Total No. of Questions : 5]

SEAT No. :

PC-4364

[Total No. of Pages : 3

[6347]-101
M.Sc.
STATISTICS
STS 501-MJ: Fundamental of analysis and calculus
(2023 Pattern) (Semester - I) (4 Credits)

Time : 3 Hours] ***[Max. Marks : 70***

Instructions to the candidates :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt all questions : **[5 × 2 = 10]**

a) Prove or disprove: Set of all rational numbers is a countable set.

b) Calculate limit supremum and limit infimurn for $a_n = (-1)^n \left(\frac{n+5}{n}\right)$.

c) If E° denote the set of all interior points of a set E, then prove that E° is always open.

d) Define convergence of sequence and uniformly continuous function.

e) Define concave function with illustration.

Q2) Attempt any Three of the following: **[3 × 5 = 15]**

a) Prove or disprove: Every Cauchy sequence is bounded.

b) Prove that $\lim_{n \to \infty} \frac{1}{\sqrt{n!}} = 0$.

c) What are countable and uncountable sets? Check whether the following sets are countable set or uncountable set :

i) Set of integers

ii) Set of all positive multiples of five.

iii) Set of real numbers in [0,1]

d) Show that arbitrary union of open set is open.

P.T.O.

Q3) Attempt any Three of the following: **[3 × 5 = 15]**

a) State and prove mean value theorem.

b) Determine the convergence of the given series.

i) $\Sigma \frac{(-1)^{n-1} x^n}{n}$

ii) $\Sigma(-1)^n x^{2n}$

c) Define compact set. Show that [0, 1] is compact set.

d) Define limit point of a sequence and give an example of a bounded sequence with exactly two limit points.

Q4) Attempt any Three of the following: **[3 × 5 = 15]**

a) If r is rational (r $\neq$ 0) and x is irrational, then prove that $r + x$ and rx are irrational. Also prove that $\sqrt{12}$ is irrational number.

b) Obtain the radius of convergence for the following series.

i) $\Sigma \frac{n^n}{n!} Z^n$

ii) $\Sigma \frac{2^n}{n!} Z^n$

c) Suppose $a_1 \geq a_2 \geq \ldots\ldots \geq 0$. Then show that the series Σa_n converges if and only if the series $\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \ldots\ldots$ converges.

Hence prove that $\Sigma \frac{1}{n^p}$ converges if $p \geq 1$ and diverges if p $\leq$ 1.

d) Define power series and radius of convergence. Obtain the radius of convergence for the following series : $\sum_{n=1}^{\infty} \frac{(-1)^n n^2 x^n}{5^n \sqrt{n^5}}$

Q5) Attempt any one of the following: **[1 × 15 = 15]**

a) i) Define the following Terms : **[8]**

A) Absolute continuity

B) Functions of bounded variation

C) Concave function

D) Directional derivative

ii) Show that if $\underline{F} : \mathbb{R}^n \to \mathbb{R}^m$ is a function that is continuously differentiable in an open neighborhood of a point $a \in \mathbb{R}^n$, then the Jacobian matrix $J(\underline{F})$ exists at that point where the elements of the Jacobian matrix are the first partial derivatives of the components of $\underline{F}$. **[7]**

b) i) Discuss Cauchy criterion for convergence with illustration. **[8]**

ii) If $S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ... + \frac{1}{n!}$, for n = 1,2,3,..... Prove that $2 < \lim_{n\to\infty} S_n < 3$. **[7]**

♦ ♦ ♦

Total No. of Questions : 5]

SEAT No. :

PC-4365

[Total No. of Pages : 3

[6347]-104

M.Sc. (Part - I)

STATISTICS

STS - 510 - MJ : Optimization Techniques

(2023 Pattern) (Semester - I) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions :** **[5 × 2 = 10]**

a) Write the concept of cutting plane method.

b) Describe the rules for entering variable in case of tie for the simplex method of linear programming problem?

c) Narrate about branch and bound method concept and methodology.

d) What do you understand about the network model?

e) Explain the meaning of prohibited transportation routes.

Q2) **Attempt any THREE of the following :** **[3 × 5 = 15]**

a) Explain the procedure of Kuhn-Tucker conditions to get an optimal solution to the quadratic problem.

b) Describe the theory of Karmarkar's polynomial time algorithm.

c) Write Dijkstra' s shortest route algorithm.

d) The mathematical model for daily data of two products of manufacturing firm has the quadratic form as given below:

Max $Z = 12x + 21y + 2xy - 2x^2 - 2y^2$

Subject to constraints

$8 - y \geq 0$

$10 - x - y \geq 0$

$x, y \geq 0$

Find the solution to the problem by Beale's method when $\lambda_1 = 0$ & $\lambda_2 \neq 0$.

P.T.O.

Q3) Attempt any Three of the following: **[3 × 5 = 15]**

a) What is the scope of network model.

b) Obtain the value of given problem by using dynamic programming: Max $Z = y_1 . y_2 . y_3$ subject to the constraint $y_1 + y_2 + y_3 = 10$ and $y_1, y_2, y_3 \geq 0$.

c) Explain the concept of duality in linear programming problem.

d) State the advantages of linear programming problem.

Q4) Attempt any THREE of the following: **[3 × 5 = 15]**

a) Explain the degeneracy in transportation problem and resolution for it.

b) Describe the general formation of non-linear programming problem.

c) Write the steps of Gomory's Pure-Integer Programming Algorithm.

d) Find the solution of the following linear programming problem

Maximize $Z = 40x_1 + 35x_2$

subject to the constraints

$$2x_1 + 3x_2 \leq 60,$$
$$4x_1 + 3x_2 \leq 96,$$
$$x_1, x_2 \geq 0$$

Q5) Attempt any One of the following: **[1 × 15 = 15]**

a) i) Consider the following project with the estimates of time in weeks:

Activity	A	B	C	D	E	F	G	H	I
Predecessor Activity	-	A	A	B	C	D,E	D,E	F	G
Optimistic time	2	3	8	9	8	16	19	2	1
Most Likely time	4	6	10	12	9	21	22	5	3
Pessimistic time	6	9	12	15	10	26	25	8	5

I) Draw the network diagram. **[4]**

II) Compute out the mean, variance, time required to complete the project. **[3]**

III) What is the probability that the project will not exceed 45 weeks? **[2]**

ii) Explain the advantages and disadvantages of dynamic programming problem. **[6]**

b) i) Write the Floyd's algorithm of shortest route problem. **[6]**

ii) Obtain the initial basic feasible solution of the following transportation problem using

Vogel's approximation method and check optimality of the solution: **[9]**

Destination / Source	D_1	D_2	D_3	D_4	Supply
S_1	19	30	50	10	7
S_2	70	30	40	60	9
S_3	40	8	70	20	18
Demand	5	8	7	14	

▽ ▽ ▽ ▽

Total No. of Questions : 5]

SEAT No. : ☐

PC-5100

[Total No. of Pages : 3

[6347]-105R
M.Sc.
STATISTICS
STS - 511 - MJ : Statistical Quality Control
(2023 Pattern) (Semester - I)

Time : 3 Hours] ***[Max. Marks : 70***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt each the following:** **[5 × 2 = 10]**

a) Justify the use of 3σ control limits for the control chart when characteristic under study is not normally distributed.

b) Define the terms Producer's risk and Consumer's risk.

c) State Renewal Reward theorem and give its one real life application.

d) Give any two real life examples of irrelevant components.

e) Derive the hazard rate for the Pareto distribution.

Q2) **Attempt any Three questions out of Four questions.** **[3 × 5 = 15]**

a) What is the reliability of the power system if it is operated by 3 independent and identical dry cells connected in series with the individual reliability of 0.9?

b) Explain block replacement policy and age replacement policy. Also, discuss the drawbacks of this policy.

c) Prove the following implications

DMRL $\Rightarrow$ NBUE and NBUE $\Rightarrow$ HNBUE

d) Write a short note on s-shaped property of reliability function.

P.T.O.

Q3) **Attempt any Three questions out of Four questions.** **[3 × 5 = 15]**

a) Obtain the probability of acceptance of double sampling plan. Hence find ATI.

b) State the Sigmund's approximation formula for ARL and determine the parameters of CUSUM chart.

c) Define C_p C_r and C_{Pk}. Also give the interpretation of it.

d) Explain the construction of Conforming Run Length (CRL) chart for process fraction defective.

Q4) **Attempt any Three questions out of Four questions.** **[3 × 5 = 15]**

a) Explain Multivariate control charts for mean vector when dispersion matrix is known. Also, state its limitations.

b) Prove that any subset of associated random variables are associated.

c) Obtain minimal path set for the 2-out of-3 coherent system.

d) Define IFRA class of life distribution. Hence prove the class of life distribution belongs to IFRA iff $\overline{F}(\alpha t) \geq [\overline{F}(t)]^{\alpha}$, $0 < \alpha \leq 1$.

Q5) **Attempt any One question out of Two questions.** **[1 × 15 = 15]**

a) The data in the following table represent individual observations on viscosity (x_i) taken hourly from a chemical process. The target value of molecular weight is 10 and the process standard deviation is $\sigma = 1$. Set up an EWMA control chart for $\lambda = 0.10$ and L = 2.7 and interpret from the plot. **[15]**

Period i	x_i	Period i	x_i
1	9.45	11	9.03
2	7.99	12	11.47
3	9.29	13	10.51
4	11.66	14	9.4
5	12.16	15	10.08
6	10.18	16	9.37
7	8.04	17	10.62
8	11.46	18	10.31
9	9.2	19	8.52
10	10.34	20	10.84

b) i) Define Module of Coherent System. Hence find modules of the coherent system (C, ϕ), where C = {1, 2, 3, 4, 5} and

$$\phi(\underline{X}) = X1[1-(1-X2)(1-X3)][1-(1-X4)(1-X5)]$$ **[8]**

ii) Obtain the hazard rate for the proportional hazard rate family and linear hazard rate family. **[7]**

▽ ▽ ▽ ▽

Total No. of Questions : 5]

SEAT No. :

PC-4367

[Total No. of Pages : 3

[6347]-106
M.Sc. - I
STATISTICS
STS-512 MJ : Actuarial Statistics
(2023 Pattern) (Semester - I) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt each of the following :** **[5 × 2 = 10]**

a) Define time until death r.v.. Obtain its survival function.

b) Define Makeham's force of mortality.

c) Examine whether following can serve as survival function.

$$S(x) = \frac{1}{(1+x)^2}$$

d) Suppose the life length r.v. X has a distribution with survival function : $s(x) = 1 - \frac{x^2}{100} \quad 0 \le x \le 10$.

Find ${}_5P_4$.

e) Explain deferred probabilities.

Q2) **Attempt any three of the following :** **[3 × 5 = 15]**

a) Suppose the life length r.v. X is modelled by a distribution with force of mortality specified below:

$$\mu_x = \begin{cases} 0.01 & ;0 \le x < 15 \\ 0.02 & :15 \le x < 25 \\ 0.03 & ;\ge 25 \end{cases}.$$

Find the p.d.f. of T(20).

b) Suppose the life length r.v. is modelled by a uniform distribution over the interval (0, w). Find d_x.

c) Explain in brief n year term life insurance, whole life insurance and n year endowment insurance. Also obtain expressions for net single premiums for each.

d) If $\mu_x = \frac{2}{x+1} + \frac{2}{100-x}$;$0 \le x \le 100$

Find the expected number of deaths which occur between ages 1 and 4 in a life table with a radix of 10,000.

Q3) Attempt any three of the following : **[3 × 5 = 15]**

a) Suppose the life length r.v. X has a distribution with survival function :

$$s(x) = 1 - \frac{x^2}{100} \quad 0 \le x \le 10 .\ ]$$

Find the distribution of K(4). Also obtain its expectation e_4.

b) Find the amount to which Rs. 10000 will accumulate after 10 years if the rate of interest is,

i) 5% as the force of interest.

ii) 5% as the effective rate of interest.

iii) 5% per annum payable quarterly.

c) Prove that, $a_{x:\overline{n}|} = \sum_{k=1}^{n} v^k {}_k p_x$.

d) Define annuity. Explain annuity certain immediate and annuity certain due. Obtain present value of these annuities.

Q4) Attempt any three of the following : **[3 × 5 = 15]**

a) Let $F(x) = 1 - \left(1 - \frac{x}{105}\right)^{1/5}$;$0 \le x \le 105$;

Calculate,

i) Probability that a life aged 30 survives to atleast age 70.

ii) The median future lifetime at age 50.

iii) The complete expectation of life at age 50.

b) Find the present value on March 1,2024 of the following cash flow :

5000 received on March 1, 2025, 7000 paid on March 1, 2026, 4000 received on March 1, 2027, 3000 paid on March 1, 2028 and 6000 received on March 1, 2029 when the effective rate of interest is 0.04.

c) Prove that, $S_{\overline{n|}} < \bar{S}_{\overline{n|}} < \ddot{S}_{\overline{n|}} \ \forall_n$

d) Explain loss at issue random variable. Also explain equivalence principle.

Q5) Attempt any one of the following : **[15]**

a) Suppose the life length r.v. X is modelled by a distribution with force of mortality specified below: **[15]**

$$\mu_x = \begin{cases} 0.04 & ; 0 \le x \le 15 \\ 0.08 & ; 15 \le x \le 25 \\ 0.12 & ; 25 \le x \le 35 \\ 0.18 & ; x \ge 35 \end{cases}$$

Find $1000\bar{A}_x$, for x = 25, 30, 35 and 40.

b) i) Show that $a_{x:\overline{n|}} < \bar{a}_{x:\overline{n|}} < \ddot{a}_{x:\overline{n|}}$ **[8]**

ii) Suppose survival mode is defined by the following values of p_x. **[7]**

x	0	1	2	3	4
p_x	0.9	0.8	0.6	0.3	0

I) What is the corresponding values of $s(x)$ for x = 0,1,2,3,4,5?

II) Usiug radix l_0 = 10000, find values of l_x and d_x.

III) Find ${}_3d_0$, ${}_2q_1$, ${}_3p_1$ and ${}_3q_2$

⌘⌘⌘

Total No. of Questions : 5]

SEAT No. :

PC-4368

[Total No. of Pages : 3

[6347]-201
M.Sc.
STATISTICS
STS 551-MJ: Modern Statistical Inference
(2023 Pattern) (Semester - II) (4 Credits)

Time : 3 Hours] ***[Max. Marks : 70***

Instructions to the candidates :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt all questions : **[5 × 2 = 10]**

a) Define sufficient statistic with an illustration.

b) Distinguish between most powerful test and uniformly most powerful test.

c) Define the term estimability and hence show that if there exist two unbiased estimators of θ,then there exist infinitely many unbiased estimators of θ.

d) Define shortest expected length with illustration.

e) Define the terms:

i) Test function

ii) Type I and type II error.

Q2) Attempt any Three of the following: **[3 × 5 = 15]**

a) State and prove Cramer-Rao Inequality.

b) Define Pitman family. Find minimal sufficient statistic for θ *for a* $(\theta)\downarrow$ and $b\ (\theta)\uparrow$.

P.T.O.

c) Let $X_1, X_2, ..., X_n$ be a random sample from Bernoulli (p) find the MVBUE of p.

d) Let $X_1, X_2, ..., X_n$ is a random sample from discrete distribution with

$P(X_1 = 1) = \frac{2(1-\theta)}{2-\theta}, P(X_2 = 2) = \frac{\theta}{2-\theta}$ where $0 < \theta < 1$ is unknown.

Find the estimator of θ by using method of moments.

Q3) Attempt any Three of the following: **[3 × 5 = 15]**

a) If $T(\underline{X})$ is complete sufficient statistic then show that $T(\underline{X})$ is independent of every ancillary statistic $S(\underline{X})$.

b) Define Fisher information function and show that, in the regular case of point estimation, that $V\left(\frac{\partial \log f(x,\theta)}{\partial\theta}\right) = -E\left(\frac{\partial^2 \log f(x,\theta)}{\partial\theta^2}\right)$.

c) Let $X_1, X_2, ..., X_n$ be a random sample from *Binomial* $(1,\theta)$ for $0 \le \theta \le \frac{1}{2}$ find maximum likelihood estimator of θ.

d) Let $X_1, X_2, ..., X_n$ be a random sample from $N(\mu, \sigma^2)$ obtain $(1-\alpha)$ level shortest expected length confidence interval for μ when σ^2 is known.

Q4) Attempt any Three of the following: **[3 × 5 = 15]**

a) State and prove Rao-Blackwell Theorem.

b) Let $X_1, X_2, ...X_n$ be a random sample from $N(\mu, \sigma^2)$ μ is known find $I_n(\sigma^2)$

c) Show that $N(\mu, \sigma^2)$ has MLR property obtain UMP test of size α for testing H_0: $\mu \le \mu_0$ against H_1: $\mu > \mu_0$.

d) Let X has binomial distribution with parameter n and p. Suppose that n is given and the unknown parameter p has prior distribution, which is uniform on the interval [0, 1]. By considering the squared error loss function and observation X = n. Find Bayes estimator of p and the median of posterior distribution of p.

Q5) Attempt any one of the following: **[1 × 15 = 15]**

a) i) Let $X_1, X_2, ..., X_n$ be a random sample from *Geometric* (p) UMVUE of p. **[7]**

ii) Define uniformly most accurate confidence interval. Let $X_1, X_2, ..., X_n$ be a random sample from $N(\mu, \sigma^2)$, σ^2 is known. For testing H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$. Obtain uniformly most accurate $1 - \alpha$ level confidence sets of μ. **[8]**

b) i) State and prove Neyman Pearson lemma. **[8]**

ii) Let $f(x, y, \lambda, p) = \binom{x}{y} p^y (1-p)^{(x-y)} \dfrac{e^{-\lambda}\lambda^x}{x!}, y = 0,1,2,..., x; x = 0,1,2,...;$

$\lambda > 0$ and $0 < p < 1$ be joint pmf of (X,Y). Show that it belongs to two parameter exponential family and obtain minimal sufficient statistic $(T_1, T_2)'$ for $(p, \lambda)'$. **[7]**

♦ ♦ ♦

Total No. of Questions : 5]

SEAT No. :

PC-4369

[Total No. of Pages : 2

[6347]-202
M.Sc (Statistics)
STS-552-MJ : Regression Analysis and Applications
(4 credits) (2023 Pattern) (Semester - II)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions.** **[5 × 2 = 10]**

a) For the given data we obtain the eigenvalues of $X'X$ matrix as $\lambda_1 = 3.169$, $\lambda_2 = 1.006$, $\lambda_3 = 0.763$, $\lambda_4 = 0.553$, $\lambda_5 = 0.317$, $\lambda_6 = 0.192$. Calculate the condition number and state whether the problem of multicollinearity is harmful or not?

b) Write the given model in general linear model set up.

$E(Y_1) = \theta_1 + \theta_2$, $E(Y_2) = \theta_1 - \theta_3$ and $E(Y_3) = \theta_1 + \theta_2$

c) Define link function for the generalized linear model.

d) Explain the no intercept term model.

e) State and prove any two properties of direct regression estimators.

Q2) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Obtain $100(1 - \alpha)$ % joint confidence region for repressor parameters β_0 and β_1 of the multiple linear regression model.

b) Explain the reverse regression method with a suitable example. Also, obtain the least square estimators of the regression coefficients for it.

c) Define Polynomial regression model. Estimate the parameters of orthogonal polynomial regression model and give its variance.

d) Obtain the least square estimators of the regression coefficients for reverse regression model.

P.T.O.

Q3) Attempt any 3 questions out of 4 questions. **[3 × 5 = 15]**

a) Discuss the method of principle component regression for dealing with the problem of multicollinearity.

b) Discuss Wald test in logistic regression model.

c) For the generalized linear model, explain the following link function.

i) Binomial function

ii) Inverse binomial

iii) Inverse gamma

d) In case of near or high multicollinearity, explain the all possible consequences which can be encountered.

Q4) Attempt any 3 questions out of 4 questions. **[3 × 5 = 15]**

a) Explain the Gauss-Newton iteration method of parameter estimation in nonlinear regression models.

b) How various residual plots are useful in cheking validity of assumptions made in analysis of experimental data

c) What is pure error? When one can have an estimate of it?

d) Describe the method of maximum likelihood estimation to estimate the parameters of the poisson regression model.

Q5) Atttempt any 1 question out of 2 questions. **[1 × 15 = 15]**

a) i) Define: Intrinsic Model and Extrinsic Model. **[2]**

ii) Describe Box and Cox transformation in regression model. **[3]**

iii) Describe the problem of autocorrelation? **[3]**

iv) For the simple linear regression model, with first order autoregressive errors, Discuss the Durbin-Watson test to detect the presence of autocorrelation in errors. **[7]**

b) i) Discuss Score test in logistic regression model. **[4]**

ii) Explain the logistic regression model with single explanatory variable. Also, obtain the maximum likelihood equation for it. **[8]**

iii) Define the following terms: **[3]**

I) Mallows Cp-statistic

II) PRESS residuals

III) Leverage points

Total No. of Questions : 5]

SEAT No. :

PC-4370

[Total No. of Pages : 3

[6347]-203
M.A./M.Sc.
STATISTICS
STS - 553 - MJ : Multivariate Analysis & Applications
(2023 Pattern) (Semester - II) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of Statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) **Attempt all questions.** **[5 × 2 = 10]**

a) Define Mahalanobis D^2 statistic. Give relation between Mahalanobis D^2 and Hoteling T^2 statistic.

b) Define (i) Sample mean vector (ii) Wishart matrix.

c) Write note on Principal Component Analysis.

d) Distinguish between student's t and Hoteling T^2 statistic.

e) Obtain moment generating function of multivariate normal distribution.

Q2) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Let $X=\begin{bmatrix} 14 & 12 & 18 & 20 \\ 4 & 6 & 8 & 2 \end{bmatrix}$ obtain sample mean vector and sample variance-covariance matrix.

b) State and solve any two properties of Wishart distribution.

c) Let $\underline{X} \sim N_p(\mu, \Sigma)$ if partition of random vectors $[X^{(1)}, X^{(2)}]$ then obtain the marginal distribution of $X^{(1)}$.

d) Write a note on Scree plot.

P.T.O.

Q3) Attempt any 3 questions out of 4 questions. **[3 × 5 = 15]**

a) Let X ~ $N_3(0, \Sigma)$ where $\Sigma = \begin{bmatrix} 1.0 & 0.8 & -0.4 \\ 0.8 & 1.0 & -0.56 \\ -0.4 & -0.56 & 1.0 \end{bmatrix}$ find the conditional distribution of $\begin{bmatrix} X_2 \\ X_3 \end{bmatrix} / X_1$.

b) Derive the characteristic function of Wishart distribution.

c) Let $\underset{\sim}{X} \sim Np(0,1)$ and if A and B are real symmetric matrices of order p. then E(X′A X) = tr A.

d) State the differences between hierarchical clustering and non-hierarchical clustering method.

Q4) Attempt any 3 questions out of 4 questions. **[3 × 5 = 15]**

a) Let $\underset{\sim}{X} \sim Np(0, \Sigma)$ and C has the Wishart distribution then $\frac{n-p+1}{p} X'C^{-1}X \sim F_{p,\, n-p+1}$ derive the distribution of T^2 under the null hypothesis $H_0 : \mu = \mu_0$.

b) Obtain M.L.E. of $\underset{\sim}{\mu}$ and Σ when $\underset{\sim}{X} \sim Np\,(\underset{\sim}{\mu}, \Sigma)$.

c) Write note on non-Gaussian distribution? Give the pdf of multivariate Beta distribution.

d) Let $\underset{\sim}{X} \sim N_3(\mu, \Sigma)$ with $\underset{\sim}{u}' = [2, -3, 1]$ and $\Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$ obtain the Conditional distribution of X_3 given $X_1 = 3$ and $X_2 = 2$.

Q5) Attempt any 1 question out of 2 questions. **[1 × 15 = 15]**

a) i) Determine the principal components Y_1, Y_2 and Y_3 for the covariance matrix $\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ Also, calculate the proportion of total population variance explained by the first principal component. **[10]**

ii) $\underline{X} \sim N_2(\underline{0}, \Sigma)$ where $\Sigma = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$ Find $P(X_1 < X_2)$. **[5]**

b) i) Analyse the following data using one way MANOVA **[10]**

Treatments	**Observations**
1	$\begin{bmatrix} 9 \\ 3 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} \begin{bmatrix} 9 \\ 7 \end{bmatrix}$
2	$\begin{bmatrix} 0 \\ 4 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
3	$\begin{bmatrix} 3 \\ 8 \end{bmatrix} \begin{bmatrix} 1 \\ 9 \end{bmatrix} \begin{bmatrix} 2 \\ 7 \end{bmatrix}$

ii) Define canonical correlation and canonical variables give the applications of canonical correlation. **[5]**

▽ ▽ ▽ ▽

Total No. of Questions : 5]

SEAT No. :

PC-4371

[Total No. of Pages : 3

[6347] - 204

M.Sc. (Statistics)

STS-560-MJ: Advances in Generalized Linear Models

(2023 Pattern) (Semester - II) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) **Attempt all of the following questions :** **[5 × 2 = 10]**

a) Define concordance and discordance in binary logistic regression.

b) Define 'Deviance' for a Generalized Linear Model (GLM).

c) Explain 'Baseline Category Logits' for nominal response variables.

d) State any two differences between Logit and Probit model.

e) Explain the concept of zero inflation in data analysis.

Q2) **Attempt any 3 questions out of 4 questions :** **[3 × 5 = 15]**

a) For a GLM with canonical link function, explain how the likelihood equations imply that the residual vector $e = (y-\hat{\mu})$ is orthogonal with C(X).

b) For the one-way layout for Poisson counts, derive the likelihood-ratio statistic for testing $H_o: \mu_1 = \cdots = \mu_c$.

c) Explain Probit model in detail.

d) Explain the concept of Generalized Estimating Equations (GEE).

P.T.O.

Q3) **Attempt any 3 questions out of 4 questions :** **[3 × 5 = 15]**

a) Show that the exponential dispersion family representation for the gamma distribution.

b) Discuss the following goodness of model fit measures

   i) Pearson Chi-square   ii) Akaike Information Criteria

c) Explain the concept of the Beta-Binomial model in statistical modeling.

d) Discuss the need for Poisson regression model and its use. State the assumptions of Poisson regression model and also provide the tests for determining statistical significance of regression coefficients.

Q4) **Attempt any 3 questions out of 4 questions :** **[3 × 5 = 15]**

a) Show that a gamma mixture of Poisson distributions yields the negative binomial distribution.

b) Explain GLM and its components.

c) Compare the empirical Bayes and hierarchical Bayes modeling approaches in Bayesian statistics.

d) Explain the Quasi-Likelihood Method in statistical modeling. Discuss its advantages over traditional likelihood-based methods, and give with one example.

Q5) **Attempt any 1 question out of 2 questions :** **[1 × 15 = 15]**

a) i) The following are output table from a Poisson regression model.**[8]**

	Coef	Std error	z	P> |z|	[0.025	0.975]
Intercept	1.203	0.123	9.779	0.000	0.962	1.445
Age	-0.034	0.009	-3.765	0.000	-0.052	-0.016
Income	0.027	0.005	5.322	0.000	0.017	0.037

A) Interpret the coefficient estimates for the Intercept, Age, and Income variables.

B) Assess the statistical significance of each coefficient using the provided p-values.

C) Determine whether Age and Income variables are statistically significant predictors of the outcome.

D) Discuss the implications of the confidence intervals for the coefficient estimates.

ii) Explain Poisson regression model in detail and state any two applications. **[7]**

b) i) Discuss Cumulative logit model for multinomial response model with example. **[8]**

ii) Compare the marginal modeling with Generalized Estimating Equations (GEE) in the context of analyzing correlated or clustered data and discuss their respective strengths, limitations, and applications. **[7]**

Total No. of Questions : 5]

SEAT No. :

PC-4372

[Total No. of Pages : 4

[6347]-205
F.Y. M.Sc (Statistics)
STS-561-MJ : STATISTICAL METHODS IN EPIDEMIOLOGY
(2023 Pattern) (Semester - II)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*

Q1) **Attempt all the questions :** **[5 × 2 = 10]**

a) Differentiate between non-communicable and communicable disease.

b) Define: Latent period and incubation period

c) Draw a well- labeled diagram of chain of infection

d) What is composite sampling in epidemiology, and how does it differ from individual sampling?

e) Suppose the incidence rate of lung cancer is 46 new cancers per 100,000 person- years, and the prevalence is 23 per 100,000 populations.Calculate the average duration of disease.

Q2) **Attempt Any Three of the following :** **[3 × 5 = 15]**

a) The NCHS reported that the mean total cholesterol level in 2002 for all adults was 203. Let n = 3539 participants who attended the seventh examination of the Offspring in the Framingham Heart Study. The statistics on total cholesterol levels of participants: n = 3310 $\bar{x}$ = 200.3, and s = 36.8. Is there statistical evidence of a difference in mean cholesterol levels in the Framingham Offspring as compared to the national mean? i.e we want to assess whether the sample mean of 200.3 in the Framingham sample is statistically significantly different from 203 (i.e., beyond what we would expect by chance).

P.T.O.

b) Illustrate all the modes of direct transmission.

c) Define specificity in the context of diagnostic testing and explain its significance.

d) Explain the SEIR model.

Q3) Attempt Any Three of the following : [3 × 5 = 15]

a) Tabie shows data from a cohort study of oral contraceptive (OC) use and Myocardial Infarction (Ml) among women aged 16-49 years.

			MI	
OC Use		Yes	No	Total
	Yes	27	455	482
	No	77	1831	1908
	Total	104	2286	2390

Calculate Absolute Risk for the above table and give your interpretation.

b) How can agent-based models contribute to our understanding of the spatiotemporal spread of infectious diseases compared to compartmental models

c) Describe case-control study design with the help of diagram and give an example of it.

d) Define ROC curve and explain how it is constructed.

Q4) Attempt Any Three of the following : [3 × 5 = 15]

a) A clinical trial is conducted to compare an experimental medication to placebo to reduce the symptoms of asthma. Two hundred participants are enrolled in the study and randomized to receive either the experimental medication or placebo. The primary outcome is a self-reported reduction of symptoms. Among 100 participants who receive the experimental medication, 38 report a reduction of symptoms as compared to 21 participants of 100 assigned to the placebo. Generate a 95% Cl for the difference in proportions of participants reporting a reduction of symptoms between the experimental and placebo groups..

b) Discuss the key health outcomes that should be considered when assessing the impact of lockdowns on disease transmission and healthcare utilization during a pandemic such as COVID-19.

c) Define sensitivity in the context of diagnostic testing in epidemiology.

How is it calculated, and what does it represent?

d) A public health laboratory is tasked with screening a population of 1000 individuals for a certain infectious disease. Due to limited resources and time constraints, the laboratory decides to implement pooled testing. Under certain

Assumptions :

i) The prevalence of the disease in the population is estimated to be 5%.

ii) The laboratory decides to pool samples in groups of 10 individuals (pool size = 10) Determine the number of pools needed and estimate the expected number of positive pools.

Q5) **Attempt Any One of the following :** **[1 × 15 = 15]**

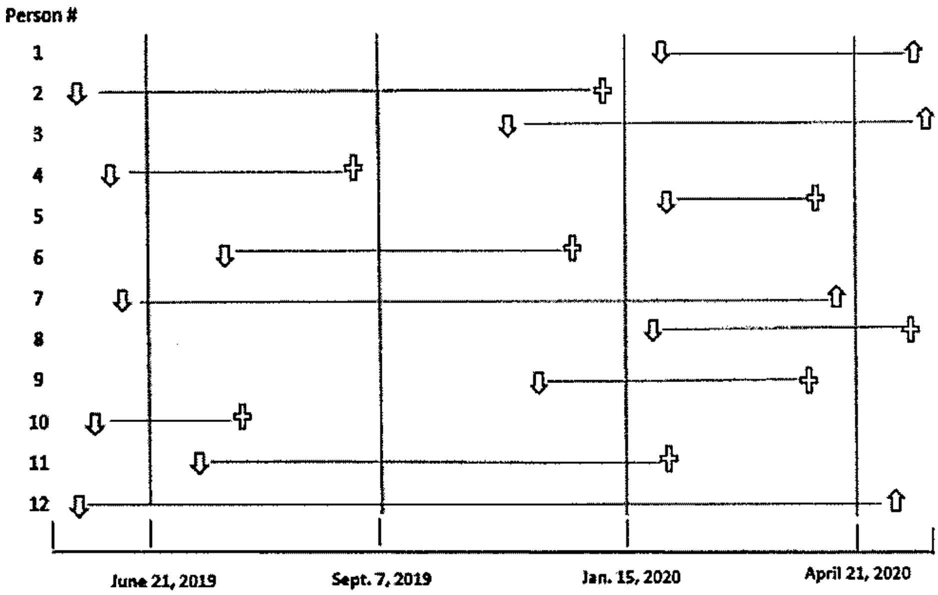
a) i) Explain the role of Markov Chain Monte Carlo (MCMC) methods in Bayesian inference for latent variable. **[7]**

ii) Write any four applications of SIR model. **[5]**

iii) Draw the cohort study design. **[3]**

b) i) Define Reproduction number (Ro) and determining the factors of Ro **[5]**

ii) Figure (attached below) represents 12 new cases of illness over about 10 months in a population of 40 persons. Each horizontal line represents one person. The down arrow indicates the date of onset of illness. The solid red line represents the duration of illness. The up arrow and the cross represent the date of recovery and date of death, respectively.



Figure, calculate the following :

1) Point prevalence on January 15, 2020

2) Period prevalence from June 21, 2019 to April 21, 2020

3) Incidence rate from June 21, 2019 to April 21, 2020 using the population alive on September 7, 2019 as the denominator. Express the rate per 100 (round up to whole number) **[7]**

iii) Define epidemiological triad: Agent, Host and Environment. **[3]**

⌘⌘⌘

Total No. of Questions : 5]

SEAT No. :

PC-4373

[Total No. of Pages : 3

[6347]-206

M.Sc.

STATISTICS

STS - 562 - MJ : Discrete Data Analysis (Equivalent Course)

(2023 Pattern) (Semester - II) (4 Credits)

Time : 3 Hours] ***[Max. Marks : 70***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions.** **[5 × 2 = 10]**

a) Give any two measures of association for contingency table.

b) Explain the concept of marginal Odd's ratio.

c) Define the terms joint independence and marginal independence.

d) Explain the terms with illustration :

i) Sensitivity

ii) Specificity.

e) Explain multinomial model.

Q2) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Estimate the parameter of Multinomial distribution using MLE method.

b) Explain likelihood ratio test and construct the confidence interval using it.

c) Construct the confidence interval for Binomial parameter using Score test.

d) Obtain the parameter of Poisson regression model using MLE method.

P.T.O.

Q3) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Write a short note on adjacent category model.

b) Explain the concept of generalized linear model for cluster responses.

c) Distinguish between population-averaged and subject-specific models for longitudinal data.

d) Explain how linear trend is alternative to independence.

Q4) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Write a short note on the Cochran-Mantel-Haenszel test for $2 \times 2 \times k$ contingency table.

b) Distinguish between population-averaged and subject-specific models for longitudinal data.

c) Write a short note on Baseline logit model.

d) When the 2000 General Social Survey asked subjects whether they would be willing to accept cuts in their standard of living to protect the environment, 344 of 1170 subjects said "yes." Estimate the population proportion who would say "yes" and conduct a significance test to determine whether a majority or minority of the population would say "yes." Also, Construct and interpret a 99% confidence interval for the population proportion who would say "yes".

Q5) **Attempt any 1 question out of 2 questions.** **[1 × 15 = 15]**

a) i) Compare the marginal modeling with Generalized Estimating Equations (GEE) in the context of analyzing correlated or clustered data and discuss their respective strengths, limitations and applications. **[8]**

ii) Refer the following Table on x = mother's alcohol consumption and Y = whether a baby has sex organ malformation. With scores (0, 0.5, 1.5, 4.0, 7.0) for alcohol consumption, ML fitting of the linear probability model has the output:

Parameter	Estimate	Standard error	Likelihood ratio 95% confidence limits	
Intercept	0.00255	0.0003	0.0019	0.0032
Alcohol	0.00109	0.0007	–0.0001	0.0027

State the prediction equation and interpret the intercept and slope. Also, use the model fit to estimate the probabilities of malformation for alcohol levels 0 and 7.0 and relative risk comparing those levels. **[7]**

b) i) A coin is flipped twice. Let Y = number of heads obtained, when the probability of head for a flip equal π. **[7]**

I) Assuming $\pi = 0.50$, specify the probabilities for the possible values for Y and find the distribution's mean and standard deviation.

II) Find the binomial probabilities for Y when π equals 0.60 and 0.40.

III) Suppose $y = 1$ and π is unknown. Calculate and sketch likelihood function. Comment on it.

ii) 800 boys are classified according to socio-economic status (S), whether they are participated in Boys Scout (B) and whether they have been labeled as Juvenile Delinquent (D) as follows:

Socio-economic status(S)	Boys Scout (B)	Delinquent (D)	
		Yes	No
Low	Yes	11	043
	No	42	169
Medium	Yes	14	104
	No	20	132
High	Yes	08	196
	No	02	059

Test the association between three variables. **[8]**

▽▽▽▽

Total No. of Questions : 4]

SEAT No. :

PC-4374

[Total No. of Pages : 2

[6347]-1001
M.Sc (Part - I)
STATISTICS
STS-501-MJ : Fundamentals of Analysis and Calculus
(2 credits) (2024 Pattern) (Semester - I)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) **Attempt all questions.** **[5 × 1 = 5]**

a) Write the conditions for L 'Hospital's rule.

b) What is the relationship between continuity and uniform continuity?

c) What is the definnition of a Cauchy sequence?

d) State the chain rule.

e) Find the limit infimum of $a_n = \sin x$.

Q2) **Attempt any TWO of the following:** **[2 × 5 = 10]**

a) Prove or disprove: Every Cauchy sequence is bounded.

b) Discuss convergence of log series. Also find it's radius of convergence.

c) Show that arbitrary union of open set is open.

Q3) **Attempt any TWO of the following:** **[2 × 5 = 10]**

a) State and prove mean value theorem.

b) Determine the convergence of the given series.

i) $\sum \frac{(-1)^{n-1} x^n}{n}$ ii) $\Sigma\, (-1)^n\, x^{2n}$

c) Define compact set. Show that [0,1] is compact set.

P.T.O.

Q4) Attempt any ONE of the following: **[1 × 10 = 10]**

a) i) Prove or disprove : If series $\Sigma\, a_n$ is convergent then series $\Sigma\,|\,a_n|$ is also convergent. **[4]**

ii) Define power series and radius of convergence. Obtain the radius of convergence for the following series: $\sum_{n=1}^{\infty}\frac{(-1)^n n^2 x^n}{5^n \sqrt{n^5}}$ **[6]**

b) i) Discuss Cauchy criterion for convergence with illustration. **[7]**

ii) Prove or disprove: Set of irrational numbers in uncountable. **[3]**

Total No. of Questions : 5]

SEAT No. :

PC-4375

[Total No. of Pages : 3

[6347]-1002
M.Sc. - I
STATISTICS
STS - 502 - MJ : Linear Algebra
(2023 Pattern) (Semester - I) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions :** **[5 × 2 = 10]**

a) Define null space and nullity.

b) Explain the term elementary operation of a matrix with example.

c) State any two properties of MP inverse.

d) Find the eigen values of $A = \begin{bmatrix} 7 & -1 \\ 2 & 6 \end{bmatrix}$.

e) Write down the quadratic form of $A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 2 \end{bmatrix}$.

Q2) **Attempt any Three of the following :** **[3 × 5 = 15]**

a) State and prove Cayley Hamilton theorem.

b) Let S be a set with two or more vectors in a vector space V. Then show that S is linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the rest of the vectors in S.

c) For which value of k will the following system have no solution and infinitely many solutions?

$4x + y + (k^2 - 14) = k + 2$

$x + 2y - 3z = 4$

$3x - y + 5z = 2$

d) Define g-inverse and obtain two distinct g-inverses for the following matrix. $A = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 3 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

P.T.O.

Q3) **Attempt any Three of the following :** **[3 × 5 = 15]**

a) Define linear combination of the vectors. Hence show that (–1, –4, 1) is a linear combination of $\overrightarrow{v_1} = (1,2,3)$ and $\overrightarrow{v_2} = (0,-1,2)$.

b) Explain Gram-Schmidt orthogonalization process.

c) Define the following terms.

i) Row rank

ii) Column Rank

iii) Basis

iv) Dimension

v) Spanning set

d) Define row space of a matrix. Hence find a basis for the row space of

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

Q4) **Attempt any Three of the following :** **[3 × 5 = 15]**

a) Define Vector space and its subspace with an example.

b) If $\vec{v}_1, \vec{v}_2 .... \overrightarrow{v_r}$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, ...\lambda_r$ then show that $\{\vec{v}_1, \vec{v}_2 .... \overrightarrow{v_r}\}$ is linearly independent.

c) Find the value of k for which the vector (1, –2, k) in $\mathbb{R}^3$ is a linear combination of vectors (3, 0, –2) and (2, –1, –5).

d) $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation given by $T(X_1, X_2, X_3) = (X_1 + X_2 + X_3, 2X_2 + X_3)$ Find the matrix of T with respect to the basis B = {(2, 2, 1), (0, 1, 0), (1, 0, 1)}.

Q5) **Attempt any One of the following :** **[1 × 15 = 15]**

a) i) Defme spectral decomposition of a real symmetric matrix. Hence obtain for $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$.

ii) Show that characteristic vector associated with distinct characteristic roots of a symmetric matrix is orthogonal.

[10+5]

b) i) Find geometric and algebraic multiplicity of a matrix given below. Also check whether A is diagonalizable?

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

ii) Let $S = \{\vec{v_1}, \vec{v_2}, \ldots \vec{v_r}\}$ be a set of vectors in a vector space V. Then show that L(S) is a subspace of V. Hence show that L(S) is the smallest subspace of V containing S.

[8+7]

▽▽▽▽

Total No. of Questions : 5]

SEAT No. :

PC-4376

[Total No. of Pages : 3

[6347]-1003
M.Sc. (Part - I)
STATISTICS
STS-503-MJ : Probability Distribution
(4 Credits) (2023 Pattern) (Semester - I)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) **Attempt all questions.** **[5 × 2 = 10]**

a) Let X_1 and X_2 be two independent random variables each having the same distribution : $P(X_i = k) = \frac{1}{2^{k+1}}$, $k = 0,1,2,....,$ then find $P(X_1 + X_2 = 5)$.

b) Define σ – *field*. How does it differ from a *field* of sets.

c) Let X be a random variable with MGF

$$M_X(t) = \frac{1}{12} + \frac{1}{6}e^t + \frac{1}{3}e^{2t} + \frac{1}{4}e^{-t} + \frac{1}{6}e^{-2t};\ t \in \mathrm{R}.\text{Then, calculate } 8\mathrm{E}(X).$$

d) If $E(X^2) < \infty$ then $V(X) \geq Var[E(X \mid Y)]$ with equality iff X is a function Y.

e) Define expectation of a non-negative random variable.

Q2) **Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Give an example of random variable which has exactly finite moment upto 5^{th} order but for which the higher order moments do not exist. Justify your answer.

P.T.O.

b) A continuous random variable X has cumulative distribution function

$$F(x)=\begin{cases}0, & x<0\\ 1-e^{-x^2}, & x\geq 0\end{cases}$$

Derive the probability density function of X and then find the mean, variance, median and mode of the distribution. Give a rough sketch of the probability density function.

c) Describe the relationship between a σ - *field* and a probability measure. How does a probability measure behave with respect to countable unions, intersections, and complements of sets in the σ-field? Illustrate your answer with a real-world example, such as rolling a die or flipping a coin.

d) Prove or disprove: The set of points of discontinuity of a distribution function is at most a countably infinite set.

***Q3)* Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) State and prove characteristic properties of bivariate distribution function.

b) Define a term quantile function? Discuss properties like monotonicity and right-continuity.

c) Let X, Y, Z are independent random variable distributed uniform on the interval (0,1). Find distribution of X + Y and compute $P\{X > YZ\}$.

d) Define the Stieltjes moment problem. Explain the relationship between moments and measures in the Stieltjes moment problem.

***Q4)* Attempt any 3 questions out of 4 questions.** **[3 × 5 = 15]**

a) Let $\underline{X}$ be a random vector with $N_n(\underline{0}, I_n)$ distribution. Show that the two quadratic forms $\underline{X}'A\underline{X}$ and $\underline{X}'B\underline{X}$ are independent iff $AB = 0$, where A and B are symmetric idempotent matrices.

b) Let X ~ Poisson(λ) distribution. Show that $P(P(s))$ is probability generating function of random variable Y, where $P(s)$ is probability generating function of random variable X. Obtain $P(Y = 0)$ and $E(Y)$.

c) What is a mixture distribution? Find the mixture of two normal distributions. Hence obtain its mean and variance.

d) Let the random vector $X = (X_1, X_2, X_3)$ have the joint pdf

$$f_X(x_1,x_2,x_3)=\begin{cases}\dfrac{81}{4}x_1^2x_2^2x_3^2, & -1\leq x_1,x_2,x_3\leq 1\\ 0, & \textit{Otherwise}\end{cases}$$

Then obtain the variance of the random variable $(X_1 + X_2 + X_3)$.

Q5) Attempt any 1 questions out of 2 questions. **[1 × 15 = 15]**

a) i) Let X_1, X_2, ..., X_3 be n i.i.d. N(0, 1) random variables. Obtain necessary and sufficient condition for independence of linear form $\underline{l}'\,\underline{X}$ and quadratic form $\underline{X}'A\underline{X}$ where $\underline{X} = \left(X_1, X_2, ..., X_n\right)'$. **[5]**

ii) Prove independence of $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $(n-1)S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2$ by using the result in (i) **[3]**

iii) Find the cumulative distribution function for the random variable having the probability density function $f(x) = 2\beta x e^{-\beta x^2}$, $x > 0$ where β is a positive constant. Hence, find the probability density of $Y = X^2$. **[7]**

b) i) Define non-central Chi-square distribution and obtain its characteristic function. **[6]**

ii) $(X, Y) \sim BVN\left(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2\right)$ then find $P(X > \mu_x, Y < \mu_y)$ **[7]**

iii) Define: Multivariate Beta and Poisson distribution. **[2]**

Total No. of Questions : 4]

SEAT No. :

PC-4377

[Total No. of Pages : 2

[6347]-1004
M.Sc. - I
STATISTICS
STS 510-MJ: Optimization Techniques
(2024 Pattern) (Semester - I) (2 Credits)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates :

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) Attempt all questions : **[5 × 1 = 5]**

a) State the Bellman's principle of optimality.

b) Write one use of zero-one integer programming problem.

c) Find the cut for the equation : $\frac{5}{8} = x_2 + \frac{1}{5}x_3 - \frac{3}{4}s_1$ (x_2 – *source row*).

d) What is cut capacity in maximal flow model?

e) Define state variable in dynamic programming.

Q2) Attempt any two of the following: **[2 × 5 = 10]**

a) Explain concept of branch and bound method. Also write its advantages.

b) Write the linear transformation of shortest route method.

c) Consider the following dynamic programming,

$f_n(1) = p_1 \log p_1 + p_2 \log p_2 + ... + p_n \log p_n$

Subject to constraint $p_1 + p_2 + ... + p_n = 1$

and $p_i \geq 0$ for all i.

Show that $f_n(1)$ is minimum when $p_1 = p_2 = ... = p_n = 1/n$

P.T.O.

Q3) Attempt any two of the following: **[2 × 5 = 10]**

a) State the advantages and disadvantages of dynamic programming problem.

b) The mathematical model for daily data of two products of manufacturing firm has the quadratic form as given below :

Max $Z = 12x + 21y + 2xy - 2x^2 - 2y^2$

Subject to constraints

$8 - y \geq 0$

$10 - x - y \geq 0$

$x, y \geq 0$

Find the solution to the problem by Beale's method.

c) Write the algorithm for minimum spanning tree problem.

Q4) Attempt any one question. **[1 × 10 = 10]**

a) i) Explain the term -Addition of a cut makes the previous non-integer optimal solution infeasible.

ii) State the applications of maximum flow problems.

b) i) Write the formulation of minimum cost flow problem.

ii) What is Dijkstra's shortest route algorithm? Write the steps in this algorithm.

♦ ♦ ♦

Total No. of Questions : 4]

SEAT No. :

PC-4378

[Total No. of Pages : 2

[6347]-1005
M.Sc. - I
STATISTICS
STS 511-MJ: Statistical Quality Control
(2024 Pattern) (Semester - I) (2 Credits)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt all questions : **[5 × 1 = 5]**

a) Explain the term; Fast Initial Response.

b) Give the expression for one-sided upper and lower CUSUIMs.

c) State the relation between parameters of tabular CUSUM and V-mask CUSUM.

d) A process is centred with specification limits 115 ± 10 and the estimated variance 25. Compute C_p and C_{pk}.

e) Define the signal-to-noise ratio.

Q2) Attempt any Two of the following: **[2 × 5 = 10]**

a) Explain the construction and working of EWMA chart for monitoring the process mean.

b) Obtain the relation between number of defectives and capability index C_p.

c) Discuss the advantages of CUSUM and EWMA chart over Shewhart Control charts.

P.T.O.

Q3) Attempt any two of the following: **[2 × 5 = 10]**

a) A process is in control with. $\overline{\overline{X}} = 120$ and $\overline{R} = 4$, n = 5. The process specifications are 115 ± 10. Compute C_p C_{pk} and C_{pm}. ($d_2 = 2.326$)

b) Write a short note on Gauge capability analysis

c) Explain the working of Synthetic control chart.

Q4) Attempt any one of the follwing : **[1 × 10 = 10]**

a) i) Explain the construction and working of Hotelling T^2 control chart for process mean vector when dispersion matrix is known and unkonwn.

ii) Explain the working of double sampling plan with curtailed inspection. **[6+4]**

b) i) Define the Average run length and give the expression for it for Shewhart Control chart.

ii) Using CUSUM chart, check whether the following process is under control or not, if target value is 175, $k = 0.5$, $h = 4.77$, $\hat{\sigma} = 4$. The observations are as follows; 160, 186, 190, 250, 158.5.

[6+4]

♦ ♦ ♦

Total No. of Questions : 4]

SEAT No. :

PC-4379

[Total No. of Pages : 2

[6347]-1006
M.Sc. - I
STATISTICS
STS - 512 - MJ : Actuarial Statistics
(2024 Pattern) (Semester - I) (2 Credits)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of statistical tables and scientific calculator is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) **Attempt all questions:** **[5 × 1 = 5]**

a) Suppose force of interest per annum is 0.07. Find d.

b) Define the curtate future life time random variable.

c) In how many years will a sum of money double itself at compound interest with effective rate $i = 0.05$?

d) Show that $e_x = P_x (1 + e_{x+1})$.

e) Explain the term : Deferment of benefit.

Q2) **Attempt any Two questions out of Three questions.** **[2 × 5 = 10]**

a) Suppose the life length random variable X is modeled by a distribution with survival function as specified below

$$S(x) = \begin{cases} 1-\dfrac{x^2}{100}, & 0 \le x \le 10 \\ 0, & otherwise \end{cases}$$

Find $F(x)$, ${}_2P_4$, probability density function of T (4) and median future life time of a person of age 5.

b) State the Gompertz's law of modelling life length random variable. Derive the distribution of $k(x)$ assuming that life length random variable is modelled by Gompert'z law.

c) Explain the terms:

i) Loss at issue random variable

ii) Equivalence principle of premiu

P.T.O.

Q3) Attempt any Two questions out of Three questions. **[2 × 5 = 10]**

a) Prove that, $A_x^{(m)} = \frac{i}{i^{(m)}} A_x$.

b) For a whole life insurance of 1 on (41) with death benefit payable at the end of year of death, you are given: $i = 0.05$, $= p_{40} = 0.9972$, $A_{41} - A_{40} = 0.00822$, ${}^2A_{41} - {}^2A_{40} = 0.00433$, Z is the present value random variable for this insurance, Find Var (Z).

c) Define annuity. Also, derive the expression of actuarial present value for n- year temporary life annuity and whole life annuity.

Q4) Attempt any One question out of Two questions. **[1 × 10 = 10]**

a) i) Assume that premiums are calculated on the basis of the equivalence principle. If ${}_kL$ is the prospective : loss random variable for a fully discrete whole life insurance of 1000 issued to (x). It is given that $A_x = 0.125$, $A_{x+k} = 0.4$, ${}^2A_{x+k} = 0.2$, $d = 0.05$. Then calculate E(${}_kL$) , Var(${}_kL$) and the aggregate reserve at time k for 100 policies of this type. **[7]**

ii) Write a short note on retrospective reserve. **[3]**

b) i) Define Annually decreasing n-year term insurance. **[2]**

ii) For a 2 year term insurance of 10 issued to (60). It is given that benefits are payable at the end of year of death and $i = 0.05$. Calculate Var(Z), given the following extract from a select and ultimate life table with select period 2. **[8]**

x	$l_{[x]}$	$l_{[x]+1}$	l_{x+2}	$x + 2$
60	78900	77200	75100	62
61	76400	74700	72500	63
62	73800	72000	69800	64

▽ ▽ ▽ ▽

Total No. of Questions : 5]

SEAT No. :

PC-4380

[Total No. of Pages : 2

[6347]-1007
M.Sc.
STATISTICS
STS 541-MJ: Research Methodology
(2024 Pattern) (Semester - I) (4 Credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt all questions : **[5 × 2 = 10]**

a) Explain convolution method briefly to generate random sample.

b) Define the following terms: i) Reflexive relation ii) Transitive relation

c) Define the term intersection of an indexed family of sets.

d) What is digit frequency test?

e) Write a short note on empirical research.

Q2) Attempt any Three out of 4 questions : **[3 × 5 = 15]**

a) Give the testing procedure to test the random number using Empirical Test.

b) Describe the stages involved in the research process. How do they relate to each other?

c) What is the purpose of a literature review in research? How do you conduct a literature review?

d) What are the different types of research designs? Explain the strengths and limitations of each.

Q3) Attempt any three questions out of four questions. **[3 × 5 = 15]**

a) Write a short note on alias method.

b) If A, B, and C are subsets of some universal set U, then prove that $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

c) If A, B, and C are subsets of some universal set U, then prove that $A - (B \cup C) = (A - B) \cap (A - C)$.

d) Let X be a random variable defined as $X = F^{-1}(U)$ where U be Uniform (0,1) random variable. Show that $X = F^{-1}(U)$ has distribution function $F(F^{-1}(U))$.

Q4) Attempt any three questions out of four questions. **[3 × 5 = 15]**

a) Give an algorithm for simulating a multivariate normal random variable.

b) Explain the advantages and disadvantages of Inverse Transformation Method.

c) Write a R-code to simulate a random sample from the distribution with density $f(x) = 3x^2$, $0 < x < 1$ using inverse transform method.

d) Suppose R is an equivalence relation on a set A. Then the set $\{[a]: a \varepsilon A)$ of equivalence classes of R forms a partition of A.

Q5) Attempt any one question out of two questions. **[1 × 15 = 15]**

a) i) Let A, B and C be sets. Then show that : **[7]**

$A \times (B \cup C) = (A \times B) \cup (A \times C)$

ii) For statements P. Q and R : **[8]**

A) Show that, $(P \vee Q) \rightarrow R \equiv (P \rightarrow R) \wedge (Q \rightarrow R)$.

B) The statement $\sim (P \rightarrow Q)$ is logically equivalent to $P \wedge \sim Q$.

b) i) ‘What is ggplot? Explain how it is useful for data visualization. **[3]**

ii) Use the rejection method to generate a random variable having density function, **[5]**

$f(x) = 20x(1 - x)^3, 0 < x < 1$.

iii) Let A, B and C be non-empty sets and assume that $f: A \rightarrow B$ and $g: B \rightarrow C$. Then show that :

A) If f and g are both injections, then $(g \circ f): A \rightarrow C$ is an injection.

B) If f and g are both surjections, then $(g \circ f): A \rightarrow C$ is surjection.

C) If f and g are both bijections, then $(g \circ f): A \rightarrow C$ is bijection. **[7]**

♦ ♦ ♦

Total No. of Questions : 4]

SEAT No. :

PC-4381

[Total No. of Pages : 2

[6347]-3001
M.Sc. - II
STATISTICS
STS - 601 - MJ : Probability Theory
(2024 Pattern) (Semester - III) (2 Credits)

Time : 2 Hours] ***[Max. Marks : 35***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of Statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions :** **[5 × 1 = 5]**

a) Define expectation and Moments of a random vector.

b) Define empirical distribution function of a random vector.

c) State Strong law of large numbers.

d) State Liapounov's form of Central limit theorem.

e) Define convergence in probability.

Q2) **Attempt any Two of the following :** **[2 × 5 = 10]**

a) State and prove Cr inequality.

b) State and prove Jordan decomposition theorem.

c) $X_n \xrightarrow{P} C \Leftrightarrow X_n \xrightarrow{L} C$ that is $F_n(x) \to F(x)$ where c is constant and

$$F(x) = \begin{cases} 0 & if\ x < c \\ 1 & if\ x \geq c \end{cases}$$

P.T.O.

Q3) Attempt any Two of the following : **[2 × 5 = 10]**

a) Consider a following distribution function

$$F(x) = \begin{cases} 0 & if \ x < 0 \\ \frac{x}{4} + \frac{1}{4} & if \ 0 \le x \le 1 \\ 1 & if \ x \ge 1 \end{cases}$$

Obtain the decomposition of $F(x)$.

b) What is cumulative distribution function state its properties for bivariate random vector.

c) State and prove Bore1 0-1 law.

Q4) Attempt any One of the following : **[1 × 10 = 10]**

a) State and prove Khinchin's Weak law of large numbers.

b) i) Show that, $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X$.

ii) Compare convergence in distribution and almost sure convergence.

▽▽▽▽

Total No. of Questions : 7] **SEAT No. :**

PC-4382 **[Total No. of Pages : 3**

[6347]-3002

M.A./M.Sc (Statistics)

STS602-MJ: STOCHASTIC PROCESSES

(2023 & 2024 Pattern) (Semester - III) (4 credits)

Time : 3 Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions.** **[5 × 2 = 10]**

a) Write any applications of stochastic processes.

b) Define Weiner process and its Properties.

c) Explain extinction probabilities.

d) Give the Transition Probability Matrix (TPM) for random walk with absorbing barries and birth-death chain.

e) Let $\{X_n$, n=0, 1,2,....$\}$ be a Markov chain with state space S = {0, 1, 2,} and transition probability matrix Compute $P[X_2 = 1 \mid X_0 = 2]$

$$P = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0.4 & 0.6 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

P.T.O.

Q2) Attempt any Three of the following : **[3 × 5 = 15]**

a) Explain the postulates of Yule - Furry process and find an expression for $P_n(t)$.

b) Derive the relation between Poisson Process and Binomial Distribution.

c) Define a renewal process also state the elementary renewal theorem with its application.

d) Suppose that probability of dry after rainy day is $\frac{3}{4}$ and that rainy day after dry day is $\frac{2}{3}$ Let X_n be state of process after n^{th} day write state space and one step TPM of process $\{X_n, n \geq 1\}$ also find probability that second day is dry given that initial day is dry.

Q3) **Attempt any Three of the following :** **[3 × 5 = 15]**

a) Let $\{X_n, n = 0,1.2,....\}$be a Branching process with $X_0 = 1$. Find the mean and variance of X_n in terms of those of the offspring distribution.

b) Write short notes on the following

i) Interarrival time in Poisson process

ii) Periodic States

c) Define Brownian motion. Write it as a function of Standard Brownian motion. Define Geometric and integrated Brownian motion. Show that Brownian motion process can be obtained as the limit of a random walk. state the assumptions and results which you have used.

d) In a Poisson process with a rate of 4 arrivals per minute, what is the probability that there will be exactly 3 arrivals in a 30 -second interval?

Q4) **Attempt any Three of the following :** **[3 × 5 = 15]**

a) Let {B(t), t > 0} is a standard Brownian motion then compute the conditional distribution of B(s) providedB(t_1) = A and B(t_2) = B, where $0 < t_1 < s < t_2$.

b) Show that State j is Persistent if and only if $\sum_{n=0}^{\infty} pij^{(n)} = \infty$.

c) Customer arrive at a service station according to a poisson process of rate $\lambda = 3$ customer per hour. Suppose 3 customers arrived during the 1^{st} 30 minutes? What is the probability that only 4 customers arrived during the hour? What is the probability that there is no customer in 1^{st} 10 minutes.

d) Derive the generating function relations satisfied by a Branching process.

Q5) **Attempt any One of the following :** **[1 × 15 = 15]**

a) i) A Markov chain on states {1,2,3,4} has transition probability matrix

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

Classify the states are as recurrence or transient. **[8]**

ii) Obitan E[X(t)], where X(t) is a linear birth and death process. **[7]**

b) i) If {X_n, n=0,1,2,...} is the Galton -Watson Branching process, obtain E(X_n) and Var (X_n). **[8]**

ii) Find the expected duration of game in the gambler ruin problem.**[7]**

Total No. of Questions : 5] **SEAT No. :**

PC-4383 **[Total No. of Pages : 3**

[6347]-3003

M.Sc. (Part - II)

STATISTICS

STS - 603 - MJ : Design and Analysis of Experiments

(2023 & 2024 Pattern) (Semester - III) (4 Credits)

Time : 3 Hours] ***[Max. Marks : 70***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt each of the following :** **[5 × 2 = 10]**

a) Explain 3 basic principles of design of experiments.

b) Check whether following block design is connected:

Block-I	Block-II	Block-III
A	A	C
B	B	
	D	

c) Write the hypothesis and ANOVA table of One-way ANOVA.

d) Define BIBD, with usual notations prove that : $\lambda(v - 1) = r\,(k - 1)$

e) Define the following :

i) Rotatable CCD

ii) Spherical CCD

P.T.O.

Q2) Attempt any three of the following : **[3 × 5 = 15]**

a) Explain One-way ANOVA with repeated measures.

b) Obtain parameters of the following PBIBD:

Blocks	Treatments		
1	1	2	5
2	1	3	6
3	1	4	7
4	8	9	10
5	6	7	10
6	3	4	10
7	2	3	8
8	2	4	9
9	5	7	9
10	5	6	8

c) What is Confounding? Why confounding is used even at the loss of information on confounded effects. Explain types of confounding.

d) What is Simplex Lattice Design. Explain a [4, 3] Simplex Lattice Design.

Q3) Attempt any three of the following : **[3 × 5 = 15]**

a) Construct one-half fraction of a 2^4 design with highest possible resolution. Write down its alias structure.

b) Explain following method for comparing pairs of treatment means:

i) Duncan's Multiple Range Test

ii) Dunnets Test

c) Give the statistical analysis of 3^2 factorial experiment in terms of linear and quadratic effects.

d) Write a note on response surface methodology.

Q4) Attempt any three of the following : **[3 × 5 = 15]**

a) Give the statistidal analysis of 2^3 factorial experiment through RBD.

b) Given the following block, find out the interaction confounded

acde	bcd	e	abec	ad	bde	ab	c

c) Explain resolution III, IV and V chsigns with example.

d) Write a note on Taguchi design.

Q5) **Attempt any one of the following :**

a) i) A group of 6 different rats with their swim speed on different days as a result of a change in the water temperature given in the following table: **[8]**

	Water temperature			
Rats	20ºC	24ºC	27ºC	32ºC
1	39	38	34	33
2	29	25	20	20
3	36	37	29	24
4	25	29	18	19
5	21	27	24	22
6	34	33	30	31

The question is whether or not changing the water temperature affects the swimming speed of rats. Use Friedman test for analysis.

ii) Explain the following Tests: **[7]**

I) Levene's Test

II) Bartlett's Test

b) i) Show that RBD and BIBD are connected block design. **[8]**

ii) Analyze the following data : **[7]**

Block - I	Block - II	Block - III	Block - IV
A = 4	C = 3	A = 5	B = 7
B = 6	D = 4	D = 4	C = 3

▽ ▽ ▽ ▽

Total No. of Questions : 4]

SEAT No. :

PC-4384

[Total No. of Pages : 3

[6347]-3004

M.Sc. (Part - II) (Statistics)

STS-610 MJ : SURVIVAL ANALYSIS

(2023 & 2024 Pattern) (Semester - III) (2 Credits)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Choose the correct alternative to each of the following questions :** **[5 × 1 = 5]**

i) The hazard rate is constant for ________.

a) Gamma (4,4) b) U (0, 4)

c) Exponential(4) d) Weibull (4, 2)

ii) The Cauchy functional equation holds true for the ______

a) Gamma distribution b) Cauchy distribution

c) Weibull distribution d) Exponential distribution

iii) In an experiment with Type I censoring scheme, the random variable of interest is the ________.

a) termination time of the experiment

b) no. of units failed before termination time

c) no. of units involved in the experiment

d) life time of the units at the experiment.

iv) A lifetime distribution belongs to IFRA class iff,

a) Hazard rate is increasing function of time ‘t’

b) Hazard rate is decreasing function of time ‘t’

c) Hazard rate average function is decreasing function of time ‘t’

d) Hazard rate average function is increasing function of time ‘t’

P.T.O.

v) The Kaplan - Meier Estimator of the survival function is also known as :

a) Sum-Limit estimator b) Sum-Product estimator

c) Product-Limit estimator d) Maximum-Sum estimator

Q2) Attempt any 2 questions out of 3 questions : [2 × 5 = 10]

a) Prove the following implications :

IMRL $\rightarrow$ NWUE and NBUE $\rightarrow$ HNBUE

b) Describe the Cox's Proportional Hazards model of regression for complete data.

c) Show that 'no ageing' property is characterized by exponential equilibrium distribution function.

Q3) Attempt any 2 questions out of 3 questions : [2 × 5 = 10]

a) Derive the maximum likelihood estimator of parameter of exponential distribution for the type II censored data.

b) The life time (T) of a certain component has hazard rate r (t) = 2; t > 0. Find the mean residual life function and compare it with expected life E (T). Comment on the result.

c) Find estimator of variance of the actuarial estimator of the survival function.

Q4) Attempt any 1 question out of 2 questions : [1 × 10 = 10]

a) i) State and prove Cauchy functional equation.

ii) The following failure and censor times (in operating hrs.) were recorded on 12 turbine

vanes: 142, 149, 320, 345+, 560, 805, 1130+, 1720, 2480+, 4210+, 5280, 6890.

(+ indicates censored observation). Censoring was a result of failure mode other than wear out. Compute the Kaplan-Meier estimate of the survival function.

[5+5]

b) i) Prove that F is IFR if and only if equilibrium distribution function is concave function of t

ii) Following table shows the failure time of two machines, new and old.

	Failure times (day)
New machine	250, 476+, 355, 200, 355+
Old machine	191, 563, 242, 285, 16, 16, 16, 257, 16

(+ indicates censored times).

Test whether the new machine is more reliable than the old one by using log rank test.

⌘⌘⌘

Total No. of Questions : 4]

SEAT No. :

PC-4385

[Total No. of Pages : 2

[6347]-3005
M.Sc.
STATISTICS
STS - 611- MJ : Asymptotic Inference
(2023 & 2024 Pattern) (Semester - III) (2credits)

Time : 2 Hours] ***[Max. Marks : 35***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) **Attempt all questions.** **[5 × 1 = 5]**

a) Give Crammers regularity conditions.

b) What is the relation between CAN estimator and MLE in case of one parameter exponential family?

c) Define BAN estimator.

d) Define joint consistency.

e) State one parameter Canonical Form.

Q2) Attempt any TWO of the following: **[2 × 5 = 10]**

a) Find CAN estimator of θ when random sample of size n is drawn from $(U(\theta,\theta + 1))$

b) Explain in detail locally most powerfull test with example.

c) Discuss Super - efficient estimator with illustration.

P.T.O.

Q3) Attempt any TWO of the following: **[2 × 5 = 10]**

a) Explain likelihood ratio test and Wald test in briefly.

b) Consider density function $f(x,\theta) = \theta\, x^{\theta}\, e^{-x^{\theta}}$; $x > 0$ and $\theta > 0$. Obtain CAN estimator for θ based on percentile method. Also obtain its asymptotic variance.

c) Let $X_1, X_2, ... X_n$ is random sample from U (θ – 1, θ + 1). Find maximum likelihood estimate of θ.

***Q4)* Attempt any ONE of the following:** **[1 × 10 = 10]**

a) i) Prove consistency property of MLE of parameter o distribution belongs to one parameter exponential family. **[7]**

ii) State Cramer - Huzurbazar theorem. **[3]**

b) i) Explain Wald test in briefly. **[3]**

ii) Explain method of percentile to obtain CAN estimator. Also find CAN estimator of θ when random sample of size n is drawn from $(U(\theta,\theta + 1))$ **[7]**

Total No. of Questions : 4]

SEAT No. :

PC-4386

[Total No. of Pages : 2

[6347]-3006
M.Sc. (Part - II)
STATISTICS
STS 612-MJ: Machine Learning
(2024 Pattern) (Semester - III) (4 Credits)

Time : 2 Hours] *[Max. Marks : 35*

Instructions to the candidates :

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt all questions : **[5 × 1 = 5]**

a) Write the main types of feature selection techniques?

b) What is the Fl-score, and why is it important?

c) What is pruning in the context of Decision Trees?

d) Define Average Linkage in Agglomerative Hierarchical Clustering.

e) What are common methods to handle missing data in Machine Learning?

Q2) Attempt any Two of the following: **[2 × 5 = 10]**

a) Explain the different types of Machine Learning. What are the key differences between them?

b) How does statistics play a role in Machine Learning, and why is it critical?

c) Given a set of 5 objects {A, B, C, D, E} and their distance matrix, create clusters using Agglomerative Hierarchical Clustering.

	A	B	C	D	E
A	0	2	6	10	9
B	2	0	5	9	8
C	6	5	0	4	5
D	10	9	4	0	3
E	9	8	5	3	0

P.T.O.

Q3) Attempt any two of the following: **[2 × 5 = 10]**

a) What are activation functions in Machine Learning?

b) How classifier performance is assessed using a confusion matrix and what related measures can be derived from it?

c) A company wants to predict whether customers will purchase a product (Buy) based on three features: Age (Young, Middle-aged, Senior) Income (Low, Medium, High) ,Student (Yes, No). Find the root node by using ID3 Algorithm.

Customer	Age	Income	Student	Buy (Target)
1	Young	High	No	No
2	Young	High	Yes	Yes
3	Middle-aged	High	No	Yes
4	Senior	Medium	No	Yes
5	Senior	Low	No	No
6	Senir	Low	Yes	Yes
7	Middle-aged	Low	Yes	Yes
8	Young	Medium	No	No
9	Young	Low	Yes	Yes
10	Senior	Medium	Yes	Yes
11	Young	Medium	Yes	Yes
12	Middle-aged	Medium	No	Yes

Q4) Attempt any one of the follwing : **[1 × 10 = 10]**

a) i) What is a Decision Tree in Machine Learning? Explain how it works and the key components of a Decision Tree.

ii) Explain the concepts of Ensemble Learning. How do Bagging and Boosting differ in terms of methodology and objectives?

b) i) What are Artificial Neural Networks (ANNs), and how do they function in Machine Learning?

ii) What is the Apriori Algorithm and how is it used for association rule mining?

♦ ♦ ♦