KANNUR UNIVERSITY

(Abstract)

M.Sc. Statistics with Data Analytics programme under Credit Based Semester System (CBSS) offered at Don Bosco Arts & Science College Angadikkadavu - Model Question Paper of 1st Semester Core Courses - Implemented w.e.f 2022 Admission - Orders issued.

ACADEMIC C SECTION

AcadC/16764/MSc/ Stati- Data Analysis/2022 Dated: 17.06.2023

Read:-1. U.O. No. Acad C/16764/M.Sc./Stati-Data Analysis/2022 dtd.07.02.2023

2. U.O. Note No. Ex.C1/BQPS/M.Sc Statistics with Data Analytics Progm/2022 Dated 23.02.2023.

3. Letter No.acad.C/16764/M.Sc/Stati-Data Analysis/2022 dtd.04.03.2023

4. Email dtd.24.05.2023 from the Principal, Don Bosco Arts & Science College Angadikadavu .

ORDER

1. The Scheme, Syllabus and Model Question Papers of M.Sc. Statistics with Data Analytics Programme was implemented in the University w.e.f. 2022 admission, vide Paper read 1 above.
2. The Examination Branch vide Paper read 2 above, intimated that the U.O. referred 1 above does not contain the model Question Paper for all courses and requested to provide Model Question Paper for all courses of MSc Statistics with Data Analytics Programme and also sought clarification for the discrepancy in the course code of the Elective Paper Data Mining (it is mentioned as MST3E04 in the approved syllabus and MST3E01 in the Model Question Paper included in the approved Syllabus).
3. As Kannur University is not having the syllabus of aforementioned subject and in the circumstances of nonexistence of Board of studies, the Principal, Don Bosco Arts & Science College, Angadikadavu was submitted the syllabus of the Programme. Hence as ordered by the Registar, the Principal, Don Bosco Arts & Science College, Angadikadavu was requested to provide the Model Question Paper for all courses of the Programme and also sought clarification for the discrepancy in the Course code for the Elective course Data Mining in the Model Question Paper and in the approved Syllabus (ref 3)
4. The Principal, Don Bosco Arts & Science College, Angadikadavu, vide Paper read 4 above, submitted the Model Question Paper for the First semester Core courses of M.Sc. Statistics with Data Analytics Programme.
5. The Vice Chancellor after considering the matter in detail and exercise of the powers of the Academic Council conferred under section 11(1), chapter III of Kannur University Act 1996, ***accorded sanction to implement the Model Question Paper of First Semester M.Sc. Statistics with Data Analytics under CBSS w.e.f 2022 admission.***
6. The Model Question Papers of M.Sc. Statistics with Data Analytics Programme (CBSS)

are uploaded on the website of the University .

7. Orders are issued accordingly.

Sd/-

Narayanadas K

DEPUTY REGISTRAR (ACAD)

For REGISTRAR

To: 1.The Principal, Don Bosco Arts & Science College, Angadikadavu.

Copy To: 1. The Examination Branch (Through PA to CE)

2. EXCI, Computer Programmer, Web Manager, (to upload in web site)
3. PS to VC/PA to PVC/ PA to Registrar
4. DR/ARI Academic
5. EG1/EX CI (Exam)
6. SF/DF/FC

Forwarded / By Order

SECTION OFFICER

KANNUR UNIVERSITY
CIVIL STATION P.O.
KANNUR
Pin-670 002
THAVAKKARA

MODEL QUESTION PAPER

FIRST SEMESTER M.Sc DEGREE EXAMINATION

Branch: Statistics with Data Analytics

MST1C01: - Mathematical Methods for Statistics

Time : 3 Hours **Maximum Marks : 80**

Part A

(Answer ALL questions. Each question carries 2marks).

1. Define closed set. Give an example.
2. What is the kind of discontinuity of the function $f(x) = (\sin 2x)/x$; $x \neq 0$; $= 0$; if $x = 0$ at the origin?
3. Using Lagrange's mean value theorem prove that if $f'(x) = 0$ for all $x \in [a, b]$, then $f(x)$ is a constant on $[a, b]$.
4. A function f is defined on R by $f(x) = x$; $0 \leq x < 1$; $= 1$; $x \geq 1$. Examine whether the derivative of $f(x)$ at $x = 1$ exists.
5. Explain dimension of a vector space
6. Explain idempotent matrix.
7. Distinguish between ordinary inverse and generalized inverse.
8. Define index and signature of the quadratic form

(8 X 2 = 16 marks)

Part B

(Answer any FOUR questions. Each question carries 4 marks)

9. Establish Cauchy's principle of convergence of sequence of real numbers.
10. Explain classification of quadratic forms with an example.
11. State Alembert's Ratio Test. Test the convergence of the infinite series.
12. State and prove Rolle's theorem of differential calculus.
13. Define extreme values. Show that the function $f(x, y) = (y-x)^4 + (x-2)^4$ has a minimum at (2, 2).
14. Describe the method of finding inverse of a square matrix with an example

(4 X 4 = 16 marks)

Part C

(Answer any FOUR questions. Each question carries 12 marks)

15. Define limit point of a set. Give an example. Also state and prove Bolzano- Weierstrass theorem.
16. Explain linear dependence and independence. Prove that linear dependence and independence in a system of vectors is not changed by scalar multiplication of the Vectorsbynon-zeroscalar.
17. State and prove a necessary and sufficient condition for Riemann-Stieltje's integrability.
18. (i) State and prove fundamental theorem of integral calculus.

(ii) Show that every continuous function is integrable.

19. Explain basis and orthogonal basis. Also explain the Gram Schmidt orthogonalization process.
20. Explain diagonalization of a quadratic form. Prove that if A is a real symmetric matrix, then $P'AP = \Lambda$, where P is an orthogonal matrix and Λ is a diagonal matrix.

(4 X 12 = 48 marks)

MODEL QUESTION PAPER

FIRST SEMESTER M.Sc DEGREE EXAMINATION

Branch: Statistics with Data Analytics

MST1C02: - Probability Theory

Time : 3 Hours **Maximum Marks : 80**

PART A

Answer all questions. Each question carries 2marks.

1. Define monotone field. Show that a σ field is a monotone field.
2. Write a short note on induced probability space.
3. If $\{A_n\}$ is a sequence of arbitrary events, show that $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$
4. If X is random variables, show that $|X|$ is a random variable.
5. Let X have the probability mass function $P\left[X = (-1)^{j+1}\frac{3^j}{j}\right] = \frac{2}{3^j}; j = 1,2,3...$ What can you say about the existence of $E(X)$?
6. Let X be an integer-valued random variable with probability generating function P(s). Find the probability generating function of $Y = aX + b$,where a and bare non-negative integers.
7. Let $X_n \xrightarrow{p} X$ and g be a continuous function defined on $\Re$. Then show that $g(X_n)$ converges to $g(X)$ in probability.
8. If $X_1, X_2, ..., X_n$ is a sequence of Bernoulli random variables with probability success p, write down the central limit theorem result.

(8 X 2 = 16 marks)

PART B

Answer any 4 questions. Each question carries 4 marks

9. Let $\{A_n\}$ be a non-decreasing sequence of events. Then prove that $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$.
10. Define axiomatic definition of probability. State and prove addition theorem on probability.
11. Define distribution function of a random variable and show that distribution function is a right continuous function.
12. Let X be a nonnegative continuous random variable of the continuous type with probability density function $f(x)$ and let $\alpha > 0$,then find probability density function of $Y = X^{\alpha}$ using transformation technique.
13. Let X be random variable such that $P(X=0) = \frac{1}{4}$, $P(X=1) = \frac{1}{2}$, $P(X=2) = \frac{1}{4}$. Write down its distribution function.
14. Suppose $\{X_n\}$ is a sequence of random variables with probability mass function $P(X_n = 1) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$. Examine the convergence in probability of $\{X_n\}$.

(4 X 4 = 16 marks)

PART C

Answer any 4 questions. Each question carries 12 marks.

15. **(i)** State and prove Bayes' theorem.

(ii) A random variable X has the probability density function $f(x)=\begin{cases} x, & 0 \le x < 1 \\ 2-x, & 1 \le x < 2 \end{cases}$ find the moment generating function of X. Hence obtain mean and variance of X.

16. **(i)** Define expectation of a random variable. If X is a non-negative random variable with distribution function $F(x)$, prove that $E(X)=\int_0^{\infty}(1-F(x))dx$.

(ii) Define characteristic function of a random variable. If X and Y are two independent random variables with characteristic functions $\Phi_X(t)$ and $\Phi_Y(t)$ respectively, find the characteristic function of $X+Y$.

17. **(i)** If $E|X|^r < \infty$ then show that $E|X|^s < \infty$ for $0 < s \le r$.

(ii) State and prove Jensen's inequality. Using Jensen's inequality show that $E(\log X) \le \log(E(X))$, where X is a positive valued random variable, provided both expectations exist.

18. **(i)** Define convergence in probability. If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$, show that $X_n + Y_n \xrightarrow{p} X+Y$.

(ii) Let X be a random variable with $E|X|^k < \infty$ for some k, then show that $n^k P(|X| > n) \to 0$, as $n \to \infty$.

19. **(i)** Establish weak law of large numbers for a random sample $X_1, X_2, \ldots, X_n$ drawn from a population with mean μ and variance σ^2.

(ii) Let $\{X_k\}$ be a sequence of independent random variables with values $-2^k, 0$ and 2^k and probabilities $P(X_k = \pm 2^k) = 2^{-(2k+1)}$; $P(X_k = 0) = 1-2^{-2k}$. Examine whether weak law of large numbers holds for the sequence.

20. **(i)** Let $\{X_n\}$ be a sequence of random variables with distribution functio

$F_{X_n}(x)=\begin{cases} 1-(1-\frac{1}{n})^{n.x}, & x>0 \\ 0 & \text{otherwise} \end{cases}$. Show that $\{X_n\}$ converges in distribution to an exponential distribution

(ii) State Lindberg-Levy form of central limit theorem. Let $X_1, X_2, \ldots, X_n$ be a sequence of independentexponential random variables, with probability density function $f(x)=e^{-x}; x \ge 0$ and let $Y = X_1 + X_2 + \ldots + X_{100}$. Give an approximate value for $P(Y > 110)$ using central limit theorem.

(4 X 12 = 48 marks)

MODEL QUESTION PAPER

FIRST SEMESTER M.Sc DEGREE EXAMINATION

Branch: Statistics with Data Analytics

MST1C03: Distribution Theory

Time : 3 Hours **Maximum Marks : 80**

Part A

(Answer ALL questions. Each question carries 2 marks).

1. Define pgf and obtain the pgf of Negative Binomial distribution
2. Define generalized power series distribution and obtain its mgf.
3. Define log normal distribution and obtain its mean.
4. A truncated Poisson distribution is given by the mass function, $P(X = x) = e^{-\lambda}\lambda^{x}/\{(1 - e^{-\lambda}).x!\}$, x = 1, 2, 3, ..., find the mean of the distribution.
5. Define conditional distribution.
6. For a random variable X, Define the r^{th} moment about 'a'.
7. Prove the addictive/Reproductive property of binomial distribution.
8. Find the mean of the Poisson distribution with parameter λ.

(8 X 2 = 16 marks)

Part B

(Answer any FOUR questions. Each question carries 4 marks)

9. Let $X_1, X_2, X_3, \ldots, X_{k-1}$ have a multinomial distribution. Find the mgf of the distribution and hence obtain the marginal distribution of X_1
10. State and prove lack of memory property of X ~ Geometric distribution
11. Explain recurrence relation for row moments
12. Write down the properties of probability mass function.
13. Derive the moment generating function of Poisson distribution. Hence, find its variance and the third central moment.
14. Establish the relationship between Chi-square, t and F distributions

(4 X 4 = 16 marks)

Part C

(Answer any FOUR questions. Each question carries 12marks)

15. (i) Define hyper geometric distribution
    (ii) Find the factorial moments of the hyper geometric distribution and hence or otherwise derive the mean and variance of the distribution.
16. Derive the recurrence relation of cumulants of power series distribution. Show that binomial distribution and negative binomial distribution are special cases of power series distributions
17. (i) Define Standard Weibull distribution and obtain its rth raw moment.

(ii) If X_i, i = 1,2,3,…,n are i.i.d.r.v's having Weibull distribution with three parameters, show that the variable Y = min(X_1, X_2, …, X_n) , also has Weibull distribution and identify the parameters.

18. Define non central F variate, derive its pdf and obtain the mean. Also deduce the pdf of central F distribution from the pdf of the non central F distribution.

19. (i) Define order statistics. Derive the distribution of rth order statistic based on a random sample of size n from a population.

    (ii) Derive the distribution of largest sample in case of uniform (0, θ) population.

20. (i) Derive the asymptotic distribution of sample median

    (ii) Derive the standard error of sample variance

………………

(4 X 12 = 48 marks)

MODEL QUESTION PAPER
FIRST SEMESTER M.Sc DEGREE EXAMINATION

Branch: Statistics with Data Analytics

MST1C04: Statistical Programming using R

Time : 3 Hours **Maximum Marks : 80**

Part A

(Answer ALL questions. Each question carries 2 marks).

1. Write the syntax to obtain the following matrix, A =[-8 3 0 4 -1 2 -5 7 3].
2. Obtain the syntax to get multiple rows and columns in R.
3. Write the syntax to obtain the rank of a matrix in R.
4. Write down the features of R programming.
5. Define scatter plot.
6. Define switch statement with syntax.
7. Define Shapiro-Wilk Normality test.
8. Write down different types of sampling methods

(8 X 2 = 16 marks)

Part B

(Answer any FOUR questions. Each question carries 4 marks)

9. Define different types of functions in R.
10. Define operators in R.
11. Explain any 4 graphs in R with syntax.
12. Obtain the syntax to plot pdf and cdf of a Poisson random variable with $\lambda = 1$
13. Create two matrixes in R and perform matrix multiplication.
14. Write the syntax to get transpose, inverse and power of a matrix in R.

(4 X 4 = 16 marks)

Part C

(Answer any FOUR questions. Each question carries 12 marks)

15. Answer the following
    i) Describe binomial distribution.
    ii) Obtain and plot pdf and cdf where p is fixed and n is varying x=1,2,3,.....10 ; n=30,40,50,60,70,80; p=0.05. Write the conclusion also.
16. Explain decision making statements in R
17. Write down R code for the following questions.
    i) Ten soldiers visit a riffle range for two consecutive weeks.For the first week their scores are 67,24,57,55,63,54,56,68,33,43 and during the second week they score in the same order 70,38,58,58,56,67,68,72,42,38. Examine if there is any significant difference in their performance.
    ii) Consider the following data set consisting of 10 independent measurements

80,120,110,115,83,88,95,99,107,100.Let us test whether this sample has come from a normal population with mean 100.

18. Define Independent sample t-test, Paired sample t-test and its corresponding non parametric tests with R codes.
19. Explain different sampling techniques in Sampling Methods in detail.
20.
    i) What are the various kinds of allocations in sampling methods
    ii) Explain two stage cluster sampling, and systematic sampling.

(4 X 12 = 48 marks)