

Seat No.: _____

Enrolment No. _____

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA - SEMESTER-III • EXAMINATION – SUMMER 2015

Subject Code: 2630003

Date: 12/05/2015

Subject Name: Statistical Methods (SM)

Time: 02:30 pm – 5:00 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 Standard deviation = 8 and Coefficient of variation = 64% then compute mean. **02**
- (ii) Define (i) Mutually Exclusive Events (ii) Exhaustive Events. **02**
- (iii) Assume that you have a binomial experiment with $p = 0.8$ and a sample size of 100. What is the value of variance of this distribution? **01**
- (iv) What is the Z value for a 97.5% confidence interval estimation? **02**

(b) (i) The following data show the number of hours worked by 400 statistics students. **04**

Number of Hours	Frequency
0 - 9	20
10 - 19	80
20 - 29	200
30 - 39	100

Referring above data Find:

1. The number of students working 19 hours or less.
 2. The percentage of students who work at least 10 hours per week.
 3. The cumulative relative frequency for the class of 20 – 29.
 4. The midpoint of the last class.
- (ii) An MCA applies for job in two companies X and Y. The probability of his being selected in company X is 0.7 and being rejected at Y is 0.5. The probability of at least one of his applications being rejected is 0.6. What is the probability that he will be selected in one of the companies? **03**

Q.2 (a) The hourly wages of a sample of eight individuals is given below. **07**

Hourly wage (in Rs.) : 27, 25, 20, 10, 12, 14, 17, 19

Compute mean, median, standard deviation, 25th percentile.

(b) (i) State Chebyshev's theorem. Using it, determine the percentage of the data within the range 32 to 48 for a sample with a mean of 40 and a s.d. of 7. **04**

(ii) The probability distribution of a random variable X is given below. **03**

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

Find

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

OR

(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.

- Q.3 (a)** Ten percent of the screws produced by automatic machine are defective. A random sample of Twenty screws is selected. **07**
- What is the probability that the sample contains exactly 3 defective?
 - What is the probability that the sample contains at most 4 defective?
 - What is the probability that the sample contains at least 5 defective?
 - What is the expected number of defective in the sample?
- (b)** The life expectancy of computer terminals is normally distributed with a mean of 4 years and a standard deviation of 10 months. **07**
- What is the probability that a randomly selected terminal will last more than 5 years?
 - What percentage of terminals will last between 2.5 and 4.5 years?
 - 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 |
- Compute the expected frequencies using Poisson distribution.
 - Test whether or not the Poisson model is a valid model for the above data.
- (b)** Assume that test scores from a college admissions test are normally distributed, with a mean of 450 and s.d. of 100. **07**
- What percentage of people taking the test score between 400 and 500?
 - Suppose someone receives a score of 630. What percentage of the people taking the test score better? What percentage score worse?
 - If a particular university will not admit anyone scoring below 480, what percentage of the persons taking the test would be acceptable to the university?

- Q.4 (a)** (i) Differentiate between cluster sampling and stratified sampling methods. **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) Describe Type – I and Type – II errors. **03**
- (ii) A new brand of chocolate bar is being market tested. Four hundred of the new chocolate bars were given to consumers to try. The consumers were asked whether they liked or disliked the chocolate bar. You are given their responses below. **04**

Response	Frequency
Liked	300
Disliked	100

Construct a 98% confidence interval for the true proportion of people who liked the chocolate bar.

- (b)** The student council of a college claims that the mean number of hours per week a student spends in the library during the final exam week is 12.5 hours. An administrator of the library chooses a sample of eight students who volunteer to provide the number of hours they have spent in the library during the final week of Fall '02. The hours are:
10.4, 9.1, 13.6, 6.4, 7.9, 8.6, 11.8, 12.6 **07**
- What can be said about the student council's claim based on the above dataset for the Fall '02 semester? Use $\alpha = 0.05$.

- Q.5 (a)** Assume you have noted the following prices for books and the number of pages that each book contains. **07**

Book	A	B	C	D	E	F	G
Pages (x)	500	700	750	590	540	650	480
Price (y) (in \$)	7	7.5	9	6.5	7.5	7	4.5

- (i) Use the method of least squares to compute an estimated regression line between the price and the number of pages.
 (ii) Compute SST, SSR, and SSE.
 (iii) Compute the 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

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)** A group of 2000 individuals from 3 different cities were asked whether they owned a foreign or a domestic car. The following contingency table shows the results of the survey. **07**

Type of Car	CITY		
	Ahmedabad	Baroda	Surat
Domestic	60	200	480
Foreign	140	600	520

At $\alpha = 0.05$ test to determine if the type of car purchased is independent of the city in which the purchasers live.
