



Daffodil International University
Department of Computer Science and Engineering
Faculty of Science and Information Technology
Mid-term Examination, Spring 2024
Course Code: STA227, Course Title: Statistics and Probability

Time: 01:30 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)	Suppose, an economist wanted to know the average income of the people of Bangladesh. So, he divided the whole population into 495 sub-districts and 50 sub-districts were selected randomly. Then income of every household of that 50 sub-districts were recorded. Find population, sample, variable, type of variable, level of measurement and employed sampling technique from the above scenario.	[3]	CO1														
	b)	Choose the appropriate measures of central tendency for nominal and ordinal data? Define which measure is applicable at all levels of measurements?	[2]															
2.	a)	The quiz marks of 10 students are 13, 9, 12, 10, 14, x, 15, 11, 14, 9 Identify the value of x when the average is 11.5 for the 10 students.	[2]	CO2														
	b)	Suppose, following data represents the amount of time (in minutes per day) spending on social media of 15 students in a shelf: 250, 320, 280, 310, 400, 290, 300, 330, 180, 290, 480, 285, 275, 260, 295.																
	i)	Construct an appropriate graph using the above data.	[3]															
	ii)	Identify if there is any “Outlier” in the above dataset. Also represent with a modified box-Plot.	[4]															
	iii)	Identify the best measure of central tendency in this case with appropriate explanation.	[1]															
3.	a)	The weekly sales of phone of two brands Samsung and Oneplus are recorded as below: <table border="1"><tr><td>OnePlus</td><td>59</td><td>75</td><td>95</td><td>100</td><td>56</td><td>85</td></tr><tr><td>Samsung</td><td>80</td><td>90</td><td>150</td><td>125</td><td>82</td><td>95</td></tr></table>	OnePlus	59	75	95	100	56	85	Samsung	80	90	150	125	82	95		CO3
	OnePlus	59	75	95	100	56	85											
	Samsung	80	90	150	125	82	95											
	Compare brand seems to be more consistent in respect to sales. Solve using relative measure of dispersion.	[5]																
b)	In a survey of 10 CSE students, we gathered data on their weekly coding practice (in hour). The recorded times are as follows: 3.0, 2.9, 3.7, 4.0, 3.1, 3.6, 4.3, 3.4, 2.2, 2.1, Analyze the shape of the distribution by using the formula of skewness and comment on the shape.	[5]																

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

Measures of Location

Quartile

$$Q_i = \frac{i \times n}{4}$$

Deciles

$$D_i = \frac{i \times n}{10}$$

Percentile

$$P_i = \frac{i \times n}{100}$$

Inner fence

$$Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR$$

Outer fence

$$Q_1 - 3 \times IQR, Q_3 + 3 \times IQR$$

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

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

Sample standard deviation, $\sqrt{s^2}$

Coefficient of variation for

population, $C.V = \frac{\sigma}{\mu} \times 100$

Coefficient of variation for

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

Shape of the distribution

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

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