National University of Computer & Emerging Sciences, Karachi
Spring-2022 CS-Department
Final Examination 10th June 2022, 8:30 am to 11:30 am

Course Code: MT2005	Course Name: Probability & Statistics
Instructor Name / Names: Mr. Muhammad Jamil Usmani, Mr. Muhammad Amjad, Dr. Syed Muhamad Fahad Riaz,	
Student Roll No:	Section No:

Instructions:
- Return the question paper.
- Read each question completely before answering it. There are **10 questions and 04 pages**
- In case of any ambiguity, you may make an assumption. But your assumption should not contradict any statement in the question paper.
- All the answers must be solved according to the sequence given in the question paper.

Time: 3 hrs.   **Total Points** = 100(50% weightage)

Question # 01:   [2+6]

Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations L_1, L_2, L_3, and L_4 will be operated 40%, 30%, 20%, and 30% of the time. If a person who is speeding on her way to work has probabilities of 0.2, 0.1, 0.5, and 0.2, respectively, of passing through these locations,

I.   Draw the tree diagram.
II. What is the probability that she will receive a speeding ticket?

Question # 02:   [5+3]

a) A shipment of 7 computer sets contains 2 defective sets. An institute makes a random purchase of 3 of the sets. If X is the number of defective sets purchased by institute.

I.   Find the probability distribution of X.
II. Find the cumulative distribution function of the random variable X representing the number of defectives then using $F(x)$ $find$ $P(0 < X \le 2)$

b) The probability distribution of the discrete random variable X is

$$f(x) = \binom{3}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3.$$

Fina the mean and variance of X

Question # 03:   **[6]**

    I.   Installation of some software package requires downloading 82 files. On the average, it takes 15 sec to download one file, with a variance of 16 sec2.
            What is the probability that the software is installed in less than 20 minutes?

    II.   Suppose that the average household income in some country is 900 coins, and the standard deviation is 200 coins. Assuming the Normal distribution of incomes, compute the proportion of "the middle class," whose income is between 600 and 1200 coins.

Question # 04:   **[3+5]**

If an unauthorized person accesses a computer account with the correct username and password (stolen or cracked), can this intrusion be detected? Recently, a number of methods have been proposed to detect such unauthorized use. The time between keystrokes, the time a key is depressed, the frequency of various keywords is measured and compared with those of the account owner. If there are significant differences, an intruder is detected.
A longtime authorized user of the account makes an average 0.2 seconds between keystrokes.
One day, as someone typed the correct username and password.
The following times between keystrokes were recorded when a user typed the username and password:
.24, .22, .26, .34, .35, .32, .33, .29, .19, .36, .30, .15, .17, .28, .38, .40, .37, .27 seconds

    I.   Construct a 99% confidence interval for the mean time between keystrokes assuming normal distribution of these times.
    II.   Is this evidence of an unauthorized attempt at a 5% level of significance

Question # 05:   **[3+5]**

CD writing is energy consuming; therefore, it affects the battery lifetime on laptops. To estimate the effect of CD writing, 30 users are asked to work on their laptops until the "low battery" sign comes on. Eighteen users without a CD writer worked an average of 5.3 hours with a standard deviation of 1.4 hours. The other twelve, who used their CD writer, worked an average of 4.8 hours with a standard deviation of 1.6 hours. Assuming Normal distributions with equal population variances.

    I.   Construct a 95% confidence interval for the battery life reduction caused by CD writing.
    II.   Does a CD writer consume extra energy, and therefore, does it reduce the battery life on a laptop at $\alpha = 0.05$ [Critical value = 1.645]

Question # 06:   **[7+3]**

Several neurosurgeons wanted to determine whether a dynamic system (Z-plate) reduced the operative time relative to a static system (ALPS plate). The data displayed in the table below on operative times, in minutes, for the two systems.
At the 5% significance level, does the data provide sufficient evidence to conclude that the mean operative time is less with the dynamic system than with the static system? Also calculate the confidence interval. Assume population variances are unequal. [critical value: -1.74]

   Dynamic: 370, 360, 510, 445, 295, 315, 490, 345, 450, 505, 335, 380, 325, and 500.

   Static: 430, 445, 455, 455, 490, and 535.

Question # 07:   **[3+7]**

Do students sleep more in Karachi or in Lahore? Data from surveys in elementary mathematics classes were collected from the different schools of both cities. Their summary for the number of hours' students sleep is given as:

	N	Mean	St. Dev.
Karachi	173	6.93	1.71
Lahore	190	7.11	1.95

   I.   Calculate the 95% confidence interval of mean.
   II. Is there sufficient evidence to conclude that the mean hours of sleep are different at the two schools at 5% level of significance?

Question # 08:   **[1+1+3+2+8]**

   Number of fires and number of acres burned are as follows:

Fires:	72	69	58	47	84	62	57	45
Acres:	62	41	19	26	51	15	30	15

   I.   Identify dependent and independent variables.
   II.   Draw a scatter plot.
   III. Compute the regression model of type $\mu_{Y|x} = \beta_0 + \beta_1 x$
   IV.   Determine coefficient of correlation "r"
   V.   Verify the identity SST = SSR+SSE

Question # 09: [8+4]

The average monthly electric power consumption (Y) at a certain manufacturing plant is considered to be linearly dependent on the average ambient temperature (x_1) and the number of working days in a month (x_2). Consider the one-year monthly data given in the table below. **[16]**

Y	210	206	260	244	271	285	270	265	234	241	258	230
X_1	20	21	24	25	24	26	25	25	24	25	25	23
X_2	23	21	24	25	24	26	25	25	24	25	25	23

I. Estimate unknown parameter of multiple linear regression equation $\mu_{Y|x1,x2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

II. Determine the multiple coefficients of determination (R^2), where

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

Question # 10: [10+5]

The data in the table given below represent the number of hours of relief provided by five different brands of headache tablets administered to 25 subjects experiencing fevers of 38°C or more.

Tablet

A	B	C	D	E
5.2	9.1	3.2	2.4	7.1
4.7	7.1	5.8	3.4	6.6
8.1	8.2	2.2	4.1	9.3
6.2	6.0	3.1	1.0	4.2
3.0	9.1	7.2	4.0	7.6

I.   Perform the analysis of variance (ANOVA)
II. Test the hypothesis at the 0.05 level of significance that the mean number of hours of relief provided by the tablets is the same for all five brands. [Critical value = 2.84]
III. Discuss the result.