


National University of Computer and Emerging Sciences, Lahore Campus

	Course Name:	Statistical and Mathematical Methods for Data Science	Course Code:	DS 501
	Program:	MS Data Science	Semester:	Fall 2019
	Duration:	60 Minutes	Total Marks:	40
	Paper Date:	September 24, 2019	Weight	
	Section:	N/A	Page(s):	3
	Exam Type:	Midterm Exam 1		

Student : Name: _____ **Roll No.** _____

- Instruction/Notes:**
1. Solve in the space provided. Extra sheets are NOT allowed
 2. One A4 handwritten help sheet is allowed in the exam
 3. Mobile phones, laptops, smart watches etc. are NOT allowed
 4. Use of calculators is allowed
 5. Sharing calculators is NOT allowed
 6. In case of any ambiguity make a reasonable assumption.
- Good luck!

QUESTION 1 **(Marks: 7+3)**

In a study it is found that 40% tall people are obese. An overall 20% people are tall.

- a. If a person is selected at random then what is the probability that this person is **both tall and not obese**? Show formula and working.
- b. Suppose that the proportion of obese people is 0.1 then can we say that height and obesity are independent? Show all working to make your decision? (A simple yes or no IS NOT ACCEPTABLE)

QUESTION 2**(Marks: 10)**

Suppose that 70% cold drinks get sold on a sunny day. Only 20% cold drinks get sold on a non-sunny day. The probability of a sunny day is 0.4. What is the overall fraction of cold drinks that get sold? A mathematical formula and working both are required.

QUESTION 3**(Marks: 10)**

Find $P(C=1|\mathbf{x}=(1,0))$ using naive Bayes' assumption. Show all working and formulas.

x_1	1	0	1	0	1
x_2	0	0	1	1	0
C	1	1	1	0	0

QUESTION 4 (no working required for any part)**(Marks: 10)**

a. What is the unbiased estimate of covariance between x and y?

x	-2	-1	3	0
y	0	-3	4	2

Answer: _____

b. The correlation of a random variable z with itself is: _____

c. If A and B are mutually exclusive exhaustive events then: $P(A \cup B) =$ _____ $P(A, B) =$ _____d. We have two independent variables x and y. $\text{var}(x)=4$ and $\text{var}(y) = 9$. A point in 2D space is written as (x,y). What is the Mahalanobis distance between (1,3) and (2,5)?

Answer: _____

e. Given the following table of joint distribution of x and y (each cell indicates $P(x,y)$):

	y = 0	y = 1
x = 0	0.2	0.4
x = 1	0.1	0.3

 $P(x=0|y=1) =$ _____