variance

$$s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{n-1}$$

Population standard deviation,

$$\sqrt{s^2}$$

Coefficient of variation for population,  $C.V = \frac{\sigma}{\mu} \times 100$

Coefficient of variation for sample,  $C.V = \frac{s}{\bar{x}} \times 100$

### Measures of Location

$$\text{Quartile, } Q_l = \frac{l(n+1)}{4}$$

$$\text{Decile, } D_l = \frac{l(n+1)}{10}$$

$$\text{Percentile, } P_l = \frac{l(n+1)}{100}$$

You can also use

Quartile

$$Q_l = \frac{l \times n}{4}$$

Deciles

$$D_l = \frac{l \times n}{10}$$

Percentile

$$P_l = \frac{l \times n}{100}$$

### Hypothesis testing

Z test statistics

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

t test statistics

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

### Shape of the distribution

$$\text{Coefficient of Skewness, Sk} = \frac{3 \times (\text{Mean} - \text{Median})}{\text{Standard deviation}}$$

$$\text{Kurtosis } \beta_2 = \frac{\mu_4}{\mu_2^2}$$