

Seat No _____

Enrollment No _____

Gujarat Technological University

MCA - SEMESTER-III EXAMINATION – SUMMER 2017

Subject Code: 2630003

Date: 02-06-2017

Subject Name: Statistical Methods (SM)

Time: 02:30 pm - 05:00 pm

Total Marks: 70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary and specifically write your assumptions.
3. Figures to the right indicate full marks.
4. The following values may be used wherever required:
Given: z value at 95% confidence level is 1.96 (for 2 tailed case) and 1.645 (for one-tailed case).
Given: t value at 95% confidence level is 2.571 (for 2 tailed case) and 2.015 (for one-tailed case) for df = 5. These values will be 2.447 and 1.943 respectively for df = 6.
Given: t value at 95% confidence level is 1.976 (for 2 tailed case) and 1.655 (for one-tailed case) for df = 149 or 150.
Given: t value at 99% confidence level is 3.012 (for 2 tailed case) and 2.650 (for one-tailed case) for df = 13. These values will be 2.921 and 2.583 respectively for df = 16.
Probability value corresponding to z between 0 & 2.04 in the Normal distribution table is 0.4793

- Q. 1** State whether the following statements are TRUE or FALSE. Give reason briefly. **14**
- (i) Mean is affected by each of the given data values while Median is not affected.
 - (ii) If Inter-quartile Range (IQR) is significantly less than the Standard Deviation, it indicates the existence of Outliers with good certainty.
 - (iii) The correlation coefficient of (+0.5) shows better correlation than the correlation coefficient of (-0.9).
 - (iv) The variance of a set of negative values will be negative.
 - (v) If the sample size is increased, the standard error is decreased.
 - (vi) If $P(A) = 0.3$, $P(B) = 0.4$, and $P(A \cup B) = 0.7$, we can conclude that events A and B are mutually exclusive.
 - (vii) In Poisson's probability distribution, the probability of only the occurrence of an event is known while the probability of non-occurrence of that event is not known. While in Binomial probability distribution, both the probabilities are known.
- Q. 2 (a)** There are two MCA colleges named (1) MCA-Delights and (2) MCA-Pride. Both **7**
colleges have 70 students studying in 6th semester of MCA. The mean CPI of college (1) is much lower than the mean CPI of college (2). A placement company was surprised to find that there are more number of students of college (1) having $CPI \geq 6.5$ as compared to college (2). The question and the task to you is to build CPI data for 70 students of college (1) and 70 students of college (2) to demonstrate that although Mean CPI of college (1) students is much lower than the mean CPI of college (2) students, but the number of students getting $CPI \geq 6.5$ is higher in college (1) than in college (2). [Hint: Try to fit students of the two colleges differently into the 4 groups: Very low CPI, Moderate CPI but less than 6.5, Little

better CPI ≥ 6.5 , High CPI ≥ 8 so as to meet the requirements of this question]

- (b) The cross tabulation of response frequencies for 159 statistics students, who answered questions about their political views and parent dominance on an anonymous survey, is given below. Work out probabilities to show the extent of Liberal Views being influenced by Parent's Dominance. Write your answer clearly. 7

| Political Views | ----- Dominant Parent ----- | | | Row Total |
|----------------------------|-----------------------------|-------------|------------|------------|
| | Mother (M) | Neither (N) | Father (F) | |
| Liberal (L) | 8 | 9 | 4 | 21 |
| Middle (D) | 32 | 44 | 29 | 105 |
| Conservative (C) | 11 | 7 | 15 | 33 |
| Column Total =====> | 51 | 60 | 48 | 159 |

[Hint: Compute probabilities such as $P(L)$, $P(L | M)$, $P(L | N)$, etc]

OR

- (b) Popcorn kernels are believed to take between 100 and 200 seconds to pop in a certain microwave. What sample size (number of kernels) would be needed to estimate the true mean seconds to pop with an error of ± 5 seconds and 95% confidence? Explain your assumption about σ . [Hint: Assume that the time range covers 4σ] 7

- Q. 3 (a)** The time required to pass through security screening at the airport can be annoying to travelers. The mean wait time during peak periods at Cincinnati / Northern Kentucky International Airport is 12 minutes. Assume the time to pass through security screening follows an exponential distribution. 7
- What is the probability it will take less than 10 minutes to pass through security screening during a peak period?
 - What is the probability it will take more than 20 minutes to pass through security screening during a peak period?
 - What is the probability it will take between 10 and 20 minutes to pass through security screening during a peak period?

- (b) Bride magazine reported that the mean cost of a wedding is Rs. 5,70,000. Assume that the population standard deviation is Rs. 1,80,000. Bride's plan to use an annual survey to monitor the cost of wedding. Use 95% confidence. 7
- What is the recommended sample size if the desired margin of error is Rs.30,000?
 - What is the recommended sample size if the desired margin of error is Rs. 15,000?
 - What is the recommended sample size if the desired margin of error is Rs. 6,000?

OR

- Q. 3 (a)** A shipment of 10 items has one defective and nine non defective items. In the inspection of the shipment, a sample of items will be selected and tested. If not a single defective item is found in the sample, then only the shipment of 10 items will be accepted; otherwise the shipment will be rejected. 7
- If a sample of three items is selected, what is the probability that the shipment

will be accepted?

(2) If a sample of four items is selected, what is the probability that the shipment will be accepted?

(3) If a sample of five items is selected, what is the probability that the shipment will be accepted?

(4) From the above, state whether any definite relationship exists between the sample size and the probability of acceptance of shipment? If yes, what is the relationship?

(b) According to BSNL, 70% of Internet users connect their computers to the Internet by normal telephone lines. 7

(i) What is the probability that a sample proportion from a simple random sample of 350 Internet users will be within ± 0.05 of the population proportion?

(ii) What is the probability that a sample proportion from a simple random sample of 350 Internet users will be 0.75 or greater?

Q. 4 (a) A company has recently created a new hair dryer A with fewer parts than the current hair dryer B. 300 units of each type of hair dryer were tested. 50 units of type A and 75 units of type B failed in a performance test. Can you conclude that new hair dryer is more reliable at 5% level of significance? 7

(b) The sales (in thousand Rs) data of an item in six shops before and after a special promotional campaign are as under: 7

| Shops | A | B | C | D | E | F |
|-----------------|-----|-----|-----|-----|-----|-----|
| Before campaign | 550 | 250 | 350 | 500 | 500 | 400 |
| After campaign | 600 | 220 | 300 | 550 | 580 | 450 |

Did the campaign make any significant difference in sales at 5% level of significance?

OR

Q. 4 (a) Define the chi-square test. 7

A die is thrown 151 times and the following results are obtained.

| | | | | | | |
|------------------|----|----|----|----|----|----|
| Number turned up | 1 | 2 | 3 | 4 | 5 | 6 |
| Frequency | 20 | 23 | 29 | 18 | 31 | 30 |

Test the hypothesis that the die is unbiased at 5 % level of significance

(b) When the background music was slow, the mean amount of bar purchases for a sample of 17 restaurant patrons was \$30.47 with a standard deviation of \$15.10. When the background music was fast, the mean amount of bar purchases for a sample of 14 patrons in the same restaurant was \$21.62 with a standard deviation of \$9.50. Assuming unequal variance, at 1% significance level, is the true mean higher when the music is slow? 7

Q. 5 (a) Power impacts for 8 observations were prepared at different compaction pressure and Hardness of the tablet, the following data obtain. 7

| | | | | | | | | |
|-------------|---|---|---|---|---|---|---|---|
| Observation | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------------|---|---|---|---|---|---|---|---|

| | | | | | | | | |
|-------------------|------|------|-----|-----|-----|-----|-----|-----|
| Pressure (tons) X | 0.25 | 0.75 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 4.0 |
| Hardness (Kg.) Y | 1.0 | 1.3 | 1.9 | 2.6 | 2.8 | 3.3 | 4.2 | 5.3 |

Determine the regression equation estimating Y as a linear function of X. Using this equation, estimate the Hardness value for Pressure = 5.0 tons.

- (b) (i) State Central Limit Theorem. How does it help in Testing of Hypothesis? **4**
(ii) Define Type-I and Type-II errors. Which error type (Type-I or Type-II) is tested through level of significance? **3**

OR

- Q. 5** (a) A department of transportation study on driving speed and mileage for mid-size fuel-efficient cars resulted in the following data. **7**

| | | | | | | | | |
|---------------------------|------|------|----|------|------|----|------|------|
| Driving Speed (Kms/hr) X: | 60 | 65 | 70 | 75 | 80 | 85 | 90 | 95 |
| Mileage (Kms/liter) Y: | 21.5 | 21.3 | 21 | 20.7 | 20.4 | 20 | 19.5 | 18.9 |

Determine the regression equation estimating Y as a linear function of X. Using this equation, estimate the Mileage if the speed is 100 Kms / hr.

- (b) (i) Define Simple Random Sampling, and Stratified Sampling methods. Give one example for each. **4**
(ii) Define Probability. Give the formula to determine probabilities of a union of two events in general. How is this formula simplified if the events are known to be mutually exclusive? **3**
