I/E/TX

Reg. No: 23BBS0025

VIT®
Vellore Institute of Technology

Final Assessment Test – May 2024

Course: **MAT2005 - Data Science and Statistical Modelling**

Class NBR(s): **3856 / 3866** Slot: **B2**

Time: **Three Hours** Max. Marks: **100**

- **KEEPING MOBILE PHONE/ELECTRONIC DEVICES EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE**
- **DON'T WRITE ANYTHING ON THE QUESTION PAPER**

General Instructions:

t, f, chi-squared, sign test, ıns test, Mann Whitney, Wilcoxon signed rank test tables are permitted

Answer any TEN Questions

(10 X 10 = 100 Marks)

1. Following data gives the cost and sales of ten products in Rupees: **[10]**

Cost	39	65	62	90	82	75	25	98	36	78
Sales	47	53	58	86	62	68	60	91	51	84

Find Karl Pearson coefficient of correlation for this data and interpret it.

2. A sample with sample size 7 for three variables y, x_1, x_2 is such that **[10]**

$\sum x_1 = 28,\ \sum x_2 = 101,\ \sum y = 195,\ \sum x_1 x_2 = 454,\ \sum x_1 y = 940,\ \sum x_2 y = 3100,$
$\sum x_1^2 = 140,\ \sum x_2^2 = 1549,\ \sum y^2 = 6375.$

Find the regression line for y in terms of x_1, x_2.

3. (i) What is a point-estimate? What are some criterions for good estimates? **[10]**

(ii) Consider variables $x_1, x_2, \ldots x_n$ be random variables drawn from normal distribution. Is the estimator $T = \sum x_i/(n+1)$ a unbiased estimator? Explain your answer.

4. Suppose $x_1, x_2, \ldots x_n$ be random variables with density function, **[10]**
$f(x, \sigma) = (1/(2\sigma))\exp(-|x_i|/\sigma)$. Find the maximum likelihood estimate for σ.

5. Explain the concept of sufficient statistics with the help of an example. **[10]**

6. Suppose that $x_1, x_2, \ldots x_n$ form a random sample from a Poisson distribution with density function $f(x, \theta) = \theta^x \exp(-\theta)/(x_i!)$. Show that $T = \sum x_i$ is a sufficient statistic for θ. **[10]**

7. Two independent samples of 8 and 7 items gave the following values: **[10]**

Sample A	9	11	13	11	15	9	12	14
Sample B	10	12	10	14	9	8	10	

Examine whether there is a significant difference in the means of the two samples at 5% significance level.

8. A sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 secured third class, 90 obtained second class and 20 got first class. Are these figures commensurate with the general result which is in the ratio of 4:3:2:1 for the various categories respectively? Use level of significance of 5%. **[10]**

9. Three samples each of size 5 were drawn from three uncorrelated normal populations. Test using analysis of variance whether the population means are equal at 5% level. [10]

Sample 1	10	12	9	16	13
Sample 2	9	7	12	11	11
Sample 3	14	11	15	14	16

10. The following data gives heights of 15 students in a class in centimetres: [10]

165, 172, 186,176, 170, 160, 156, 180, 174,160, 170, 178, 175,182,169.

Using sign test can we say that the median height of the students is 170 cm at 10% level?

11. The scores of 10 people in an IQ test were measured before and after training: [10]

Before	40	50	45	42	42	35	48	38	45	40
After	43	47	50	48	43	41	45	42	48	45

Using Wilcoxon signed rank test can we say that there was any effect of the training on the scores of the students at 5% level.

12. The marks of 8 students in Maths and Science are as given in table below: [10]

Maths	70	75	80	75	65	60	90	75
Science	85	80	90	95	70	75	95	90

Calculate Spearman rank correlation coefficient and interpret it.

⇔⇔⇔ I/E/TX⇔⇔⇔