|  |  |  |  |
| --- | --- | --- | --- |
| **DEERWALK INSTITUTE OF TECHNOLOGY** | | | |
| **FINAL TERM EXAMINATION, SEM III** | | **SUBJECT** | **STA - 210 STATISTICS II** |
| **PASS MARK** | **24** | **FULL MARK** | **60** |
| **TIME** | **3 HRS** | **DATE** | **31 DECEMBER, 2021** |
| INSTRUCTIONS  *Students should be seated in the examination hall 10 minutes prior to the conduction of examination.*  *Fill in the required details carefully on the front page of the answer sheet.*  *Malpractice observed during the examination will lead to an expulsion.*  *Raise your hand towards Invigilator in case you need any assistance during the examination.*  *Good luck and all the best.*  **Attempt Any Eight Questions [8×5=40]**   1. What do you mean by partial correlation coefficient? State the relationship between simple and partial correlation coefficients when there are three variables. If r12 = 0.5, r23 = 0.1 and r13 = 0.4, compute r12.3 and r23.1. 2. The weights of 4 people before they stopped smoking in kilogram, are as follows:  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Before | 66 | 80 | 69 | 59 | 52 | 72 | 75 | 81 | | After | 71 | 82 | 68 | 62 | 56 | 67 | 73 | 77 |   Use the Wilcoxon matched pair sign rank test for paired observation to test the hypothesis, at 0.05 level of significance, that giving up smoking has no effect on a person’s weight against the alternative that one’s weight increases if he or she quits smoking.   1. A random sample of 15 adults living in a small town is selected to estimate the proportion of voting favoring a certain candidate for mayor. Each individual was also asked if he or she was a college graduate. By letting Y and N designate the responses of “yes” and “no” to the education question, the following sequence was obtained:  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | N | N | N | N | Y | Y | N | Y | Y | N | | Y | N | N | N | N |  |  |  |  |  |   Use the runs test at the 0.05 level of significance to determine if the sequence supports the contention that the sample was selected at random.   1. The average time taken by server to execute an algorithm varies from time to time. From the past experience it is known that the time taken is normally distributed with standard deviation of 6.7 minutes. The IT manager wishes to estimate the average by drawing a random sample such that the probability is 0.95 that the mean of the sample will not deviate by more than 1 minute from the population mean. What should be sample size? 2. A random sample of 10 bulbs has the following life in months; 24, 26, 32, 28, 20, 20, 23, 34, 30 and 43. Obtain the 95% confidence limit for the population mean life of bulbs. 3. In some town, each day is either sunny or rainy. A sunny day is followed by another sunny day with probability 0.7, whereas a rainy day is followed by a sunny day with probability 0.4. Weather conditions in this problem represent a homogeneous Markov chain with 2 states: state 1 = “sunny” and state 2 = “rainy.” Transition probability matrix of sunny and rainy days is given below.   P =  Compute the probability of sunny days and rainy days using the steady-state equation for this Markov chain.   1. Telephone calls arrive at telephone booth following Poisson distribution at an average time of 5 minutes between one and next. The length of phone call is assumed to be exponentially distributed with an average of 4 minutes. (i) What is probability that a person arriving at booth will have to wait? (ii) What is average length of queue that forms from time to time? What is probability that it takes more than 15 minutes altogether to wait for phone and complete call? 2. Consider the partially completed ANOVA table below. Complete the ANOVA table and answer the followings. What design was employed? How many treatments were compared? How many observations were analyzed? At the 0.05 level of significance can one conclude that the treatments have different effects? Why?      1. Write short notes on    1. Cochran Q test    2. Statistical assumptions of linear regression model   **Attempt Any Two Question [2×10=20]**   1. Suppose you are given following information with n = 28 Multiple regression model = 5 + 18X1 + 20X2 Sample size (n) = 28 Total sum of squares (TSS) = 250 Sum of squares due to error (SEE) = 100 Standard error of regression coefficient of X1 (Sb1) = 3.2 Standard error of regression coefficient of X2 (Sb2) = 5.5    1. Predict the value of Y for X1 = 15 and X2 = 25    2. Test the significance of regression coefficient of X2 at 5% level of significance.    3. Compute the coefficient of multiple of determination.    4. Test over all significance of regression model at 5% level of significance. 2. Differentiate between parametric and non-parametric test.   Test the hypothesis of no difference between the ages of male and female employees of a certain company using the Mann-Whitney U test for the sample data. Use the 0.05 level of significance.   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Males | 31 | 25 | 38 | 33 | 42 | 40 | 44 | 26 | | Females | 44 | 30 | 34 | 47 | 35 | 32 | 35 | 47 |  1. What is Randomized Block Design? Carry out ANOVA for following design.  |  |  |  |  | | --- | --- | --- | --- | | A ( 11) | C (10) | A (13) | B (12) | | C (12) | B ( 18) | B (9) | A (14) | | B ( 14) | A (10) | C (10) | C (16) |   Also calculate the relative efficiency of the design with respect to CRD. | | | |