

No. of Printed Pages : 6 **MSTL-003(Set-I)**

**POST GRADUATE DIPLOMA IN
APPLIED STATISTICS (PGDAST)**

Term-End Practical Examination

December, 2024

MSTL-003 : BIOSTATISTICS LAB

Time : 3 Hours

Maximum Marks : 50

Note : (i) Attempt any **two** questions.

(ii) Solve the questions in Microsoft Excel.

(iii) Use of Formulae and Statistical Tables

Booklet for PGDAST is allowed.

(iv) Mention necessary steps, hypotheses,
interpretation, etc.

1. (a) A study was conducted on 32 patients admitted in a nephrology department of a hospital to investigate the Chronic Kidney Disease (CKD) and its associated risk factors. The data recorded on alcohol consumption and status of chronic kidney disease are given in the following table :

Alcohol Consumption	Chronic Kidney Disease	
	Yes	No
Yes	14	1
No	9	8

Test whether there is an association between alcohol consumption and kidney disease at 5% level of significance. 10

- (b) Suppose 30 patients suffering from brain tumor, included in a study, were assigned radiotherapy. To study the survival pattern, they all were followed upto death.

The data so obtained are as follows : 15

Patient ID	Survival Time	Patient ID	Survival Time	Patient ID	Survival Time
P01	35	P11	8	P21	27
P02	1	P12	27	P22	5
P03	2	P13	6	P23	24
P04	1	P14	16	P24	8
P05	16	P15	4	P25	4
P06	23	P16	7	P26	7
P07	30	P17	19	P27	31
P08	11	P18	9	P28	19
P09	7	P19	9	P29	12
P010	5	P20	10	P30	14

- (i) Estimate the survival function, CDF, PDF and hazard function.
- (ii) Plot survival and CDF curves.
2. In a study, different doses (in mg) of a medicine were prescribed to the patients who were experiencing migraine. The amount of dose (x_i), total number of patients who received medication (n_i) and number of patients whose migraine were cured after two hours of taking medications (y_i) were recorded. The data are given as follows :

Amount of Dose (x_i)	Total number of patients (n_i)	Number of cured patients (y_i)
5	240	60
10	180	48
15	120	40
20	200	80
25	260	104

For this data :

- (i) Fit a logistic regression model.

- (ii) Draw a scatter plot for the fitted logistic model.
 - (iii) Test the significance of the individual model coefficients β_0 and β_1 at 1% level of significance.
 - (iv) Determine the Nagelkerke pseudo R-squared.
3. A study was conducted to study the effect of walking on the blood sugar levels of the diabetic patients. 40 diabetic patients were randomly selected. The random blood sugar levels (in mg/dL) and average duration of walk (in minutes) of these patients are recorded. The data are as follows :

25

S. No.	Random blood sugar (mg/dL)	Duration of walk (in minutes)
1	302	43
2	328	47
3	390	51
4	364	49
5	407	33
6	374	50

7	390	51
8	307	45
9	264	38
10	382	51
11	407	53
12	328	47
13	302	43
14	441	56
15	325	46
16	307	45
17	462	57
18	300	43
19	374	50
20	470	58
21	447	56
22	347	48
23	364	49
24	418	54
25	308	45
26	278	39
27	297	41
28	382	51

29	247	38
30	335	48
31	308	45
32	428	55
33	335	48
34	300	43
35	390	52
36	418	54
37	347	48
38	390	52
39	325	46
40	287	40

- (i) Prepare a scatter plot to get an idea about the relationship among the variables.
- (ii) Fit a linear regression model and perform its related analysis at 1% level of significance.
- (iii) Does the fitted regression model satisfy the linearity and normality assumptions ?

× × × × × × ×