LOYOLA COLLEGE (AUTONOMOUS) CHENNAI – 600 034



M.Sc. DEGREE EXAMINATION - STATISTICS

SECOND SEMESTER - APRIL 2025



PST 2504 - CATEGORICAL DATA ANALYSIS

Date: 02-05-2025	Dept. No.	Max. : 100 Marks
Time: 01:00 PM - 04:00 PM		

SECTION A

Answer ANY FOUR of the following

 $4 \times 10 = 40 \text{ Marks}$

- 1. Explain the various data measurement in statistics with examples.
- 2. A researcher is studying the number of customer complaints received at a call centre. Based on past data, the average number of complaints per day is believed to be $\lambda = 4$. However, after implementing a new customer service policy, the researcher wants to test whether the average number of complaints has increased. A random sample of $\mathbf{n} = \mathbf{10}$ days is taken, and the number of complaints recorded per day is as follows: $X_1=6$, $X_2=5$, $X_3=7$, $X_4=4$, $X_5=5$, $X_6=6$, $X_7=3$, $X_8=5$, $X_9=6$, $X_{10}=7$. Use the **Score Test** to determine whether there is significant evidence to reject H_0 at $\alpha = \mathbf{0.05}$ significance level.
- 3. Explain the following with examples:
 - a). Case-Control study
- b). Cohort Study
- c). Clinical Trails
- d) Cross-Sectional Study.
- 4. Describe the components of GLMs (Random component, systematic component and link function).
- 5. A clinical trial records the response (Success = 1, Failure = 0) of 100 patients before and after receiving a new drug. The results are:

Response Before	Response After	Count
Recovery	Recovery	50
Recovery	No recovery	15
No recovery	Recovery	25
No recovery	No recovery	10

Estimate the **marginal proportion** of success before and after treatment.

- 6. Explain the statistical model of multinomial and ordinal logistic regression model.
- 7. What are the measures used for the goodness of fit in logistic regression model. Explain in detail.
- 8. Two judges evaluate 100 research papers as either "Accept" or "Reject." Their ratings are given in the table below:

	Judge B: Accept	Judge B: Reject	Total
Judge A: Accept	30	10	40
Judge A: Reject	15	45	60
Total	45	55	100

- (a) Calculate the **Observed Agreement**.
- (b) Calculate the **Expected Agreement** under chance agreement.
- (c) Compute Cohen's Kappa (κ) and interpret the result.

SECTION B

Answer ANY THREE of the following

 $3 \times 20 = 60 \text{ Marks}$

- 9. a) Derive the Wald test statistic and confidence interval for Poisson distribution.
 - b) A supermarket manager wants to determine whether the average number of customers arriving at the checkout counter per hour has increased from the historical average λ =20 customers per hour. In a random sample of n=40 hours, the totall number of customers observed is Y=900. Test the null hypothesis H₀: λ =20 at a significane level of α =0.05, and calculate a 95% confidence interval for λ .
- 10. Derive the test statistic for Likelihood Ratio test for binomial distribution and hence estimate the confidence interval.
- 11. Compute Concordant pairs and Discordant pairs for the following cross-classification table on job satisfaction (Y) by income (X).

_	Job Satisfaction						
Income (\$)	Extremely Dissatisfied	Little Dissatisfied	Moderately Satisfied	Very Satisfied			
<15000	1	3	10	6			
15000 - 25000	2	3	10	7			
25000 - 40000	1	6	14	12			
>40000	0	1	9	11			

- 12. a) Estimate the likelihood ratio test for multinomial distribution
 - b) Estimate M.L.E of parameters of multinomial distribution.
- 13. Explain the following:
 - a). Gamma test and it's properties b) Yule's Q test and it's properties.
 - c) Estimate the sensitivity and specificity for the following data using KS Statistics:

Predicted Probability	0.42	0.56	0.37	0.99	0.72	0.76	0.12	0.85	0.45	0.56	0.32	0.19	0.95
Y	0	1	0	1	1	1	0	0	0	1	0	0	1