VIT
Vellore Institute of Technology

**FALL Semester 2022-2023**
**Continuous Assessment Test – II**
**Programme Name: MCA**
**Course Code & Name:**
**MAT5007 –Applied Statistical Methods**

**Number:5832** **Slot: D2** **Exam Duration: 90 minutes** **Maximum Marks: 50**

**(ANSWER ALL QUESTIONS -5x10=50 Marks)**

1. The incident of defective in 200 samples of 6 is shown in the following table

| No of defective per sample | 0 | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|---|
| No of samples | 36 | 70 | 61 | 25 | 7 | 1 | 0 | 200 |

Assuming these results follows a binomial distribution, compute the theoretical binomial probabilities and frequencies.

2. Fit a Poisson distribution for the following data and also test for the goodness of fit.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| f | 314 | 335 | 204 | 86 | 29 | 9 | 3 | 0 |

3. The average ticket price for a major league baseball game was \$11.98 in 1998 (USA Today, November 11, 1998). Adding the cost of food, parking, and souvenirs, the average cost for a family of four to attend a game was approximately \$110.00. Assume the normal distribution applies and that the standard deviation is \$20.00. What is the probability the cost will exceed \$100.00?

   a) What is the probability a family will spend \$90.00 or less?
   b) What is the probability a family will spend between \$80.00 and \$130.00?

4. The Education testing service conducted a study to investigate difference between the scores of male and female students on the scholastic aptitude test. The study identified a random sample of 562 female and 852 male students who had achieved the same high score on the mathematics portion of the test. That is, the female and male students were viewed as having similarly high abilities in mathematics. The SAT verbal scores for two samples are as follows. Female Students $\overline{x_1}$=547, S1=83 Male Students $\overline{x_2}$ =525, S2=78.Do the data support the conclusion that given a population of female students and a population of male students with similarly high mathematical abilities, the female students will have a significantly higher verbal ability? Test at .01 level of significance. What is your conclusion?

5 A sample of 1545 men and a samples of 1691 women were used to compare the amount of housework done by women and men in dual-earner marriages. The study showed that 67.5% of the men felt the division of housework was fair and 60.8% of the women felt the division of housework was fair. Is the proportion of men who felt the division of housework was fair greater than the proportion of women who felt the division of housework was fair? Support your conclusion with a statistical test using a .05 level of significance.

$n_1 = 1545$
$n_2 = 1043$

VIT
Vellore Institute of Technology
(Deemed to be University under Section 3 of UGC Act, 1956)

# SCHOOL OF ADVANCED SCIENCES

| Name of Examination | | Continuous Assessment Test – II, Fall Semester 2022-23 | | |
|---|---|---|---|---|
| **Slot: A2** | | Course Mode: Class Based Learning | **Class Number (s): VL2022230105761** | |
| Course Code: | MAT5019 | Course Title: | Business Statistics with R | |
| Emp. No.: | 16354 | Faculty Name: | Dr. Rajesh Moharana | School: SAS |
| Contact No.: | 8144862662 | Email: | rajesh.moharana@vit.ac.in | |

*General Instructions (if any):* **Non-programmable calculator and hand-written notebooks are permitted.**

| Q. No. | Question Text | Marks | Module No. | CO/ BL |
|---|---|---|---|---|
| **Answer All the Questions** | | **Total Marks: 5 x 10 Marks = 50 Marks** | | |
| 1. | Two dimensional random variables $X$ and $Y$ have the joint probability function $P(X = x; Y = y) = k(x^2 + y)$; for $x = 0, 1, 2, 3$ and $y = 0, 1$.<br>*(i)*. Evaluate "$k$".<br>*(ii)*. Find all the marginal distributions of $X$ and $Y$.<br>*(iii)*. Find the conditional distribution of $Y \mid X = 2$. | 10 | 3 | CO3 BL3 |
| 2. | The joint probability density function of two random variables $X$ and $Y$ is given by:<br>$f(x, y) = 6(1 - x - y);\ x > 0, y > 0, x + y < 1.$<br>Test whether $X$ and $Y$ are independent. For the above joint distribution, find the conditional density of $X$ given $Y = y$. | 10 | 3 | CO3 BL3 |
| 3. | $(i)$ Assuming that the typing mistakes per page committed by a typist follow a Binomial distribution, find the theoretical frequencies for the following distribution of typing mistakes.<br>No. of mistakes Per page: 0, 1, 2, 3, 4, 5, 6<br>No. of pages: 13, 25, 52, 38, 32, 16, 4<br>$(ii)$ A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused. | 10 | 4 | CO4 BL3 |
| 4. | Two numbers within brackets denote the marks of the students in Mathematics and Statistics subjects. The marks of ten students are as follows: (70,62), (64,68), (75,68), (50,45), (64,81), (80,60), (75,68), (40,48), (55,50), (64,70). Calculate the rank correlation coefficient of this data and comment on the types of correlation between Mathematics and Statistics. | 10 | 5 | CO2 BL5 |
| 5. | Marks obtained by 12 students in the college test ($y$) and the university test ($x$) are as follows:<br>$x$: 41, 45, 50, 68, 47, 77, 90, 100, 80, 100, 40, 43<br>$y$: 60, 63, 60, 48, 85, 56, 53, 91, 74, 98, 65, 43<br>Estimate the mark of a student could have obtained in the university test if he/she obtained 100 in the college test but was ill at the time of the university test? | 10 | 5 | CO2 BL5 |

VIT
Vellore Institute of Technology

# Final Assessment Test - Jan / Feb 2023

Course: MAT5007 - Applied Statistical Methods
Class NBR(s): 5820 / 5832 / 5854 Slot: D2
Time: Three Hours Max. Marks: 100

**KEEPING MOBILE PHONE/SMART WATCH, EVEN IN 'OFF' POSITION, IS TREATED AS EXAM MALPRACTICE**

**General Instruction : Scientific calculator and Statistical tables are allowed**

**Answer any TEN Questions**

**(10 X 10 = 100 Marks)**

1. Calculate the arithmetic mean and the median of the frequency distribution given below. Hence calculate the mode using the empirical relation between the three.

| Class limits | 130-134 | 135-139 | 140-144 | 145-149 | 150-154 | 155-159 | 160-164 |
|---|---|---|---|---|---|---|---|
| Frequency | 5 | 15 | 28 | 24 | 17 | 10 | 1 |

2. From the data given below, find the Standard deviation

| Ages | 20-25 | 25-30 | 30-35 | 35-40 | 40-45 | 45-50 | 50-55 | 55-60 |
|---|---|---|---|---|---|---|---|---|
| No. of Persons | 50 | 70 | 80 | 80 | 150 | 120 | 70 | 50 |

3. The following table gives the distribution of monthly income of 500 workers in a factory

| Monthly income | Below Rs 100 | 100-150 | 150-200 | 200-250 | 250-300 | 300 and above |
|---|---|---|---|---|---|---|
| No. of Workers | 10 | 25 | 145 | 220 | 70 | 30 |

Obtain Bowley's coefficient of skewness.

4. Calculation Pearson's coefficient of correlation from the following data: Take 65 and 70 as the assumed average of the variates X and Y respectively

| X | 45 | 55 | 56 | 58 | 60 | 65 | 68 | 70 | 75 | 80 | 85 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Y | 56 | 50 | 48 | 60 | 62 | 64 | 65 | 70 | 74 | 82 | 90 |

5. From the following data of the age of husband and the age of wife, from two regression lines and calculate the husband's age when the wife's age is 16.

| Husband's age | 36 | 23 | 27 | 28 | 28 | 29 | 30 | 31 | 33 | 35 |
|---|---|---|---|---|---|---|---|---|---|---|
| Wife's age | 29 | 18 | 20 | 22 | 27 | 21 | 29 | 27 | 29 | 28 |

6. The distribution of typing mistakes committed by a typist is given below. Assuming a Poisson model, find the expected frequencies.

| Mistakes per page | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| No. of pages | 142 | 156 | 69 | 27 | 5 | 1 |

7. The daily wages of 1000 workmen are normally distributed around a mean of Rs 70 and with a standard deviation of Rs 5. Estimate the number of workers whose daily wages will be

   i) Between Rs 70 and 72
   ii) Between Rs 69 and 72
   iii) more than Rs 75
   iv) Less than Rs 63
   v) more than Rs 80

8. A person buys 100 electric tubes of each of two well-known makes, taken at random from stock for testing purpose. He finds that make A has a mean life of 1300 hours with a standard deviation of 82 hours, and make B has a mean life of 1248 hours with a standard deviation of 93 hours. Discuss the significance of these results. Which make of electric tubes should the person buy?

9. The subject under investigation is the measure of dependence of Tamil on words of Sanskrit origin. One newspaper article reporting the proceedings of the constituent assembly contained 2025 words of which 729 words were declared by literary critic to be of Sanskrit origin. A second article by the same author describing atomic research contained 1600 words of which 640 words were declared by the same critic to be of Sanskrit origin. Assuming that simple sampling conditions hold, estimate the limits for the proportion of Sanskrit words in the writer's vocabulary and examine whether there is any significant difference in the dependence of this writer on words of Sanskrit origin in writing the two articles. Take $\alpha = 0.05$.

10. The following table gives for a sample of married women, the level of education and marriage adjustment score:

| | | marriage adjustment score | | | |
|---|---|---|---|---|---|
| | | Very low | Low | High | Very high |
| Level of Education | College | 24 | 97 | 62 | 58 |
| | High school | 22 | 28 | 30 | 41 |
| | Middle school | 32 | 10 | 11 | 20 |

Can you conclude from the above, the higher the level of education, the greater is the degree of adjustment in marriage?

11. The random samples were drawn from two normal populations and the following results were obtained.
Sample 1: 16,17,18,19,20,21,22,24,26,27
Sample 2: 19,22,23,25,26,28,29,30,31,32,35,36
Obtain the estimates of the means of populations and test whether two populations have the same means.

12. Four experimenters determine the moisture content of samples of a powder, each man taking a random sample from each of the six consignments. These assessments are given in the following table.

| Observer | consignment | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 9 | 10 | 9 | 10 | 11 | 11 |
| 2 | 12 | 11 | 9 | 11 | 10 | 10 |
| 3 | 11 | 10 | 10 | 12 | 11 | 10 |
| 4 | 12 | 13 | 11 | 14 | 12 | 10 |

Perform an analysis of variance on these data and discuss whether there is any significant difference between consignments or between observers.
If there is significant difference between the level of any of the factors of variation, then determine the corresponding pairs of sample means which differ significantly.
Use $\alpha = 0.05$

⇔⇔⇔

VIT
Vellore Institute of Technology

# Final Assessment Test – Jan / Feb 2023

Course: MAT1012 - Statistical Applications

Class NBR(s): 5711 / 5715 Slot: E2

Time: Three Hours Max. Marks: 100

**KEEPING MOBILE PHONE/SMART WATCH, EVEN IN 'OFF' POSITION, IS TREATED AS EXAM MALPRACTICE**

**Answer any TEN Questions**

**(10 X 10 = 100 Marks)**

1. Draw the histogram and frequency polygon for the following data.

| Weight(in kg) | 50-55 | 55-60 | 60-65 | 65-70 | 70-75 | 75-80 | 80-85 |
|---|---|---|---|---|---|---|---|
| No. of persons | 15 | 8 | 12 | 17 | 9 | 10 | 6 |

2. The following table shows the time taken (in minutes) by 100 students to travel to school on a particular day.

| Time | 0-5 | 5-10 | 10-15 | 15-20 | 20-25 |
|---|---|---|---|---|---|
| No. of students | 5 | 25 | 40 | 17 | 13 |

Construct the frequency curve and give your valid conclusions.

3. The marks of 36 students in an entrance test are given below:

| Grades(less than) | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 | 80-90 | 90-100 |
|---|---|---|---|---|---|---|---|
| No. of students | 3 | 4 | 6 | 10 | 6 | 4 | 3 |

Obtain Median and Mode of the above data.

4. Calculate: (i) Quartile deviation and (ii) Mean deviation from mean, for the following data

| Marks | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 |
|---|---|---|---|---|---|---|---|
| No. of students | 6 | 5 | 8 | 15 | 7 | 6 | 3 |

5. Compute the spearman's rank correlation coefficient for the data recorder below and give conclusions:

| Estimated Cost | 300 | 450 | 800 | 250 | 500 | 975 | 475 | 400 |
|---|---|---|---|---|---|---|---|---|
| Actual Cost | 273 | 486 | 734 | 297 | 631 | 872 | 396 | 457 |

6. For the data given below:

| X | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Y | 9 | 8 | 10 | 12 | 11 | 13 | 14 |

Obtain: (i) The regression lines of Y on X and X on Y (ii) Regression coefficients (iii) Residuals for Y (iv) Coefficient of determination.

7. Suppose that a company publishes two magazines, M1 and M2. Based on their record of subscriptions in a suburb they find that sixty percent of the households subscribe only for M1, forty five percent subscribe only for M2, while only twenty percent subscribe for both M1 and M2. If a household is picked at random from this suburb, the magazines. Publishers would like to address the following questions: What is the probability that the randomly selected household is a subscriber for (i) at least one of the magazines M1, M2, (ii) none of those magazines M1, M2, (iii) magazine M2 given that the same household subscribes for M1, and (iv) magazine M1 given that the same household subscribes for M2?

8. In an experiment of fruit setting in pointed gourd, three different methods (I, II and III) of pollination were used in 30 %, 40 %, and 30 % plots, respectively, knowing fully that the success rates of three types of pollination are 50 %, 85 %, and 90 %, respectively. What is the probability that a particular plot has been pollinated by the methods I, II and III?

9. The table below presents 9 subgroups (samples) of four measurements on inside diameter (ID) of a part processed in a turning machine. Set up **X-bar and R Charts** on this process and also construct control limits for mean and range chart.

| Sample No. | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|---|---|---|---|---|
| 1 | 5.00 | 5.01 | 4.98 | 5.00 |
| 2 | 5.01 | 4.98 | 5.00 | 5.00 |
| 3 | 5.02 | 5.01 | 5.00 | 5.00 |
| 4 | 5.00 | 5.00 | 5.00 | 5.00 |
| 5 | 4.98 | 4.98 | 5.01 | 4.99 |
| 6 | 5.02 | 4.99 | 5.00 | 4.98 |
| 7 | 4.99 | 4.99 | 4.98 | 4.98 |
| 8 | 5.00 | 5.01 | 5.02 | 5.00 |
| 9 | 4.98 | 5.00 | 5.01 | 4.98 |

$\bar{\bar{A}} + A_2\bar{R}$

$\bar{\bar{A}} - A_2\bar{R}$

$D_3\bar{R}$

$D_4\bar{R}$

*(Given: for n=4, $A_2$=0.729, $D_3$=0 and $D_4$=2.282).*

10. 20 samples each of size 10 were inspected. The number of defectives detected in each of them are given below:

| Sample No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| No. of defectives | 0 | 1 | 0 | 3 | 9 | 2 | 0 | 7 | 0 | 1 |
| Sample No. | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| No. of defectives | 1 | 0 | 0 | 3 | 1 | 0 | 0 | 2 | 1 | 0 |

Construct the number of defectives chart (***np chart***) and check whether the process is under control or not?

11. The following table gives the frequency distribution of scent and color of 200 roses in a garden. Test at 5% level of significance whether scent is independent of color of flower or not.

| Color / Scent | Pink | Red | Total |
|---|---|---|---|
| Intense | 35 | 75 | 110 |
| Light | 55 | 35 | 90 |
| Total | 90 | 110 | 200 |

(*Given:* $\chi^2_{0.05,1} = 3.841$)

12. Below are given the gain in weights (in lbs) of pigs fed on two diets A and B.

| | |
|---|---|
| Diet A | 25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25 |
| Diet B | 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22 |

Test at 5% level of significance, if the two diets differ significantly as regards their effect on increase in weight.

(*Given;* $t_{0.05}\ (12 + 15 - 2) = t_{0.05}(25) = 2.06$)