LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – **STATISTICS**

SECOND SEMESTER – **APRIL 2023**

PST2MC03 – CATEGORICAL DATA ANALYSIS

Date: 06-05-2023 Dept. No. Max. : 100 Marks

Time: 01:00 PM - 04:00 PM

SECTION A – K1 (CO1)

Answer ALL the questions **(5 x 1 = 5)**

1. **Definitions**

a) Relative risk.

b) Logits.

c) Deviance.

d) Log-linear model.

e) Matched pairs.

SECTION A – K2 (CO1)

Answer ALL the questions **(5 x 1 = 5)**

2. **Fill in the blanks**

a) _____ quantifies the strength of association between two events.

b) Systematic component specifies the _____ variables in the model.

c) In multinomial logistic regression the response variable is _____.

d) All variables in a loglinear model are essentially _____.

e) _____ model does not require marginal homogeneity.

SECTION B – K3 (CO2)

Answer any THREE of the following **(3 x 10 = 30)**

3. Explain different scales of measurements in detail.

4. a) Write short notes on partial tables.
b) State any two real-life applications of any categorical regression model. (5+5)

5. a) Define OR and give its applications
b) Distinguish between marginal and conditional odds ratio. (4+6)

6. Estimate the main effects for the following through log linear model:

VAR 2 / VAR 1	1	2	3	TOTAL
1	20	56	24	100
2	8	28	14	50
3	2	16	2	20
TOTAL	30	100	40	170

7. Explain Kappa measure of agreement.

SECTION C – K4 (CO3)

Answer any TWO of the following **(2 x 12.5 = 25)**

8. Define contingency table and explain its types with examples.

9.	An investigator randomly assigned 99 patients with stable congestive heart failure (CHF) to an exercise program (n=50) or no exercise (n=49) and followed patients twice a week for one year. The outcome of interest was all-cause mortality. Those assigned to the treatment group exercised 3 times a week for 8 weeks, then twice a week for 1 year. Exercise training was associated with lower mortality (9 versus 20) for those with training versus those without. Exercised Dead Alive Yes 9 41 No 20 29 Compute a valid measure of association and its 95% confidence interval.
10.	Explain the backward elimination procedure in logistic regression.
11.	Explain the generalized linear model for binary and count data.
	SECTION D – K5 (CO4)
	Answer any ONE of the following **(1 x 15 = 15)**
12.	Brief about partitioned Chi-Square analysis.
13.	Derive the log-linear model of independence for two-way contingency table.
	SECTION E – K6 (CO5)
	Answer any ONE of the following **(1 x 20 = 20)**
14.	(a) Distinguish between logit and probit models. (b) Explain the procedure of logistic regression. (5+15)
15.	(a) Discuss the connection between Logistic and Log-linear models. (b) Derive the logistic model (X+Z) from the log linear model (XY,XZ,YZ) when Y is a binary response variable. (8+12)

$$$$$$$