

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA Integrated – SEMESTER III – EXAMINATION – SUMMER 2017

Subject Code: 4430603

Date: 10-05-2017

Subject Name: Statistical Methods

Time: 02:30 PM to 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) Write True/False with justification (7)

1. r is negative when both the variables are decreasing.
2. The variance of a sample of variable is -10.
3. The regression analysis helps us to study correlation between the variables.
4. The normal distribution with $X=0$ and $\sigma=1$ is known as standard normal distribution.
5. As the sample size increases, standard error also increases.
6. Type I error is an error committed by the test in accepting a true null hypothesis.
7. Degrees of freedom in case of two samples of sizes 50 and 60 are 108.

(b) A high school guidance counselor collected the following data about the grade point averages (GPA) and the SAT mathematics test scores for the six seniors. (7)

GPA	2.7	3.5	3.7	3.3	3.6	3.0
SAT	450	560	700	620	640	570

- (a) Is any relationship evident between the GPA and the SAT mathematics score? Explain.
- (b) Compute and interpret the sample covariance.
- (c) Compute the sample correlation coefficient and interpret the result.

Q.2 (a)(I) Define skewness with example. (3)

(II) The closing prices of 40 common stocks follow (4)

29.63	34	43.25	8.75	37.88	8.63	7.63	30.38	35.25	19.38
9.25	16.50	38	53.38	16.63	1.25	48.38	18	9.38	9.25
10	25.02	18	8	28.50	24.25	21.63	18.50	33.63	31.13
35.25	29.63	79.38	11.38	38.88	11.50	52	14	9	33.50

- (a) Construct frequency and relative frequency distribution.
- (b) Construct cumulative frequency and relative cumulative frequency distribution.

(b) A study of job satisfaction was conducted for four occupations: Cabinetmaker, Lawyer, physical therapist and system analyst. Job satisfaction was measured on a scale of 0-100. The data obtained are summarized in the following cross tabulation. (7)

Occupation	Satisfaction Score				
	Under 50	50-59	60-69	70-79	80-89
Cabinetmaker	0	2	4	3	1
Lawyer	6	2	1	1	0
Physical therapist	0	5	2	1	2
System analyst	2	1	4	3	0

- a) Develop the joint probability table.
- b) What is the probability one of the participants studied received a satisfaction score in the 80's?
- c) What is the probability of a satisfaction score in the 80's given the study participant was a physical therapist?
- d) What is the probability one of the participants studied was a lawyer?

OR

- (b) 50% of Americans think we are in recession, even though technically we have not had two straight quarters of negative growth. For a sample of 20 Americans, make the following calculations. (7)
- a) Compute the probability that exactly 12 people think we are in recession.
 - b) Compute the probability that no more than five people think we are in recession.
 - c) How many people would you expect to say we are in recession?
 - d) Compute the variance and Standard deviation of the number of people who think we are in recession.

- Q.3** (a) The average amount parents and children spent per child on back-to-school clothes in Autumn 2001 was Rs. 527. Assume the S.d. is 160 and the amount is normally distributed. (7)
- a) What is the probability that the amount spent on a randomly selected child is more than Rs. 700?
 - b) What is the probability that the amount spent on a randomly selected child is less than Rs. 100?
 - c) What is the probability that the amount spent on a randomly selected child is between Rs. 450 and Rs. 700?
 - d) What is the probability that the amount spent on a randomly selected child is no more than Rs. 300?
- (b) Bride magazine reported that the mean cost of a wedding is \$19000. Assume that the population standard deviation is \$9400. Bride's plan to use an annual survey to monitor the cost of wedding. Use 95% confidence. (7)
- a) What is the recommended sample size if the desired margin of error is \$1000?
 - b) What is the recommended sample size if the desired margin of error is \$500?
 - c) What is the recommended sample size if the desired margin of error is \$200?

OR

- Q.3** (a) Airline passengers arrive randomly and independently at the passenger screening facility at a major international airport. The mean arrival rate is 10 passengers per minute. (7)
- a) Compute the probability of no arrivals in one minute period.
 - b) Compute the probability that three or fewer passengers arrive in one minute period.
 - c) Compute the probability of no arrivals in a 15 second period.
 - d) Compute the probability of at least one arrivals in a 15 second period
- (b) Smith Travel Research provides information on the one-night cost of hotel rooms through-out the United States. Using \$22.50 as the planning value for the population standard deviation, what sample size is recommended for each of the following cases? Use \$2 as the desired margin of error. (7)
- a) A 90% confidence interval estimate of the population mean cost of hotel rooms
 - b) A 95% confidence interval estimate of the population mean cost of hotel rooms
 - c) A 99% confidence interval estimate of the population mean cost of hotel rooms

- Q.4** (a) Two laboratories A and B carry out estimates of fat content in ice-cream made by a firm. A sample is taken from each batch, halved, and the separated halves sent to the two laboratories. The fat content obtained by laboratories is recorded below: (7)
- | | | | | | | | | | | |
|------------------|---|---|---|---|---|---|---|---|---|----|
| Batch No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Lab A | 7 | 8 | 7 | 3 | 8 | 6 | 9 | 4 | 7 | 8 |
| Lab B | 9 | 8 | 8 | 4 | 7 | 7 | 9 | 6 | 6 | 6 |
- Is there a significant difference between the mean fat content obtained by the two laboratories A and B?

- (b) A sample of 900 members has a mean 3.4 cm and standard deviation 2.61 cm. Test whether the sample is from a large population of mean 3.25 cm and standard deviation 2.61 cm. if the population is normal and mean is unknown, find 95% confidence interval for population mean. (7)

OR

- Q.4** (a) An educator claims that the average I.Q. of American College students is at most 110 and that in a study made to test the claim 150 American college students, is selected at random, had an average I.Q. of 111.2 with a standard deviation of 7.2. use a level of significance 0.01 to test the claim of the educator. (7)

- (b) Samples of two types of electric bulb were tested for length of life and the following data are obtained: (7)

Sample No	Type I	Type II
	N1=8	N2=7
Sample means	1234 hrs.	1036 hrs.
Standard Deviation	36 hrs.	40 hrs.

Test whether the difference in the means is significantly different regarding length of life of two types of bulb at 5% level of significance.

- Q.5** (a) Given are the five observations for two variables x and y. (7)

X	2	4	5	7	8
Y	2	3	2	6	4

- Develop the estimated regression equation for these data.
- use the estimated regression equation to predict the value of y when x=4.
- Compute SST, SSE, and SSR.

- (b) Three brands of battery are under study. It is suspected that the life of 3 brands is different. 5 batteries of each brand are tested with the following results. (7)

Weeks of life		
Brand 1	Brand 2	Brand 3
10	6	8
6	8	4
2	7	6
6	4	8
2	2	4

Using ANOVA, test whether lives of these brands of batteries different at 5% level of significance. ($F_{12,2}$ at 5% is 19.40)

OR

- Q.5** (a) Sale of major appliances vary with the new housing market. A trade association compiled the following data on major appliance sale and housing market. (7)

Housing Market	2	3	4	4	5	5
Appliance sales	5	6	7	8	9	10

Develop an equation for relationship between appliance sale (in thousands) and housing market (in thousands) . fit a suitable regression line.

- (b) The Following table shows the sample values of three independent normal random variables, X1,X2 and X3. assuming that they have equal variances, test the hypothesis that they have the same mean by using ANOVA. ($F_{9,2}=19.385$) (7)

X1	13	11	16	22
X2	16	8	21	11
X3	15	12	25	10
