

Reg. No.:



Y8736

Name:

University of Kerala

Second Semester FYUGP Degree Examination, June 2026

Discipline Specific Core Course

STATISTICS**UK2DSCSTA102 - QUANTITATIVE DATA ANALYTICS II**

Academic Level: 100-199

2024 Admission onwards**Time: 2 Hours(120 Mins)****Max. Marks: 56**

Part A. 6 Marks:Time 5 Minutes.(Cognitive Level :Remember(RE)/Understand(UN)) Objective Type.1 mark each, Answer all questions

Qn No.	Question	CL	CO
1	Name a discrete distribution for which mean is equal to variance.	RE	3
2	If A and B are mutually exclusive, then $P(A \cap B) = \text{-----}$	RE	3
3	What does a positive value of Yule's coefficient of association indicate about the relationship between two attributes?	UN	4
4	If the standard deviations of X and Y are 3 and 4 respectively and the covariance between them is 12, what is the correlation coefficient between X and Y?	UN	1
5	How are the variables 'amount of exercise a person does and their weight' correlated? Negatively correlated	UN	1
6	State the relationship between the correlation coefficient and the regression coefficients.	UN	2

Part B. 10 Marks.Time:20 Minutes (Cognitive Level:Understand(UN)/Apply(AP))Two-three sentences.2 marks each.Answer all questions

Qn No.	Question	CL	CO																
7	Explain Yule's coefficient of colligation and describe how it is calculated using the appropriate formula.	UN	4																
8	Explain the properties of Karl Pearson's Correlation Coefficient.	UN	1																
9	The regression coefficient of y on x is - 0.5 and correlation coefficient between x and y is - 0.80. Compute the regression coefficient of x on y.	AP	2																
10	Let X be a random variable with the following probability mass function. Determine the value of <table border="1"><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr><tr><td>P(X=x)</td><td>a</td><td>3a</td><td>5a</td><td>7a</td><td>9a</td><td>11a</td><td>13a</td></tr></table> a. Also obtain $P(X = 2)$.	x	0	1	2	3	4	5	6	P(X=x)	a	3a	5a	7a	9a	11a	13a	AP	3
x	0	1	2	3	4	5	6												
P(X=x)	a	3a	5a	7a	9a	11a	13a												
11	For a Binomial distribution the number of trials is 20 and probability of success is 0.4.Find the mean and variance of the distribution.	AP	3																

Part C. 16 Marks.Time:35 Minutes.(Cognitive Level :Apply(AP)/Analyse(AN))Short Answer.4 marks each, Answer all 4 questions,choosing among options * within each question

Qn No.	Question	CL	CO
12	A) In a survey of 2,980 people, 230 individuals were smokers and had respiratory issues, 260 were non-smokers with respiratory issues, and 2,340 were non-smokers without respiratory issues, the rest were smokers without respiratory issues. Find the missing frequency and complete the 2×2 table of attributes. OR B) Given $P(A) = 0.2$ and $P(B) = 0.3$, determine $P(A \cup B)$ if (i) A and B are independent, (ii) A and B are mutually exclusive.	AP	4, 3
13	A) If E and F are events such that $P(E)=0.6, P(F)=0.3$ and $P(E \cap F)=0.7$, what is $P(E/F)$? OR	AP	3, 3

Qn No.	Question	CL	CO
	B) The mean and variance of a binomial distribution are 6 and 2 respectively. Find (i) $P(X = 0)$ (ii) $P(X \geq 1)$		
14	A) Given the data set $X = (3, 6, 9, 12)$, $Y = (5, 10, 15, 20)$, (i) Analyse the relationship between X and Y by plotting a scatter diagram and interpret the observed pattern in the scatter plot. (ii) Explain the nature of the correlation between the variables by computing the correlation coefficient. OR B) Given that number of observations as pairs of X and Y = 16, standard deviation of X-series = 3.21, standard deviation of Y-series = 3.42, sum of the product of deviations of X and Y from their respective means = 142, (i) Calculate the coefficient of correlation between the X and Y series. (ii) Interpret the result in the context of the relationship between the two variables.	AN	1, 1
15	A) Can $10x + 2y - 12 = 0$ and $12x - 8y + 10 = 0$ be the two regression equations? Justify your answer. OR B) The two regression lines between variables X and Y are $3x + 4y - 24 = 0$ and $5x + 2y - 26 = 0$. Identify the regression lines and compute the means of X and Y.	AN	2, 2

Part D. 24 Marks. Time: 60 Minutes. (Cognitive Level : Analyse(AN)/Evaluate(EV)/Create(CR)) Long Answer 6 Marks each. Answer all 4 questions choosing among options * within each question

Qn No.	Question	CL	CO																						
16	A) A continuous random variable X has the probability density function (pdf): $f(x) = c(4 - x), 0 \leq x \leq 4$ Analyse the pdf and (i) Determine the value of c (ii) Find $P(1 \leq X \leq 3)$ OR B) A 50-year-old husband has a probability of $3/7$ of living till he is 80, while a 40-year-old wife has a probability of $5/13$ of living till she is 70. Analyze the survival probabilities by calculating the probability that (i) both will be alive at the given ages (ii) none will be alive (iii) only wife will be alive (iv) only husband will be alive (v) only one will be alive (vi) at least one will be alive	AN	3, 3																						
17	A) Given two regression equations $3x - y - 25 = 0$ and $2x - 3y + 30 = 0$. Find i) Mean values of x and y ii) Coefficient of correlation iii) s.d of x if s.d of y is 2 OR B) Given the heights (X, in inches) and weights (Y, in kilograms) of 10 individuals, (i) calculate the coefficient of correlation between these two variables. (ii) critically evaluate the meaning of the correlation value. Interpret the result so obtained. <table border="1"><tr><td>X</td><td>62</td><td>72</td><td>78</td><td>58</td><td>65</td><td>70</td><td>66</td><td>63</td><td>60</td><td>72</td></tr><tr><td>Y</td><td>50</td><td>65</td><td>63</td><td>50</td><td>54</td><td>60</td><td>61</td><td>55</td><td>54</td><td>65</td></tr></table>	X	62	72	78	58	65	70	66	63	60	72	Y	50	65	63	50	54	60	61	55	54	65	EV	2, 1
X	62	72	78	58	65	70	66	63	60	72															
Y	50	65	63	50	54	60	61	55	54	65															
18	A) In a competitive examination, 2,500 candidates appeared, out of which 500 passed. Among the total candidates, 400 had attended a coaching class and 250 of them passed the exam. (i) Construct a 2×2 contingency table and calculate Yule's coefficient of association (Q) for the attributes. (ii) Evaluate the strength and direction of the association using the value of Q. Based on your evaluation, assess whether attending coaching classes has a positive, negative or no significant effect on the likelihood of passing the exam. Justify your answer with interpretation. OR B) In a partially destroyed laboratory, the following results of a correlation analysis are partially legible: "Variance of X= 9, Regression equations : $8X - 10Y + 66 = 0, 40X - 18Y = 214$." (i) Evaluate the mean values of X and Y, (ii) Assess the correlation between X and Y by calculating the correlation coefficient, (iii) Determine the standard deviation of Y.	EV	4, 2																						

Qn No.	Question	CL	CO																
	A) Calculate the regression equations of Y on X from the data given below: <table><tr><td>Price (Rs.) (X)</td><td>10</td><td>12</td><td>13</td><td>12</td><td>16</td><td>15</td></tr><tr><td>Amount Demanded (Y)</td><td>40</td><td>38</td><td>43</td><td>45</td><td>37</td><td>43</td></tr></table> Estimate the likely demand when the price is Rs.20. OR	Price (Rs.) (X)	10	12	13	12	16	15	Amount Demanded (Y)	40	38	43	45	37	43				
Price (Rs.) (X)	10	12	13	12	16	15													
Amount Demanded (Y)	40	38	43	45	37	43													
19	B) An environmental scientist studies the number of green initiatives (X) and the reduction in carbon footprint (Y, in metric tons). Using the provided data, (i) Formulate a regression model to predict carbon reduction based on initiatives taken. (ii) Predict carbon reduction when 20 initiatives are taken. <table><tr><td>X</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>5</td></tr><tr><td>Y</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>6</td></tr></table>	X	2	3	4	5	6	7	5	Y	3	4	5	6	7	8	6	CR	2, 2
X	2	3	4	5	6	7	5												
Y	3	4	5	6	7	8	6												