

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

### Instructions:

- Return the question paper.
- Read each question completely before answering it. There are **10 questions and 03 pages**
- In case of any ambiguity, you may make an assumption. But your assumption should not contradict any statement in the question paper.
- **It is necessary to show all working steps for the hypothesis testing and regression problems.**
- All the answers must be solved according to the sequence given in the question paper.

**Time: 180 minutes**

**Total Points = 100**

**Question # 01:** **CLO-1** **[2+6 = 8]**

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,

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

**Question # 02:** **CLO-2** **[5+2+1 = 8]**

- 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. Find the following:
  - the probability distribution of  $X$ .
  - the cumulative distribution function  $F(X)$
  - $P(0 < x \leq 2)$  using part (ii).
- Find the mean and variance for the given PMF of  $X$ .

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

**Question # 03:** **CLO-2** **[6]**

- A ball bearing's diameter is an important measurement in industrial processes. It is specified that the diameter should be in the interval of 2.99 to 3.01 cm. This implies that parts that fall outside of these specifications will not be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with a mean of 3.0 and a standard deviation of 0.005. On average, what percentage of manufactured ball bearings will be scrapped?
- Suppose that the average household income in some country is \$900, and the standard deviation is \$200. Assuming the normal distribution of incomes, compute the proportion of "the middle class," whose income is between \$600 and \$1200.

**Question # 04:**

CLO-3

**[3+5 = 8]**

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:

**0.24, 0.22, 0.26, 0.34, 0.35, 0.32, 0.33, 0.29, 0.19, 0.36, 0.30, 0.15, 0.17, 0.28, 0.38, 0.40, 0.37, 0.27 seconds.**

- (a) Is this evidence of an unauthorized attempt at a 99% level of significance? **[Critical value: 2.898]**
- (b) Construct a 99% confidence interval for the mean time between keystrokes assuming normal distribution of these times.

**Question # 05:**

CLO-3

**[3+5 = 8]**

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**.

- (a) Construct a 95% confidence interval for the battery life reduction caused by CD writing. **(critical value: 2.048)**
- (b) Does a CD writer consume extra energy, and therefore, does it reduce the battery life on a laptop at  $\alpha = 0.05$  **[Critical value: 1.701]**

**Question # 06:**

CLO-3

**[3+7 =10]**

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.

- (a) 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? (critical value: 2.015)
- (b) Also calculate the 95% confidence interval with critical value 2.571. **Assume population variances are unequal.**

**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:**

CLO-3

**[7+3 = 10]**

A demographic expert wants to decide whether in Pakistan, the mean age of married men differs from the mean age of married women. Suppose that 10 married couples are selected at random and that the ages, in years, of the people chosen are as shown below.

Husband	59	21	33	78	70	33	68	32	54	52
Wife	53	22	36	74	64	35	67	28	41	44

- (a) Perform appropriate hypothesis testing procedure for this problem at  $\alpha = 0.01$ . **(critical value: 3.250)**
- (b) **construct 95% confidence interval. [Critical value: 2.262]**

**Question # 08:**

CLO-3

**[1+3+2+2+3+1+2 = 14]**

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

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

- Identify dependent and independent variables.
- Draw and interpret the scatter plot.
- Determine the coefficient of correlation.
- Compute the regression model of type  $\mu_{Y|x} = \beta_0 + \beta_1 x$
- Test the significance of correlation at  $\alpha = 0.10$  (critical value: 1.943)

**OR**

Using the above same data, estimate the following polynomial model:  $\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2$

- Compute the coefficient of determination.
- Interpret your findings.

**Question # 09:**

CLO-3

**[8+5 = 13]**

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.

<b>y</b>	210	206	260	244	271	285	270	265	234	241	258	230
<b><math>x_1</math></b>	20	21	24	25	24	26	25	25	24	25	25	23
<b><math>x_2</math></b>	23	21	24	25	24	26	25	25	24	25	25	23

- Estimate unknown parameter of multiple linear regression equation:  $\mu_{Y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$
- Determine the multiple coefficients of determination ( $R^2$ ).

**Question # 10:**

CLO-3

**[15]**

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.

Check whether all brands provide the same level of relief from headaches at  $\alpha = 0.05$ .

(Critical value: 2.84) (show all working steps with interpretation)

<b>Panadol</b>	<b>Ponstan</b>	<b>Panadol extra</b>	<b>Ponstant forte</b>	<b>Paracetamol</b>
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	3.1	1	4.2
3	9.1	7.2	4	7.6

**Some Formulae:**

$$b_1 = \frac{(\sum X_1 Y)(\sum X_2^2) - (\sum X_2 Y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2}$$

$$b_2 = \frac{(\sum X_2 Y)(\sum X_1^2) - (\sum X_1 Y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2}$$

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

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2$$

## z-table

$z$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.1
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.2
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.3
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.4
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.5
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0019	0.6
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0026	0.7
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0036	0.8
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0048	0.9
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0064	1.0
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0084	1.1
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	1.2
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	1.3
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	1.4
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	1.5
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	1.6
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	1.7
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	1.8
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	1.9
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	2.0
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	2.1
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	2.2
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	2.3
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	2.4
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	2.5
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	2.6
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	2.7
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	2.8
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	2.9
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	3.0
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	3.1
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	3.2
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	3.3
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	3.4

## SOLUTION MANUAL

### Question 2(a)

We can select  $x$  defective sets from 2, and  $3 - x$  good sets from 5 in  $\binom{2}{x} \binom{5}{3-x}$  ways. A random selection of 3 from 7 sets can be made in  $\binom{7}{3}$  ways. Therefore,

$$f(x) = \frac{\binom{2}{x} \binom{5}{3-x}}{\binom{7}{3}}, \quad x = 0, 1, 2.$$

In tabular form

$x$	0	1	2
$f(x)$	2/7	4/7	1/7

The c.d.f. of  $X$  is

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ 2/7, & \text{for } 0 \leq x < 1, \\ 6/7, & \text{for } 1 \leq x < 2, \\ 1, & \text{for } x \geq 2. \end{cases}$$

$$P(0 < X \leq 2) = P(X \leq 2) - P(X \leq 0) = 1 - 2/7 = 5/7.$$

### Q2(b)

$$E(X) = \sum_{x=0}^3 x f(x) = (0)(27/64) + (1)(27/64) + (2)(9/64) + (3)(1/64) = 3/4.$$

variance of a random variable  $X$  is

$$\sigma^2 = E(X^2) - \mu^2.$$

$$E(x^2) = 0 + \frac{27}{64} + 4 \left( \frac{9}{64} \right) + 9 \left( \frac{1}{64} \right) = \frac{9}{8}$$

$$\text{var}(x) = \sigma^2 = \frac{9}{8} - \left( \frac{3}{4} \right)^2 = \frac{9}{16}$$

### Question 3(a)

Solution. Standardize and use Table For a Normal( $\mu = 900$ ,  $\sigma = 200$ ) variable  $X$ ,

$$\begin{aligned} P\{600 < X < 1200\} &= P\left\{ \frac{600 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{1200 - \mu}{\sigma} \right\} \\ &= P\left\{ \frac{600 - 900}{200} < Z < \frac{1200 - 900}{200} \right\} = P\{-1.5 < Z < 1.5\} \\ &= \Phi(1.5) - \Phi(-1.5) = 0.9332 - 0.0668 = \underline{0.8664}. \end{aligned}$$

**Q3(b)** mean = 15, var = 16, SD =  $\sigma = 4$

$$P(x < 20) = P\left( \frac{x - \mu}{\sigma} < \frac{20 - 15}{4} \right) = P(Z < 1.25) = 0.8944$$

#### Question 4

Solution. The sample size is  $n = 18$ , the sample mean time is  $\bar{X} = 0.29$  sec, and the sample standard deviation is  $s = 0.074$ . The critical value of  $t$  distribution with  $n - 1 = 17$  degrees of freedom is  $t_{\alpha/2} = t_{0.005} = 2.898$ . Then, the 99% confidence interval for the mean time is

$$0.29 \pm (2.898) \frac{0.074}{\sqrt{18}} = 0.29 \pm 0.05 = \underline{[0.24; 0.34]}$$

$$H_0 : \mu = 0.2 \text{ vs } H_A : \mu \neq 0.2$$

at a significance level  $\alpha = 0.01$ , we have sample statistics  $n = 18$ ,  $\bar{X} = 0.29$  and  $s = 0.074$ . Compute the T-statistic,

$$t = \frac{\bar{X} - 0.2}{s/\sqrt{n}} = \frac{0.29 - 0.2}{0.074/\sqrt{18}} = 5.16.$$

The rejection region is  $\mathcal{R} = (-\infty, -2.11] \cup [2.11, \infty)$ , where we used T-distribution with  $18 - 1 = 17$  degrees of freedom and  $\alpha/2 = 0.025$  because of the two-sided alternative.

Since  $t \in \mathcal{R}$ , we reject the null hypothesis

and conclude that *there is a significant evidence of an unauthorized use of that account.*

#### Question 5

Solution. Effect of the CD writer is measured by the reduction of the mean battery life. We have  $n = 12$ ,  $\bar{X} = 4.8$ ,  $s_X = 1.6$  for users with a CD writer and  $m = 18$ ,  $\bar{Y} = 5.3$ ,  $s_Y = 1.4$  for users without it. The pooled standard deviation is

$$s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}} = \sqrt{\frac{(11)(1.6)^2 + (17)(1.4)^2}{28}} = 1.4818$$

(check: it has to be between  $s_X$  and  $s_Y$ ). The critical value is  $t_{0.025} = 2.048$  (use 28 d.f.). The 95% confidence interval for the difference between the mean battery lives is

$$(4.8 - 5.3) \pm (2.048)(1.4818) \sqrt{\frac{1}{18} + \frac{1}{12}} = -0.5 \pm 1.13 = \underline{[-1.63; 0.63]}. \quad \diamond$$

$$n = 12, \bar{X} = 4.8, s_X = 1.6; m = 18, \bar{Y} = 5.3, s_Y = 1.4; s_p = 1.4818.$$

Testing

$$H_0 : \mu_X = \mu_Y \text{ vs } H_A : \mu_X < \mu_Y$$

at  $\alpha = 0.05$ , we obtain

$$t = \frac{\bar{X} - \bar{Y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{4.8 - 5.3}{(1.4818) \sqrt{\frac{1}{18} + \frac{1}{12}}} = -0.9054.$$

The rejection region for this left-tail test is  $(-\infty, -z_\alpha] = (-\infty, -1.645]$ . Since  $t \notin \mathcal{R}$ , we accept  $H_0$  concluding that *there is no evidence that laptops with a CD writer have a shorter battery life.*  $\diamond$

Q6)

	Systems	N	Mean	Std. Deviation	Std. Error Mean
operativeTime	D	14	401.7857	78.31414	20.93033
	S	6	468.3333	38.16630	15.58133

#### Independent Samples Test

		t-test for Equality of Means					
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
							Lower Upper
operativeTime	Equal variances assumed	-1.962	18	.065	-66.54762	33.92602	-137.82354 4.72830
	Equal variances not assumed	-2.550	17.460	.020	-66.54762	26.09323	-121.48934 -11.60590

Q7)

#### Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Husband	50.0000	10	19.29882	6.10282
	Wife	46.4000	10	17.46870	5.52409

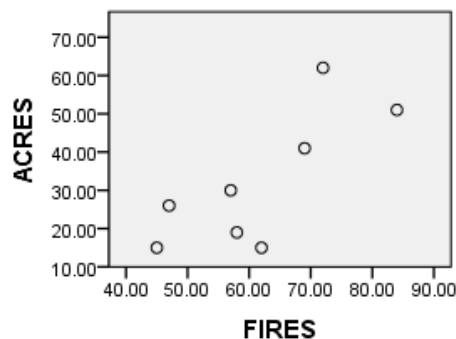
#### Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower Upper			
Pair 1	Husband - Wife	3.60000	4.97103	1.57198	.04394 7.15606	2.290	9	.048

Q8)

(a) Dependent Variable: ACRES

(b) Scatter Plot



The scatter plot indicates linear and positive relationship b/w the two variables.

(c) Correlation = 0.769

(d) Acres = -31.208 + 1.030 Fires + error

(e) Ho: The correlation is insignificant

H1: The Correlation is significant

**t-statistic** =  $r \cdot [\sqrt{n-2} / (1-r^2)] = 2.94$

**p-value** = 0.013 (reject Ho)

(f) Coeff of determination = 0.592

(g) There exist strong positive correlation b/w Fire & Acres. The correlation is significant. Furthermore the Coeff of determination suggest that the fitted regression model is explaining around 60% of the variation. Therefore the model is good.



**Q9)**  $Y = -77.380 + 8.013 X_1 + 5.527 X_2 + \text{error}$

$R = 0.854$   $R^2 = 0.730$

**Q10) ANOVA**

Ho: all means are equal

H1: at least one mean is different from others

$\alpha = 0.05$

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	63.293	3	21.098	5.934	.004
Within Groups	74.661	21	3.555		
Total	137.954	24			

**Reject Ho:** The brands have a significant impact in providing relief to headaches.