No. of Printed Pages : 5

MSTE–004

POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

June, 2025

MSTE-004 : BIOSTATISTICS—II

Time : 3 Hours *Maximum Marks : 50*

Note : *(i) Question No. **1** is compulsory.*

*(ii) Attempt any **four** questions from the remaining Question Nos. **2** to **7**.*

(iii) Use of scientific calculator (non-programmable) is allowed.

(iv) Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.

(v) Symbols have their usual meanings.

1. State whether the following statements are True *or* False. Give reasons in support of your answers : 5×2=10

   (a) Both samples should be independent and random to apply McNemar's test.

(b) For a regression model Y = 2.5 + 0.15X, 95% confidence interval for the slope is (0.0386, 0.2614) when n = 12 and SE $(\hat{\beta}_1) = 0.05$.

(c) If full and null log-likelihood functions of a logistic model are – 60 and – 62.5, the McFadden pseudo R^2 will be 0.04.

(d) The log-rank test is applied to compare the hazard functions.

(e) If relative risk for a clinical trial is 35%, the relative risk reduction will be 65%.

2. (a) The following data show the observations of two doctors who examined 350 patients :

Doctor I	Doctor II	Frequency
Y	Y	94
Y	Y	22
Y	N	7
Y	N	88
N	Y	37
N	Y	24
N	N	4
N	N	74

where Y represents presence and N represents absence of a disease. Calculate the value of Kappa statistics and its 95% confidence interval. 8

(b) Differentiate between the sensitivity and specificity of a diagnostic test. 2

3. The Systolic Blood Pressure (SBP), age (in years) and weight (in kgs) of 12 persons are given in the following table : 10

S. No.	SBP	Age	Weight
1	120	34	70
2	115	34	50
3	100	29	70
4	145	38	90
5	135	36	63
6	160	42	84
7	170	40	60
8	140	37	80
9	160	38	81
10	185	41	78
11	200	42	66
12	170	40	72

(i) Fit the multiple regression model of SBP on age and weight.

(ii) Compute the predicted values of SBP and residuals (r_i) for all the given observations.

(iii) Verify that $\sum_{i=1}^{12} r_i = 0$.

4. For the following case-control study of lung cancer and smoking status, fit a logistic model :

	Lung Cancer (Case)	No. Lung Cancer (Control)
Smoker	452	215
Non-smoker	548	785
Total	1000	1000

Also obtain the log-odds and odds ratio with the fitted logistic model. 10

5. The survival data of a randomized controlled trial corresponding to two treatment groups are given as follows :

	Survival time (in months)
Treatment 1	3, 4, 4^+, 6, 7, 9, 9^+, 11, 12, 14, 16, 18, 18, 20^+, 20^+
Treatment 2	5, 7, 8, 10^+, 12, 14, 14^+, 17, 19, 20, 20^+, 21, 22^+

Test whether there is a significant difference between the survival times of both treatments at 5% level of significance. 10

6. (a) Describe survival function with example. 5

   (b) Differentiate between Chi-square test and McNemar's test. 5

7. (a) Differentiate between coefficient of determination and adjusted coefficient of determination. 5

   (b) Explain the following in brief : 5

   (i) Probit model

   (ii) Complementary log-log model

× × × × ×