**E.G.S. PILLAY ENGINEERING COLLEGE**

|  |
| --- |
| QB |
| Regulations 2023 |
| **AIDS** |

(*Autonomous*) Nagapattinam – 611 002

Year-II / Semester-III

**Course Code: 2301MA301**

**Course Name - PROBABILITY AND STATISTICS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PART A (Marks : 02)** | | **Marks** | **CO** | **BTL** |
| **Course Outcome : Use the fundamental concepts of probability and have knowledge of standard distributions**  **which can describe real life phenomenon** | | | | |
| 1 | Calculate the constant K given the following probability distribution of discrete random variable X.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | X | 1 | 2 | 3 | 4 | 5 | | P(X) | 0.1 | 0.2 | K | 2K | 0.1 | | 2 | CO1 | L2 |
| 2 | The number of hardware failures of a computer system in a week of operations has the following probability function   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | No.of failures | 0 | 1 | 2 | 3 | 4 | 5 | 6 |  |  |  |  |  |  |  |  | | probability | 0.18 | 0.28 | 0.25 | 0.18 | 0.06 | 0.04 | 0.01 |  |  |  |  |  |  |  |  |   Obtain the mean of the number of failures in week. | 2 | CO1 | L2 |
| 3 | If X and Y are independent random variable with variance 2and 3.Obtain the variance of 3X+4Y. | 2 | CO1 | L2 |
| 4 | The cumulative distribution function(CDF)of a random variable X is | 2 | CO1 | L2 |
| 5 | Find the Mgf for the following function given by   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  | | 2 | CO1 | L2 |
| 6 | A random variable X has the density functionFind the distribution function | 2 | CO1 | L2 |
| 7 | Determine the Binomial distributions for which the mean is 4 and variance is 3. | 2 | CO1 | L2 |
| 8 | If 3% of the electric bulbs manufactured by a company are defective, find the probability that in a Sample of 100 bulbs exactly 5 bulbs are defective. | 2 | CO1 | L2 |
| 9 | The diameter of an electric cable, say X, is assumed to be a continuous random variable with Probability density function f(x) =6x (1-x), . Verify that the above is a p.d.f or not? | 2 | CO1 | L2 |
| 10 | A continuous random variable X that can assume any value between x=2 and x=5 has the density function f(x)=k(1+x). Find k. | 2 | CO1 | L2 |
| **Course Outcome: Apply the basic concepts of one and two dimensional random variables in engineering**  **Applications**. | | | | |
| 1 | Explain joint probability function | 2 | CO2 | L2 |
| 2 | If f(x, y) =. Find K. | 2 | CO2 | L2 |
| 3 | Let x and y have joint density function f(x,y)=2,0<x<y<1.  Obtain marginal density function. | 2 | CO2 | L2 |
| 4 | The joint density function of the random variables X and Y is given by  f(x,y) =8xy,0<x<y<1.Find fX(x). | 2 | CO2 | L2 |
| 5 | Determine the constant k if, the joint p.d.f of a bivariate R.V(X,Y) is given by  f(x,y)= | 2 | CO2 | L2 |
| 6 | Let X be a R.V with p.d.f f(x)= ,-1≤x≤ 1 and letY=X2.Find E(Y) | 2 | CO2 | L2 |
| 7 | Explain about Regression coefficient of Yon X and X on Y. | 2 | CO2 | L2 |
| 8 | Explain Correlation | 2 | CO2 | L2 |
| 9 | Describe the scatter diagram for the various types of correlation. | 2 | CO2 | L2 |
| 10 | Given  Find the value of C. | 2 | CO2 | L2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Course Outcome: Simulate the concept of testing of hypothesis for small and large samples in real life problems.** | | | | |
| 1 | Discuss the assumptions of t-test? | 2 | CO3 | L2 |
| 2 | Explain the chi-square value for 2x2 contingency table? | 2 | CO3 | L2 |
| 3 | Describe the applications of F test. | 2 | CO3 | L2 |
| 4 | Describe Type I Error &Type II Error. | 2 | CO3 | L2 |
| 5 | Explain about Null Hypothesis (H0) and Alternative hypothesis? | 2 | CO3 | L2 |
| 6 | State the applications of chi-square distribution. | 2 | CO3 | L2 |
| 7 | Explain about sampling distribution? | 2 | CO3 | L2 |
| 8 | Describe Critical value or Significant value? | 2 | CO3 | L2 |
| 9 | Explain about the applications of t-distribution? | 2 | CO3 | L2 |
| 10 | Describe chi-square test of goodness of fit. [ | 2 | CO3 | L2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Course Outcome: Apply the basic concepts of classifications of design of experiments in the field of statistical quality control**. | | | | |
| 1 | Explain about ANOVA | 2 | CO4 | L2 |
| 2 | Describe the uses of analysis of variance? | 2 | CO4 | L2 |
| 3 | Describe one-way classification in ANOVA? | 2 | CO4 | L2 |
| 4 | Describe two-way classification in ANOVA? | 2 | CO4 | L2 |
| 5 | Explain about the Basic principles in the design of Experiment? | 2 | CO4 | L2 |
| 6 | Describe table of one way ANOVA | 2 | CO4 | L2 |
| 7 | Describe table of two way ANOVA | 2 | CO4 | L2 |
| 8 | Construct 4 X 4 Latin square designs. | 2 | CO4 | L2 |
| 9 | Compare LSD AND RBD | 2 | CO4 | L2 |
| 10 | Explain about merits and demerits of CRD | 2 | CO4 | L2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Course Outcome: Develop exposure to the principal component analysis of random vectors and Time Series.** | | | | |
| 1 | Explain about random vectors. | 2 | CO5 | L2 |
| 2 | Explain about random matrices. | 2 | CO5 | L2 |
| 3 | Computing expected values for discrete random variable of the probability function   |  |  |  |  | | --- | --- | --- | --- | | x | -1 | 0 | 1 | | P(x) | 0.3 | 0.3 | 0.4 | | 2 | CO5 | L2 |
| 4 | Describe population correlation coefficient. | 2 | CO5 | L2 |
| 5 | Computing the standard deviation matrix from the covariance matrix A= | 2 | CO5 | L2 |
| 6 | Explain the method of fitting a straight line. | 2 | CO5 | L2 |
| 7 | Write a brief note on seasonal variations | 2 | CO5 | L2 |
| 8 | Mention the components of the time series. | 2 | CO5 | L2 |
| 9 | Discuss about irregular variation | 2 | CO5 | L2 |
| 10 | Define Time series. | 2 | CO5 | L2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PART B (Marks : 16/12) | | Marks | CO | BTL |
| **Course Outcome : Use the fundamental concepts of probability and have knowledge of standard distributions**  **which can describe real life phenomenon** | | | | |
| 1 | A random variable X has the following probability function   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Values of X | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | Probability p(x) | a | 3a | 5a | 7a | 9a | 11a | 13a | 15a | 17a |   Compute (i)the value of ‘a’  (ii) P(x3),P(X 3),P(0X5)  (iii) the distribution function of x | 16 | CO1 | L3 |
| 2 | (i)The monthly demand for Allwyn watches is known to have the probability distributions   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Demand | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | Probability | 0.08 | 0.12 | 0.19 | 0.24 | 0.16 | 0.10 | 0.07 | 0.04 |   Compute an expected demand for watches and variance.  (ii) A random variable X has the p.d.f given by f(x) =λ.e-λx ,x ≥ 0 (i)Compute MGF, the first four moments about the origin. | 10  6 | CO1 | L3 |
| 3 | The amount of time in hours that a computer function before probability density function given byx0 Estimate the probability that (i) a computer will function between 50 and 150 hours before breaking down (ii) It will function less than 100 hours? | 16 | CO1 | L3 |
| 4 | (i)The density function of a random variable X is given by f(x) =K x (2-x), 0x2. Compute K, mean, variance and rth moment.  (ii)For the triangular distribution  Compute MGF, Mean and Variance. | 10  6 | CO1 | L3 |
| 5 | (i) Utilize binomial distribution compute its MGF, Mean and Variance  (ii) If10%of the screws produced by an automatic machine are defective, Estimate the probability that out of 20 screws selected at a random, there are(i) exactly 2 defectives  (ii)at least 2 Defectives (iii) between 1and 3 defectives (inclusive). | 8  8 | CO1 | L3 |
| 6 | (i) Utilize Poisson distribution compute its MGF, Mean and Variance  (ii)The atoms of a radioactive element are randomly disintegrating. If every gram of this element, on average emits 3.9 alpha particles per second, Estimate the probability that during next second the number of alpha particles emitted from 1gm is (i) almost 6 (ii) at least 2 (iii) at least 3 and almost 6. | 8  8 | CO1 | L2 |
| **Course Outcome: Apply the basic concepts of one and two dimensional random variables in engineering**  **Applications.** | | | | |
| 1 | Mathematical marks of second year economic communication engineering students in cat-I and cat-II exams are recorded and the HOD need to do an analysis based on the student performance on the particular subject for the improvement of result, the marks are given in the following table. Calculate the marks by considering random variable and give your suggestion for the improvement. Also calculate the regression equation, correlation coefficient and mostly like marks in cat-II when marks in cat-I is 30   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | CAT - 1  marks out of 50 | 25 | 28 | 35 | 32 | 31 | 36 | 29 | 38 | 34 | 32 | | CAT - I1  marks out of 50 | 43 | 46 | 49 | 41 | 36 | 32 | 31 | 30 | 33 | 39 | | 16 | CO2 | L3 |
| 2 | If P(x, y) = K (2x+3y), x=0,1,2 and y=1,2,3. Simplify (i) P(x=0/y=2).  (ii)Find P(x +y) and P(x +y)>3. Also Estimate P(x≤1),P(y≤2), P(x≤1, y≤2) | 16 | CO2 | L3 |
| 3 | If the joint p.d.f of x and y is given by f(x, y)=  Compute (i) P(X<1).(ii) P(X<1 | Y<3). (iii) P(x +y<3) | 16 | CO2 | L3 |
| 4 | A statistical survey on the Heights of fathers and sons are taken and the observed data are given below.   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Height of father | 65 | 66 | 67 | 67 | 68 | 69 | 70 | 72 | | Height of Son | 67 | 68 | 65 | 68 | 72 | 72 | 69 | 71 |   Correlate the heights by designing two dimensional random variable and discuss the nature of correlation coefficient. | 16 | CO2 | L3 |
| 5 | If the joint density function of the random variable X and Y f(x,y)= 2-x-y, 0≤x≤1, 0≤y≤1. Examine the Correlation. | 16 | CO2 | L3 |
| 6 | The joint density function of X and Y is  Are X and Y are independent? Compute | 16 | CO2 | L3 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PART B (Marks : 16/12)** | | Marks | CO | BTL |
| **Course Outcome: Simulate the concept of testing of hypothesis for small and large samples in real life problems.** | | | | |
| 1 | (i)Ten individuals are chosen at random from a population and their heights are found to be in inches. 63,63,66,67,68,69,70,70,71,71. Discuss the suggestion that the mean heights in the universe is 66 inches  (ii) In an investigation into the health and nutrition of two groups of children of different social status, the following results are got   |  |  |  |  | | --- | --- | --- | --- | | Social Status  Healths | Poor | Rich | Total | | Below normal | 130 | 20 | 150 | | Normal | 102 | 108 | 210 | | Above normal | 24 | 96 | 120 | | Total | 256 | 224 | 480 |   Discuss the relation between the Health and their social status. | 8  8 | CO3 | L3 |
| 2 | (i)The table below gives the number of aircraft accidents that occurred during the various days of the week. Test whether the accidents are uniformly distributed over the week.   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Days: | Mon | Tue | Wed | Thu | Fri | Sat | | No. of accident: | 14 | 18 | 12 | 11 | 15 | 14 |   (ii)The average number of articles produced by two machines per day is 200 and 250 with standard deviation 20 and 25 respectively on the basis of records of 25 days of production. Can you regard both the machines are equally efficient at 1% level of significance? | 8  8 | CO3 | L3 |
| 3 | (i)Time taken by workers in performing a job are given below  Method I: 20 16 26 27 23 22 -  Method II: 27 33 42 35 32 34 38  Test at 1% level of significance whether there is any significant difference between the variance of time distribution.  (ii)The mean population of a random sample of 400 villagers in Jaipur district was found to be 400 with standard deviation of 12. The mean population of a random sample of 400 villages in Meerut district was found to be 395 with a standard deviation 15.  Is the difference between the two district means statistically significant? | 8  8 | CO3 | L3 |
| 4 | (i)A group of 10 people on diet A and another group of 8 people on diet B ,recorded that the following increase in weight   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Diet A | 5 | 6 | 8 | 1 | 12 | 4 | 3 | 9 | 6 | 10 | | Diet B | 2 | 3 | 6 | 8 | 10 | 1 | 2 | 8 | - | - |   Test the variances are significantly different.  (ii)4 coins were tossed 160 times and the following results were obtained   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | No.of heads | 0 | 1 | 2 | 3 | 4 | | Observed frequencies | 17 | 52 | 54 | 31 | 6 |   Under the assumption that the coins are unbiased, Calculte the expected frequencies of getting 0, 1, 2, 3, 4 heads and test the goodness of fit. | 8  8 | CO3 | L3 |
| 5 | (i)Before an increase in Exercise duty on Tea, 800 persons out of 1000 persons were found to be Tea drinkers and after an increase in exercise duty in Tea 800 people were Tea drinkers in a sample of 1200 people. Test whether there is any significant difference in tea consumption at 5 % level.  (ii)Two independent samples of sizes 9 and 7 from a normal population had the following values of the variables.   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Sample I | 18 | 13 | 12 | 15 | 12 | 14 | 16 | 14 | 15 | | Sample II | 16 | 19 | 13 | 16 | 18 | 13 | 15 | - | - |   Do the estimates of the population variance differ significantly at 5% level? | 8  8 | CO3 | L3 |
| 6 | 1. In a big city 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers? 2. Two horses A and B were tested according to the time(in seconds) to run a particular race with the following results:  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | Horse A | 28 | 30 | 32 | 33 | 33 | 29 | 34 | | Horse B | 29 | 30 | 30 | 24 | 27 | 29 | - |   Test Whether the horse A is running faster than B at 5% level.. | 8  8 | CO3 | L2 |

|  |
| --- |
|  |

|  |
| --- |
|  |

|  |
| --- |
|  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PART B (Marks : 16/12)** | | Marks | CO | BTL |
| **Course Outcome: Apply the basic concepts of classifications of design of experiments in the field of statistical quality control**. | | | | |
| 1 | A completely randomized design experiment with 10 plots and 3 treatments gave the following results:   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Plot No : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | Treatment: | A | B | C | A | C | C | A | B | A | B | | Yield : | 5 | 4 | 3 | 7 | 5 | 1 | 3 | 4 | 1 | 7 |   Analyze the results for treatment effects. | 16 | CO4 | L3 |
| 2 | The following are the numbers of mistakes made in 5 successive days of 4 technicians working for a photographic laboratory:   |  |  |  |  | | --- | --- | --- | --- | | Technician I  (X1) | Technician II  (X2) | Technician III  (X3) | Technician IV  (X4) | | 6  14  10  8  11 | 14  9  12  10  14 | 10  12  7  15  11 | 9  12  8  10  11 |   Test at the level of significance α = 0.01, whether the differences among the 4 sample means, can be attributed to chance. | 16 | CO4 | L3 |
| 3 | The following table give the yields on 12 samples plots under three varieties of seed   |  |  |  | | --- | --- | --- | | A | B | C | | 21 | 20 | 28 | | 23 | 17 | 22 | | 16 | 15 | 28 | | 20 | 13 | 32 |   Manipulate the average yields of land under different variances show significant difference | 16 | CO4 | L3 |
| 4 | Four doctors each test 4 treatments for a certain disease and observe the number of days each patient takes to recover. The recovery time in number of days are given as follows   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Treatment | | | | | Doctor | A | B | C | D | | 1 | 10 | 14 | 19 | 20 | | 2 | 11 | 51 | 17 | 21 | | 3 | 9 | 12 | 16 | 19 | | 4 | 8 | 13 | 17 | 20 |   Carry out an Analysis of Variance at 5 % level of significance, to test whether the sales differ with respect (i) Treatment (ii) Doctor | 16 | CO4 | L3 |
| 5 | To study the performance of 3 detergents and three different water temperatures, the following whiteness readings were obtained using specially designed equipment.  Perform two way ANOVA at 5% level.   |  |  |  |  | | --- | --- | --- | --- | |  | Detergents | | | | Water Temperature | A | B | C | | Cold | 57 | 55 | 67 | | Warm | 49 | 52 | 68 | | Hot | 54 | 46 | 58 | | 16 | CO4 | L3 |
| 6 | A farmer wishes to test the effects of four different fertilizers A, B, C D on the yield of wheat. In order to eliminate sources of error due to variability in soil fertility, he uses the fertilizers, in a Latin square arrangement as indicated in the following table, where the numbers indicate yields in business per unit area.   |  |  |  |  | | --- | --- | --- | --- | | A 18 | C 21 | D 25 | B 11 | | D 22 | B 12 | A 15 | C 19 | | B 15 | A 20 | C 23 | D 24 | | C 22 | D 21 | B 10 | A 17 |   Perform an analysis of variance to determine, if there is a significant difference between the fertilizers at level of significance | 16 | CO4 | L2 |
| **PART B (Marks : 16/12)** | | Marks | CO | BTL |
| **Course Outcome: Develop exposure to the principal component analysis of random vectors and Time Series.** | | | | |
| 1 | 1. Computing the correlation matrix from the covariance matrix A= 2. Compute the principle components to the covariance matrix | 8  8 | CO5 | L3 |
| 2 | (i)Calculate the covariance matrix for the two random variable X and Y their joint probability function is represented by the entries in the following table   |  |  |  | | --- | --- | --- | | X/Y | X=0 | X=1 | | Y=-1 | 0.24 | 0.06 | | Y=0 | 0.16 | 0.14 | | Y=1 | 0.40 | 0.00 |   (ii)Compute the principle components to the matrix A= | 8  8 | CO5 | L3 |
| 3 | 1. Compute the quadratic trend of the form y=a+bx+cx2 for the data given below:  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | Years | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | | Production | 10 | 11 | 12 | 9 | 10 | 13 | 11 |  1. Determine the equation of the straight line which best fits the following data:  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Years | 1984 | 1985 | 1986 | 1987 | 1988 | | Sales(in Rs.1000) | 35 | 56 | 79 | 80 | 40 | | 8  8 | CO5 | L3 |
| 4 | 1. Calculate the 3-yearly moving average of the data given below:  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Years | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 |  | | Sales  (millions in Rs) | 3 | 4 | 8 | 6 | 7 | 11 | 9 | 10 | 14 | 12 |  |   Draw a graph to represent the moving averages ,also predict the sale for 1993.  (ii) Calculate the seasonal variation by the ratio to trend method from the data given below   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Year** | **I Quarter** | **II Quarter** | **III Quarter** | **IV Quarter** | | 1994 | 60 | 80 | 72 | 68 | | 1995 | 68 | 104 | 100 | 88 | | 1996 | 80 | 116 | 108 | 96 | | 1997 | 108 | 152 | 136 | 124 | | 1998 | 160 | 184 | 172 | 164 | | 16 | CO5 | L3 |
| 5 | Assuming 5-yearly moving average calculate trend values from the data given below also plot the  Result on a graph:     |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Years | 1971 | 1972 | 1973 | 1974 | 1975 | 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | | Production | 105 | 107 | 109 | 112 | 114 | 116 | 118 | 121 | 123 | 124 | 125 | 127 | 129 | | 16 | CO5 | L3 |
| 6 | (i)Compute the average seasonal movement for the following series   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Year | Quarterly production | | | | | 1 | 2 | 3 | 4 | | 1974 | 3.5 | 3.9 | 3.4 | 3.6 | | 1975 | 3.5 | 4.1 | 3.7 | 4.0 | | 1976 | 3.5 | 3.9 | 3.7 | 4.2 | | 1977 | 4.0 | 4.6 | 3.8 | 4.5 | | 1978 | 4.1 | 4.4 | 4.2 | 4.5 |   (ii)Given below are the figures of production of a sugar factory   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | Year | 1974 | 1975 | 1976 | 1977 | 1978 | 1979 | 1980 | | Production | 77 | 88 | 94 | 85 | 91 | 98 | 90 |   Fit a straight by least a squares method and tabulate the trend values. | 8  8 | CO5 | L2 |