

Seat No.: _____

Enrolment No. _____

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA - SEMESTER-III • EXAMINATION – WINTER 2016

Subject Code: 630003

Date: 02/01/2017

Subject Name: Statistical Methods

Time: 10:30 am – 01:00 pm

Total Marks: 70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

Q.1 (a) Answer the following.

- (i) If the coefficient of variation is 40% and the mean is 70, then compute the variance. **02**
 - (ii) If A and B are independent events with $P(A) = 0.7$ and $P(B) = 0.8$, then compute $P(A \cup B)$. **02**
 - (iii) Assume that you have a binomial experiment with $p = 0.6$ and a sample size of 50. What is the value of variance of this distribution? **01**
 - (iv) What is the Z value for a 97.8% confidence interval estimation? **02**
- (b) (i)** The following data show the number of hours worked by 200 statistics students. **03**

Number of Hours	Frequency
0 - 9	40
10 - 19	50
20 - 29	70
30 - 39	40

Referring above data Find:

1. The number of students working 19 hours or less.
 2. The relative frequency of students working 9 hours or less.
 3. The cumulative relative frequency for the class of 10 – 19.
- (ii)** In SM exam, the probability of A is passing is $\frac{1}{2}$, probability of B passing is $\frac{1}{3}$ and probability of neither A nor B passing is $\frac{1}{4}$. Find the probability of A and B passing the exam. **04**

Q.2 (a) The following sample shows the marks of the students. **07**
Marks: 30, 44, 54, 50, 36, 42, 46, 42.

Compute mean, median, standard deviation, 75th percentile.

(b) The following data represent the length X (in cm.) of green plants and the weight Y (in gm.) of dry fiber for 8 plants: **07**

X	172	148	162	183	160	141	150	190
Y	6.4	2.3	3.5	4.7	4.1	2.9	2.8	6.6

- (i) Compute the covariance for the above data.
- (ii) Compute the correlation coefficient for the above data.

OR

(b) (i) State the % of items covered by (i) $Mean \pm \sigma$ (ii) $Mean \pm 2\sigma$ (iii) $Mean \pm 3\sigma$ **03**

- (ii) The probability distribution of a random variable X is given below.

04

X	-2	-1	0	1	2
$P(X)$	0.2	0.1	0.3	0.3	0.1

Find

- (i) Expectation of X , i.e., $E(X)$
- (ii) Variance of X , i.e., $V(X)$,
- (iii) $E(2X - 3)$
- (iv) $P(0 \leq X \leq 2)$

- Q.3 (a)** In a large university, 75% of students live in dormitories. A random sample of 5 students is selected. Use the binomial probability distribution to answer the following questions.

07

- (i) What is the probability that the sample contains exactly three students who live in the dormitories?
- (ii) What is the probability that the sample contains no students who live in the dormitories?
- (iii) What is the probability that the sample contains more than three students who do not live in the dormitories?
- (iv) What is the expected number of students (in the sample) who do not live in the dormitories?

- (b)** The life expectancy of computer terminals is normally distributed with a mean of 4 years and a standard deviation of 10 months.

07

- (i) What is the probability that a randomly selected terminal will last more than 5 years?
- (ii) What percentage of terminals will last between 2.5 and 4.5 years?
- (iii) If the manufacturer guarantees the terminals for 3 years (and will replace them if they malfunction), what percentage of terminals will be replaced?

OR

- Q.3 (a)** Consider the following distribution.

07

No. of calls	0	1	2	3	4	5	6	7 or more
Frequency	5	18	24	30	32	13	20	8

- (i) Compute the expected frequencies using Poisson distribution.
- (ii) Test whether or not the Poisson model is a valid model for the above data.

- (b)** The life time of a certain kind of batteries has a mean life of 400 hours and standard deviation as 45 hours. Assuming the distribution of life time to be normal, find

07

- (i) The percentage of batteries with a life time of at least 470 hours.
- (ii) The proportion of batteries with life time between 385 and 415 hours.
- (iii) The minimum life of the best 5% batteries.

- Q.4 (a)** (i) Describe Type – I and Type – II errors.

03

- (ii) In order to estimate the average electric usage per month, a sample of 81 houses was selected, and the electric usage was determined. Assume a population standard deviation of 450-kilowatt hours. Compute at 95% confidence, the size of the margin of error and confidence interval assuming sample mean = 1,858 KWH.

04

- (b)** In a random sample of 500 persons from Maharashtra, 200 are found to be consumers of vegetable oil. In another sample of 400 persons from Gujarat, 200 are found to be consumers of vegetable oil. Discuss whether the data reveal a significant difference between Maharashtra and Gujarat so far as the proportion of vegetable oil consumers is concerned. Use $\alpha = 0.01$.

07

OR

- Q.4 (a)** (i) Explain in brief the properties of point estimators.

03

- (ii) In a sample of 400 voters, 360 indicated they favor the incumbent governor. Determine the 95% confidence interval of voters not favoring the incumbent.

04

- (b) A tire manufacturer has been producing tires with an average life expectancy of 26,000 miles. Now the company is advertising that its new tires' life expectancy has increased. In order to test the legitimacy of the advertising campaign, an independent testing agency tested a sample of 6 of their tires and has provided the following data. 07

Life Expectancy (in Thousands of miles):	28	27	25	28	29	25
---	----	----	----	----	----	----

At 99% confidence, test to determine whether or not the tire company is using legitimate advertising. Assume the population is normally distributed.

- Q.5 (a)** The following data represent a company's yearly sales volume and its advertising expenditure over a period of 8 years. 07

Sales in Millions of Dollars (Y)	15	16	18	17	16	19	19	24
Advertising in (\$10,000) (X)	32	33	35	34	36	37	39	42

- (i) Use the method of least squares to compute an estimated regression line between sales and advertising.
(ii) Compute SST, SSR, and SSE.
(iii) Compute the coefficient of determination.
- (b) A large electronics firm that hires many workers with disabilities wants to determine whether their disabilities affect such workers' performance. Use the level of significance 0.05 to decide on the basis of the sample data shown in the following table whether it is reasonable to maintain that the disabilities have no effect on the workers' performance. 07

	Performance		
	Above average	Average	Below average
Blind	22	65	19
Deaf	17	49	24
No disability	29	95	30

OR

- Q.5 (a)** A regression and correlation analysis resulted in the following information regarding a dependent variable (y) and an independent variable (x). 07

$$\begin{aligned} \Sigma X &= 90 & \Sigma (Y - \bar{Y})(X - \bar{X}) &= 466 \\ \Sigma Y &= 170 & \Sigma (X - \bar{X})^2 &= 234 \\ n &= 10 & \Sigma (Y - \bar{Y})^2 &= 1434 \\ SSE &= 505.98 \end{aligned}$$

- (i) Find the estimated regression equation.
(ii) Compute and interpret coefficient of determination
- (b) Are the types of professional jobs held in the computing industry independent of the number of years a person has worked in the industry? Suppose 246 workers are interviewed. Use the results obtained to determine whether type of professional job held in the computer industry independent of years worked in the industry. 07

Professional Position	Year		
	0-3	4-8	> 8
Manager	6	28	47
Programmer	37	16	10
Operator	11	23	12
System Analyst	13	24	19
