

LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034**M.Sc. DEGREE EXAMINATION – STATISTICS****SECOND SEMESTER – APRIL 2016****ST 2817 – CATEGORICAL DATA ANALYSIS**

Date: 25-04-2016

Dept. No.

Max. : 100 Marks

Time: 01:00-04:00

PART - A**Answer ALL the questions****(10x2=20 Marks)**

1. Provide the general layout of a 2x2 contingency table and define joint and marginal probability.
2. The following 2x2 contingency table is from a report on Aspirin use and Myocardial Infarction. Determine odds ratio and provide your interpretation.

		Myocardial Infarction		Total
		Yes	No	
Group	Placebo	189	10845	11034
	Aspirin	104	10933	11037
Total		293	21778	22071

3. Define Likelihood ratio and Relative risk for a 2x2 contingency table.
4. Define Lambda and Cramer's V coefficient.
5. Define i) Somer's D and ii) Goodman and Kruskal Gamma.
6. Define zero inflation and over dispersion.
7. Define concordance and discordance in binary logistic regression.
8. Define deviance residual.
9. State any two drawbacks of using linear regression model in the case of dichotomous dependent variable.
10. Define Mixed Logit Model.

PART-B**Answer any FIVE the questions****(5x8=40 Marks)**

11. Explain Simpson's Paradox with an example.
12. Discuss the steps involved in construction of gains chart and provide its uses.
13. i) Discuss the need for Poisson regression model and its use. (2)
- ii) State the assumptions of Poisson Regression Model. (2)
- iii) Provide the tests for determining Statistical significance of regression coefficients. (4)
14. Discuss i) Complete separation ii) Quasi complete separation iii) overlap and iv) possible reasons for Quasi complete separation. (2+2+2+2)
15. Explain Multinomial Logistic Regression in Detail.
16. For a Negative Binomial Regression show that $E(Y_i) = \mu_i$ and $Var(Y_i) = \mu_i + (\mu_i^2 / \psi)$.
17. Discuss the following goodness of model fit measures i) Pearson Chi-square ii) Deviance iii) Akaike Information Criteria and iv) Bayesian Schwartz Information Criteria.
18. Explain Probit model in detail.

PART-C

Answer any TWO questions

(2x20=40 Marks)

19. i) Explain Binary Logistic regression in detail. (6 Marks)
ii) Discuss any one method of determining optimal cut value. (6 Marks)
iii) Explain ROC and AUC and state its use. (4 Marks)
iv) Interpretation of regression coefficient in a Binary Logistic Model. (4 Marks)
20. i) Explain Cumulative Logistic Regression in Detail. (8)
ii) Explain Log-Linear Analysis for analyzing dependency in contingency table. (12)
21. i) Explain Poisson regression mode in detail and state any two applications. (10)
ii) State the use of a) Adjacent category model b) Conditional Logit Model c) Mixed logit model
d) Negative Binomial regression e) Poisson regression. (2+2+2+2+2marks)
22. i) Define Discrete choice model and state any two uses. (4 Marks)
ii) Explain GLM and its components. (6 Marks)
ii) Discuss the methods of model validation. (8 Marks)
iv) What is the need for model validation? (2 Marks)

\$\$\$\$\$\$