School of Information Technology

Rajiv Gandhi Proudyogiki Vishwavidyalaya (UTD)

Endsem Examination CD/CB/AL 202 (**STATISTICAL METHODS**)

TIME: 3:00 Hrs TOTAL MARKS: 70

Note: Attempt any 5 questions. Each question carries equal marks.

<table>
<tr><th>S.No.</th><th>Questions</th><th>Marks</th><th>CO</th></tr>
<tr><td>Q-1</td><td>
a. Show that the line of fit of the following data is given by $y = 0.7x + 11.28$.

<table>
<tr><td>x</td><td>0</td><td>5</td><td>10</td><td>15</td><td>20</td><td>25</td></tr>
<tr><td>y</td><td>12</td><td>15</td><td>17</td><td>22</td><td>24</td><td>30</td></tr>
</table>

b. Define correlation coeffient. Calculate the coefficient of correlation between the values of x and y.

<table>
<tr><td>x</td><td>78</td><td>89</td><td>97</td><td>69</td><td>59</td><td>79</td><td>68</td><td>61</td></tr>
<tr><td>y</td><td>125</td><td>137</td><td>156</td><td>112</td><td>107</td><td>136</td><td>123</td><td>108</td></tr>
</table>
</td><td>7+7</td><td>CO1</td></tr>
<tr><td>Q-2</td><td>
a. Explain analysis of variance (ANOVA) one way and two way classification with tables and examples.

b. The equation of two regression lines obtained in a corelation analysis in a 60 observations are $5x = 6y + 24$ and $1000y = 768x - 3708$. What is the correlation coefficient and what is the ratio of variances of x and y.
</td><td>7+7</td><td>CO1</td></tr>
<tr><td>Q-3</td><td>
a. State and prove Neyman Pearson Lemma.

b. Define maximum likelihood estimator with examples. Write properties of Maximum Likelihood Estimator.
</td><td>7+7</td><td>CO2</td></tr>
<tr><td>Q-4</td><td>
a. Define estimation. Explain the properties of good estimators.

b. Suppose a random sample of size n is taken from Poisson population with probability function $f(x, \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$. Show that the most powerful critical region of size not exceeding α for testing the hypothesis $H_0 : \lambda = \lambda_0$ against $H_1 : \lambda = \lambda_1$ is of the form $\bar{x} \leq a_\alpha$ if $\lambda_0 > \lambda_1$ and $\bar{x} \leq b_\alpha$ if $\lambda_0 > \lambda_1$ where $\bar{x}$ is the sample mean and a_α and b_α are constants.
</td><td>7+7</td><td>CO2</td></tr>
</table>

School of Information Technology

Rajiv Gandhi Proudyogiki Vishwavidyalaya (UTD)

Endsem Examination CD/CB/AL 202 (STATISTICAL METHODS)

S.No.	Questions	Marks	CO
Q-5	a. Comparison of parametric and non-parametric inference with examples. Define sign test. b. Define Kendall's test. Find Kendall's rank correlation coefficient for the given data. x: 2, 7, 1, 5, 8, 10 y: 4, 5, 6, 8, 10, 9	7+7	CO3
Q-6	a. A medication is given to 11 patients who suffer from BP issues, their BP are recorded for the analysis. Use Willcoxon signed rank test to test the hypothesis that medication has no effect on BP. Before Medication: 112, 113, 118, 120, 119, 113, 110, 122, 126, 115, 119 After Medication: 116, 126, 117, 125, 126, 111, 113, 117, 126, 112, 129 b. Define Run test. A foreman for a construction company records injuries reported by workers, durinmg his shift. The following sequence show whether any injury were reported during each month in recent year. 'I' represents injuries which occur at least once in the month and 'N' represent no injury. At $\alpha = 0.05$, can we conclude that the occurence of injuries in each month is not random. II NN I NII NNN	7+7	CO3
Q-7	a. Explain ARIMA model. What is it used for? b. Define time series analysis and forecasting with examples and application.	7+7	CO4
Q-8	a. Explain nay 5 builtin functions in R programming. Define control structures with suitable examples. b. Explain with an example how to read and write data in R program. How `Im()` function is used to implement linear regression model in R programming?	7+7	CO5

Q-5 (b) data:

x	2	7	1	5	8	10
y	4	5	6	8	10	9

Q-6 (a) data:

Before Medication	112	113	118	120	119	113	110	122	126	115	119
After Medication	116	126	117	125	126	111	113	117	126	112	129