Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - I) (New) (NEP CBCS) Examination: October/November - 2025
Distribution Theory (2329101)

Day & Date: Wednesday, 29-10-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose the correct alternative. (MCQ) **08**

1) Which one of the following is not an order statistic?
   a) Maximum b) Minimum
   c) Mean d) Median

2) Let $X_1, X_2, \dots, X_n$ be a random sample from $pdf f_X(x)$ and $Y_1 \le Y_2 \le \cdots \le Y_n$ be its order statistics. If pdf of Z is $n[F_X(z)]^{n-1} f_X(z)$ then Z is _____.
   a) sample median b) sample range
   c) smallest observation d) largest observation

3) Suppose $(X_1, X_2, \dots, X_k)$ is a multinomial random variate then $Cov(X_i, X_j), i = j = 1,2, \dots, k, i \neq j$ is _____.
   a) $n\, p_i$ b) $n\, p_i p_j$
   c) $-n\, p_i p_j$ d) $n^2\, p_i p_j$

4) If X is distributed as $B(n, p), 0 < p < 1$ then _____.
   a) $-X$ is distributed as $B(-n, p)$
   b) $n - X$ is distributed as $B(n, 1-p)$
   c) $X - k$ is distributed as $B(n-k, p)$
   d) $1/X$ is distributed as $B(n, 1/p)$

5) Let X and Y are independent random variables with $N(0,1)$. The distribution of $(Y/X)^2$ is _____.
   a) Normal b) χ^2
   c) F d) t

6) Let X has a continuous distribution with $cdf\ F(x)$, then the distribution of $Y = F(x)$ is _____.
   a) exponential b) uniform
   c) normal d) none of these

7) If $X > 0$ then _____.
   a) $E[\sqrt{X}] \leq \sqrt{E(X)}$ b) $E[\sqrt{X}] \geq \sqrt{E(X)}$
   c) $E[\sqrt{X}] = \sqrt{E(X)}$ d) none of these

8) Which of the following is not a scale family?
   a) $U(0,\theta)$ b) $U(0,1)$
   c) $N(0,\sigma^2)$ d) $Exp(\theta)$

B) Fill in the blanks. **04**

a) The moment generating function of Bernoulli distribution is _____.
b) If X and Y are two independent random variables the $\rho(X,Y) =$ _____.
c) Let X and Y be independent random variable each having the $U(0,1)$ distribution. Then $Var\ (X + Y)$ is equal to _____.
d) Let $X_1, X_2, \ldots, X_n$ are *iid* $N(0,1)$ random variables and if $Y = \sum_{i=1}^{n} X_i^2$ the variance of Y is _____.

Q.2 Answer the following. (Any Six) **12**

a) Define a distribution function of a bivariate random vector (X,Y).
b) Define location-scale family. Give an example.
c) Define non-central chi-square distribution.
d) Define order statistic. Write the joint pdf of any two order statistics.
e) State the Jensen inequality.
f) Define symmetric random variable. Give one example of symmetric random variable.
g) Define bivariate Poisson distribution.
h) State the relation between distribution function of a continuous random variable and uniform random variable.

Q.3 Answer the following. (Any Three) **12**

a) Let X has $N(0,1)$ distribution. Obtain the pdf of $Y = |X|$.
b) Let X has $Poisson\ (\lambda)$ distribution. Obtain the MGF of X.
c) Write a note on mixture of distributions.
d) Let F be a distribution function of random variable X. Define $G(x) = [F(x)]^n, n$ is positive integer. Examine $G(x)$ to be a distribution function.

Q.4 Answer the following. (Any Two) **12**

a) Let (X, Y) be a discrete bivariate random vector. Define
- i) Joint p.m.f of (X, Y).
- ii) Marginal p.m.f. of X and marginal p.m.f. of Y.
- iii) Independence of X and Y.

b) State and prove Minkowski's inequality.

c) Let X and Y are jointly distributed with pdf

$$f(x,y) = \begin{cases} k & (x+2y), 0 < x < 2, 0 < y < 1 \\ 0, & otherwise \end{cases}$$

Find marginal distributions of X and Y.

Q.5 Answer the following. (Any Two) **12**

a) Let (X, Y) has $BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Obtain the conditional distribution of Y given $X = x$.

b) Obtain the MGF of multinomial distribution with k cells. Hence show that pmf of i^{th} component X_i is $B(n, p_i)$.

c) Let X is a non-negative continuous random with distribution function $F(x)$.

If $E(X)$ exist then show that $E(X) = \int_0^\infty [1 - F(u)]du$.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - I) (New) (NEP CBCS)
Examination: October/November – 2025
Estimation Theory (2329102)

Day & Date: Friday, 31-10-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose the correct alternative. **08**

1) Exponential distribution with pdf $f(x,\theta) = e^{-(x-\theta)}, x \geq \theta, \theta \in R$ is a member of _____.
   a) one parameter exponential family
   b) Pitman family
   c) power series family
   d) None of the above

2) Let $X_1, X_2, \ldots, X_n$ is a random sample of size n from $U(0,\theta)$ distribution then unbiased estimator of θ is _____.
   a) $2\bar{X}$ b) $\frac{\bar{X}}{2}$
   c) $\bar{X}$ d) $\sqrt{\bar{X}}$

3) If T_n is sufficient statistic for θ based on random sample of size n, then $\frac{\partial \log L}{\partial \theta}$ is a function of _____.
   a) θ only b) T_n only
   c) both T_n and θ d) None of the above

4) If T_n *is* an estimator of θ, then Cramer-Rao's inequality provides a lower bound on _____.
   a) $Var(T_n)$ b) $E(T_n)$
   c) $Max(T_n)$ d) $Min(T_n)$

5) Suppose T_n sufficient for θ. Then $g(T_n)$ is sufficient for $g(\theta)$ if _____.
   a) g is a real valued function
   b) g is a continuous function
   c) g is one-to-one function
   d) g is a bounded function

6) Bayes estimator of a parameter under squared error loss function is _____.
   a) posterior mean b) posterior median
   c) posterior mode d) posterior variance

7) If an estimator T_n of population parameter θ converges in probability to θ as n tends to infinity is said to be ____.
   a) sufficient b) consistent
   c) efficient d) unbiased

8) MLE of σ^2 in $N(\mu, \sigma^2)$ distribution when μ is known is _____.
   a) $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ b) $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
   c) $\frac{1}{N}\sum_{i=1}^{n} X_i^2$ d) $\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$

B) Fill in the blanks. **04**

1) Let $X_1, X_2, \ldots, X_n$ is a random sample from $U(\theta - 1, \theta + 1), \theta > 0$, moment estimator of θ is ____.
2) Conditional distribution of random variable θ given $X = x$ is called _____ distribution.
3) Let $X_1, X_2, \ldots, X_n$ *be* a random sample from $N(0, \sigma^2)$. Then sufficient statistic for σ^2 is —
4) For Cauchy distribution with location θ, the consistent estimator of θ is _____.

Q.2 Answer the following. (Any Six) **12**

a) Define sufficient statistic.
b) Define posterior distribution. Illustrate with one example.
c) Define Fisher information matrix.
d) Give an example of consistent estimator which is not MLE.
e) State necessary and sufficient condition for an estimator to be UMVUE.
f) Define maximum likelihood estimator (MLE).
g) State Chapman-Robbins-Kiefer inequality.
h) Define m-parameter exponential family of distributions.

Q.3 Answer the following. (Any Three) **12**

a) State and prove invariance property of MLE.
b) Let $X_1, X_2, \ldots, X_n$ be iid $U(0, \theta)$. Find moment estimator of θ.
c) Examine whether $\{B(n, \theta), 0 \leq \theta \leq 1\}$ family is complete.
d) Define power series family of distributions. Show that Poisson distribution belongs to power series family of distributions.

Q.4 Answer the following. (Any Two) **12**

a) State and prove Cramer-Rao inequality with necessary regularity conditions.
b) Define Fisher information in a single observation. Find the same for $B(n, \theta)$ distribution, when n is known.
c) State and prove Lehmann-Scheffe theorem.

Q.5 Answer the following. (Any Two) **12**

a) Define a consistent estimator for a vector parameter. Show that joint consistency is equivalent to marginal consistency.

b) Let $X_1, X_2, \ldots, X_n$ *is* a random sample from Poisson (λ) distribution. For estimating λ using quadratic loss function, prior density of λ is $\pi(\lambda) = e^{-\lambda}, \lambda > 0$. Derive Bayes estimator of λ.

c) Obtain UMVUE of $p(1-p)$ based on a random sample of size n from $B(1, p)$ distribution.

Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - I) (New) (NEP CBCS) Examination: October/November – 2025
Statistical Mathematics (2329107)

Day & Date: Monday, 03-11-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose the correct alternative. (MCQ) **08**

1) Set of all rational is _____.
   a) Countable b) Uncountable
   c) Finite d) None of the above

2) If A is symmetric matrix, then which of the following are symmetric matrices?
   a) A^T b) $A - I$
   c) $A + I$ d) All of the above

3) If for matrices A and B, the matrix AB is a diagonal matrix, then _____.
   a) A and B both are diagonal matrices
   b) At least one of A and B is diagonal matrix
   c) A and B may or may not be diagonal matrices.
   d) A and B are skew symmetric matrices

4) Rank of matrix A is _____.
   a) Number of independent rows in A
   b) Dimension of the row space of A
   c) Dimension of the column space of A
   d) All of the above

5) The smallest sub-space containing finite set of vectors (S) is _____.
   a) Superclass of S b) Span of S
   c) Subset of S d) Basis of S

6) Integration of a product of functions can be solved using _____.
   a) Leibnitz rule b) Integration by parts
   c) Taylor's method d) None of these

7) The $\alpha(.)$ function in R-S integral is _____.
   a) Always non negative
   b) Always monotonic non-increasing
   c) Always monotonic non-decreasing
   d) Always constant

8) A sequence $S_n = (-1)^n \; n \; \epsilon N$ is _____ sequence.
   a) Divergent b) Convergent
   c) Oscillatory d) none of these

B) Fill in the blanks. **04**

1) Least upper bound is also called as _____.
2) If $\sum_{n=1}^{\infty} a_n$ converges, then $lim_{n\to\infty} \; a_n =$ _____.
3) If A is a 5×5 matrix with rank 4, then determinant of A is _____.
4) The rank of identity matrix of order 4 is _____.

Q.2 Answer the following. (Any Six) **12**

a) Define countable set.
b) Is countable union of countable sets always open?
c) Define diagonal matrix.
d) Define sequence of real numbers.
e) What do you mean by monotonic sequence?
f) Define transpose of a matrix.
g) Define addition of two matrices.
h) Define inverse of a matrix.

Q.3 Answer the following. (Any Three) **12**

a) Show that countable union of countable sets is countable.
b) What do you mean by basis of a vector space?
c) Discuss the concept of rank of a matrix.
d) Discuss continuity of a function over a set.

Q.4 Answer the following. (Any Two) **12**

a) Describe row reduced form of a matrix in detail. Also illustrate with the help of example.
b) Prove or disprove: A monotonic increasing bounded sequence always converges.
c) Discuss-
   i) Taylor's theorem.
   ii) Mean Value Theorem.

Q.5 Answer the following. (Any Two) **12**

a) How Riemann-Stieltje's integration is found out? Discuss in brief.
b) Explain the idea of G-inverse. Also illustrate with the help of example.
c) Explain the limit superior and limit inferior of a sequence of real numbers.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - I) (New) (NEP CBCS) Examination: October/November - 2025
Research Methodology in Statistics (2329103)

Day & Date: Thursday, 06-11-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose the most correct alternative. (MCQ) 08

1) A sampling _____ is the overall plan designed to achieve the survey objectives effectively.
   a) Frame b) Scheme
   c) Strategy d) Method

2) _____ is not a characteristic of good research.
   a) Systematic b) Subjective
   c) Logical d) Empirical

3) A _____ sampling scheme selects samples using a fixed interval from a randomly chosen starting point.
   a) Random b) Quota
   c) SRSWOR d) Systematic

4) Which of the following is not a key component of a sampling strategy?
   a) Defining the population and sample size.
   b) Identifying the sampling frame.
   c) Selecting the survey questions.
   d) Choosing the sampling method

5) _____ is the difference between research methods and research methodology?
   a) Methods are tools; methodology is the overall strategy
   b) Methods are theoretical; methodology is practical
   c) Methods focus on outcomes; methodology focuses on processes
   d) There is no difference between the two

6) _____ basic principles of experimental design?
   a) Randomization, replication, and control
   b) Objectivity, accuracy, and analysis
   c) Flexibility, simplicity, and precision
   d) Sampling, data collection, and interpretation

7) In double sampling, supplementary information is often used to improve estimates through _____ estimators.
   a) Regression b) Ratio
   c) Both A and B d) None of the above

8) The _____ method selects a sample of size two for PPSWOR and provides an unbiased estimate of the population mean.
   a) Des Raj b) Murthy's
   c) Horvitz Thompson d) Lahiri's

B) Fill in the blanks. **04**

1) The procedure is generalized to three or more stages and is then termed as _____.
2) The _____ method is used to allocate sample sizes across strata to minimize overall variance.
3) The primary objective of research is to _____ problems and contribute to knowledge.
4) Report writing is significant because it helps to _____ research findings for stakeholders.

Q.2 Answer the following. (Any Six) **12**

a) Define systematic sampling.
b) Define double sampling.
c) Define treatments.
d) What is the need for sampling?
e) Define conceptual research.
f) Define sampling frame.
g) Define experiment unit.
h) Write the difference between SRS and varying probability scheme.

Q.3 Answer the following. (Any Three) **12**

a) Prove: With usual notations, in SRS the bias of regression estimator $\bar{y}_l$ is, $bias\ (\bar{y}_l) = -cov(\bar{x}, b)$
b) Discuss sampling and non-sampling errors.
c) Explain in detail the literature survey.
d) What are the criteria of good research?

Q.4 Answer the following. (Any Two) **12**

a) Explain in detail probability sampling and their types.
b) Write the methods of PPS sampling with replacement in detail.
c) Discuss the meaning of the research in detail.

Q.5 Answer the following. (Any Two) **12**

a) Explain in detail the types of research.

b) Obtain Murthy's unordered estimator corresponding to Des Raj's ordered estimator for the sample size 2.

c) Prove: In simple random sample an approximate value of bias of $\hat{R} = \frac{\bar{y}}{\bar{x}}$ is given by,

i) $B(\hat{R}) \approx RC.V(\bar{x})[C.V(\bar{x}) - \rho C.V(\bar{y})]$

ii) $B(\hat{R}) \approx \frac{1-F}{n}(C_{xx} - \rho C_{yx})R$

Where $C_{xx} = C_{x^2}, C_{yx} = C_y C_x, C_x = \frac{S_x}{\bar{x}}, C_y = \frac{S_y}{\bar{y}}$

Seat No.	

Set P

M.Sc. (Statistics) (Semester - II) (New) (NEP CBCS) Examination: October/November – 2025
Stochastic Processes (2329201)

Day & Date: Tuesday, 28-10-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

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

Q.1 A) Choose correct alternative. (MCQ) **08**

1) A Markov chain is completely specified by _____ and TPM.
a) States b) State space
c) Initial distribution d) None of these

2) The collection of all possible states of a stochastic process is called as _____.
a) State Space b) Time Space
c) Chain space d) All of these

3) A non-null recurrent aperiodic state is also called as _____.
a) Transitive state b) Binomial state
c) Ergodic state d) None of these

4) Addition of two independent Poisson processes is _____.
a) a Poisson process
b) may or may not be Poisson process
c) a Bessel process
d) None of these

5) For a symmetric random walk, probability 'p' of positive jump is _____.
a) 0.25 b) 0.5
c) 1 d) 0

6) The process $\{X(t), t > 0\}$, where $X(t) =$ number of particles in a room at time t, is an example of _____ stochastic process.
a) discrete time continuous state space
b) discrete time discrete state space
c) continuous time continuous state space
d) continuous time discrete state space

7) If states i and j are communicating states, then _____.
a) state i leads to state j b) state j leads to state i
c) either (a) or (b) d) both (a) and (b)

8) Branching process is an example of _____.

a) Discrete time discrete state space stochastic process
b) Discrete time continuous state space stochastic process
c) Continuous time discrete state space stochastic process
d) Continuous time continuous state space stochastic process

B) Fill in the blanks. **04**

a) If period of a state is one, then the state is called as _____.
b) If the probability of ultimate first return $F_{ii} < 1$, then the state i is _____.
c) If X_n denotes number of active cases of COVID on n^{th} day, then $\{X_n\}$ is _____ time discrete state space stochastic process.
d) A finite Markov chain which contains only one communication class is called as _____.

Q.2 Answer the following. (Any Six) **12**

a) Define and illustrate state space of a stochastic process.
b) Define recurrent state.
c) Define periodic state.
d) Define transient state.
e) Define branching process. Also write expression for the mean of the population size at n^{th} generation.
f) Define stochastic process.
g) Define periodic and aperiodic state.
h) Define first return probability for a state.

Q.3 Answer the following. (Any Three) **12**

a) Obtain the distribution of inter-arrival times of Poisson Process.
b) Define and illustrate communicating states.
c) State and prove Chapman-Kolmogorov equation for Markov chain.
d) Discuss the classification of stochastic process based on state space and time space.

Q.4 Answer the following. (Any Two) **12**

a) Classify the states of random walk model.
b) A Markov chain with state space $S = \{1,2,3\}$ has tpm

$$\begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.8 & 0.1 \end{bmatrix}$$

It is known that the process has started with the state $X_0 = 2$

i) $P(X_1 = 2)$
ii) $P(X_2 = 3)$
iii) $P(X_0 = 1)$
iv) $P(X_3 = 2/X_1 = 1)$

c) Discuss stationary distribution of a Markov chain in detail. Illustrate with the help of example.

Q.5 Answer the following. (Any Two) **12**

a) Prove that, Markov chain is completely specified by one step t.p.m. and initial Distribution.

b) Write down the algorithm for the simulation of Poisson process and branching process.

c) Give two definitions of Poisson Process. Show that addition of two Poison processes is a Poisson process.

Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - II) (New) (NEP CBCS) Examination: October/November - 2025
Theory of Testing of Hypotheses (2329202)

Day & Date: Thursday, 30-10-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

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

Q.1 A) Choose the correct alternative. **08**

1) In usual notations, which one of the following is the probability of rejecting H_0 when H_1 is true?

a) α b) β
c) $1-\alpha$ d) $1-\beta$

2) Based on random sample of size n from $N(0,\sigma^2)$, the distribution possesses MLR property in _____.

a) $\bar{X}$ b) ΣX_i
c) ΣX_i^2 d) $\Sigma(X_i-\bar{X})^2$

3) Which of the following is a simple hypothesis for $N(\mu,\sigma^2)$?

a) $H_0: \mu = 5, \sigma = 2$ b) $H_0: \mu = 10$
c) $H_0: \mu = 0, \sigma > 1$ d) $H_0: \mu \neq 3, \sigma = 1$

4) Let $f_\theta, \theta \in \Theta = \{\theta_0, \theta_1\}$ Then MP test is based on _____.

a) $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$
b) $H_0: \theta_0 < \theta < \theta_1$ against $H_1: \theta \leq \theta_0$ or $\theta > \theta_1$
c) $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$
d) $H_0: \theta \leq \theta_0$ or $\theta > \theta_1$ against $H_1: \theta_0 < \theta < \theta_1$

5) Consider the testing problem for $N(\theta, 1)$.

i) $P_1: H_0: \theta = \theta_0$ against $H_1: \theta > \theta_0$
ii) $P_2: H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$

The UMP level α test exists _____.

a) for P_1 but not for P_2 b) for P_2 but not for P_1
c) for both P_1 and P_2 d) neither P_1 nor P_2

6) For testing simple versus simple hypotheses MP and LRT tests are _____.

a) Different
b) the same
c) not comparable
d) equivalent in size but not with respect power.

7) The family of confidence sets $\{S(\underline{x})\}$ for parameter θ is said to be unbiased at level $(1-\alpha)$ if probability of true coverage is _____.

a) at most α b) at least α
c) at most $1-\alpha$ d) at least $1-\alpha$

8) If all frequencies of classes are same, the value of Chi-square is _____.

a) One b) Zero
c) Infinite d) All of the above

B) Fill in the blanks. **04**

1) Let $X_1, X_2, \ldots, X_n$ be iid $N(\mu\sigma^2)$ where σ^2 is known. Then pivotal quantity for confidence interval of μ is, _____.
2) Generalized NP lemma is used to construct _____ tests.
3) The degrees of freedom for a χ^2 statistic in case of contingency table of order 2×2 are _____.
4) The approximate distribution of Kruskal-Walli's test statistic is _____.

Q.2 Answer the following. (Any Six) **12**

a) Define null and alternative hypothesis. Give an example for each.
b) Distinguish between randomized and non-randomized tests.
c) Define $(1-\alpha)$ level confidence set.
d) State monotone likelihood ratio (MLR) property of a family of distributions.
e) Define
i) similar test and
ii) test having Neyman structure
f) Describe likelihood ratio test (LRT).
g) What is goodness of fit test? Give its application.
h) Define confidence set and UMA confidence set of level $(1-\alpha)$.

Q.3 Answer the following. (Any Three) **12**

a) State Neyman Pearson lemma for randomized tests.
b) A sample of size one is taken from Poisson distribution with parameter λ . Let $H_0: \lambda = 1$ and $H_1: \lambda = 2$ Consider the test function

$$\phi(x) = \begin{cases} 1, & x > 3 \\ 0, & \text{otherwise} \end{cases}$$

Find probability of Type-I error and power of the test.
c) Define UMP test. Show that UMP level α test is always an unbiased test.
d) Show that MP test need not be unique using suitable example.

Q.4 Answer the following. (Any Two) **12**

a) State the generalized Neyman-Pearson lemma. Also explain in detail any one of its applications.
b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\theta, 1)$. Obtain UMP level α test for testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$.

c) Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $N(\theta, 1)$distribution. Obtain shortest length confidence interval for σ^2.

Q.5 Answer the following. (Any Two) **12**

a) Explain the concepts of UMPU tests and show that MP and UMP tests of size α are unbiased.

b) Describe Kruskal-Walli's test.

c) Derive LRT for testing $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$ based on a sample of size n from $N(\mu, 1)$ distribution.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - II) (New) (NEP CBCS) Examination: October/November – 2025
Probability Theory (2329207)

Day & Date: Saturday, 01-11-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

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

Q.1 A) Choose the correct alternative. (MCQ) **08**

1) A class $\digamma$ is said to be closed under complementation, if $A \in \mathbb{F}$ implies _____.
   a) $A \cap B \in \mathbb{F}$, for all $B \in \mathbb{F}$ b) $A^c \in \mathbb{F}$
   c) both (a) and (b) d) None of these

2) A mapping X is said to be a random variable if _____.
   a) X is one-to-one b) X is many-to-one
   c) X is linear d) None of these

3) The probability measure is a _____ measure.
   a) Finite b) Infinite
   c) Countably infinite d) Uncountable

4) Which of the following is an elementary random variable?
   a) Binomial b) Poisson
   c) Discrete uniform d) Bernoulli

5) Lebesgue measure of a set $(-1,1)$ is _____.
   a) 0 b) 1
   c) 2 d) -2

6) If $E(X) \geq 0$, then _____.
   a) $X \geq 0$
   b) $X \geq 0$ a.s.
   c) X is not necessarily non-negative
   d) X is simple r.v.

7) If a random variable X is integrable, then _____.
   a) X^+ is integrable b) X^- is integrable
   c) $|X|$ is integrable d) All of these

8) Expectation of a simple non-negative random variable Satisfies _____.
   a) Linearity property
   b) Scale preserving property
   c) Non-negativity property
   d) All of these

B) Fill in the blanks. **04**

1) A _____ function is a countable linear combination of indicators of set.
2) Convergence _____ is the strongest mode of convergence.
3) If $\{A_n\}$ is decreasing sequence of sets, then the sequence $\{A_n^c\}$ is a _____ sequence.
4) The largest field of subsets of Ω is called as _____.

Q.2 Answer the following. (Any Six) **12**

a) Define almost sure event.
b) Define a σ-field.
c) Define a measurable function.
d) Define monotonic sequence of sets.
e) Define null set.
f) Define closure under complementation.
g) Define and illustrate class.
h) Define convergence in distribution.

Q.3 Answer the following. (Any Three) **12**

a) Show that inverse mapping preserves all the set relations.
b) Define simple random variable. Also define its expectation.
c) Discuss inverse mapping in details.
d) Define indicator function and simple function.

Q.4 Answer the following. (Any Two) **12**

a) Show that inverse image of a σ-field is also a σ-field.
b) Define expectation of simple random variable. If X is a non-negative discrete random variable, then prove that-

$$E(X) = \sum_{k=0}^{\infty} P(X > k)$$

c) Prove the linearity property of expectation. Also prove that $E(cX) = cE(X)$

Q.5 Answer the following. (Any Two) **12**

a) If X is continuous non-negative random variable, with cdf $F(.)$, prove that-

$$E(X)\int_0^\infty \bar{F}(x)dx$$

b) Define and explain concept of limit superior and limit inferior of a sequence of random variables.

c) If $I_A(\omega)$ is the indicator function w.r.t. A, then prove the following

i) $I_{A^c}(\omega) = 1 - I_A(\omega)$

ii) If $A \subset B$, then $I_A(\omega) \leq I_B(\omega)$

iii) If $A_1, A_2, \ldots, A_k$ are sets from the same field.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - III) (New) (NEP CBCS)
Examination: October/November - 2025
Multivariate Analysis (2329301)

Day & Date: Wednesday, 29-10-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

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

Q.1 A) Choose the correct alternative. (MCQ) 08

1) With usual notations, Fisher's discriminant function is given by _____.
   a) $\delta'\Sigma^{-1}\underline{X}$ b) $\delta'\Sigma\underline{X}$
   c) $\delta'\Sigma\underline{\mu_1}$ d) $\delta'\Sigma\underline{\mu_2}$

2) Wishart distribution is a multivariate generalization of _____.
   a) Normal b) chi-square
   c) T d) F

3) If X has $N_p(\mu, \Sigma)$ distribution then characteristic function of vector X is _____.
   a) $Exp\left(it'\mu - \frac{1}{2}t'\Sigma t\right)$ b) $Exp\left(it'\mu + \frac{1}{2}t'\Sigma t\right)$
   c) $Exp\left(it'\mu - \frac{1}{2}t'\Sigma^{-1}t\right)$ d) $Exp\left(it'\mu + \frac{1}{2}t'\Sigma^{-1}t\right)$

4) Which of the following statistical techniques identifies homogenous subgroups?
   a) Factor analysis
   b) Multivariate analysis of variance
   c) Cluster analysis
   d) Discriminant analysis

5) Canonical correlation is _____.
   a) Always positive b) Always negative
   c) Lies in between (-1,0) d) None of these

6) If X has $N_P(\mu, \Sigma)$ distribution with $|\Sigma| > 0$ then $(X-\mu)'\Sigma^{-1}(X-\mu)$ is distributed as _____.
   a) χ^2_{p-1} b) χ^2_p
   c) χ^2_{n-p} d) χ^2_{n-p+1}

7) Let A has $W_p(n, \Sigma)$ distribution and B is a $(q \times p)$ matrix then distribution of BAB′ is _____.
   a) $W_p(n, \Sigma)$ b) $W_q(n, \Sigma)$
   c) $W_p(n, B\Sigma B')$ d) $W_q(n, B\Sigma B')$

8) As the distance between two populations increases, misclassification error _____.
   a) decreases b) increases
   c) remains constant d) none of these

B) Fill in the blanks. **04**

1) _____ is a clustering procedure where all objects start out in one giant cluster.
2) A _____ is a graphical device for displaying clustering results.
3) If X has $N_P(\mu, \Sigma)$ distribution then linear combination $Z = a'X$ has _____ distribution.
4) In complete linkage method, we merge those two clusters (or observations) which have _____ distance.

Q.2 Answer the following. (Any Six) **12**

a) State multivariate normal density.
b) State the density of Wishart distribution.
c) Define single linkage.
d) If X has $N_P(\mu, \Sigma)$ distribution with $|\Sigma| > 0$ then what will be the distribution of $\overline{X}$, the sample mean?
e) Define generalized variance.
f) Define sample variance-covariance matrix.
g) State additive property of normal distribution.
h) State additive property of Wishart distribution.

Q.3 Answer the following. (Any Three) **12**

a) Discuss singular and non-singular normal distribution.
b) Derive the MGF of multivariate normal distribution.
c) Show that two p-variate normal vectors $\underline{X}_1$ and $\underline{X}_2$, are independent if and only if $cov(\underline{X}_1, \underline{X}_2) = 0$.
d) Define variance-covariance matrix. State its properties.

Q.4 Answer the following. (Any Two) **12**

a) Define multivariate normal distribution. State its real-life applications. Also find out MLE of population mean vector based on a sample.
b) What do you mean by hierarchical clustering? Also explain agglomerative clustering.
c) Explain the idea of discriminant analysis. What are the potential errors involved in it? Obtain the classification rule for the case of two populations with densities $f_1(\underline{x})$ and $f_2(\underline{x})$.

Q.5 Answer the following. (Any Two) **12**

a) Obtain the value of constant k in the expression of multivariate normal density.

b) Obtain the distribution of several linear combinations of components of a multivariate normal vector.

c) Explain and illustrate clustering.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - III) (New) (NEP CBCS) Examination: October/November - 2025
Regression Analysis (2329302)

Day & Date: Friday, 31-10-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

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

Q.1 A) Choose correct alternative. (MCQ) **08**

1) Autocorrelation is concerned with _____.
   a) correlation among the predictors
   b) correlation among the error terms
   c) correlation among response and predictors
   d) None of these

2) The coefficient of determination (R^2) is the square of correlation coefficient between (where Y is response) _____.
   a) Y and its predicted value
   b) Y and hat matrix
   c) regressors
   d) None of these

3) Gauss-Newton iteration method of parameter estimation can be used in _____.
   a) Linear regression model
   b) Nonparametric regression model
   c) Polynomial regression model
   d) Nonlinear regression model

4) Forward selection procedure starts with the _____ predictor variables in the model.
   a) all b) some
   c) without d) None of these

5) The multicollinearity problem in regression concerns the _____.
   a) regressors
   b) error terms
   c) response variable values
   d) regression coefficients

6) The largest condition index of $(X'X)$ is defined as _____.
   a) $\lambda_{max} - \lambda_j$ b) $\lambda_{max} - \lambda_{min}$
   c) $\frac{\lambda_{max}}{\lambda_{min}}$ d) $\frac{\lambda_{min}}{\lambda_{max}}$

7) Logistic regression model is an appropriate model when response variable is distributed as _____.
   a) Poisson b) Binomial
   c) Normal d) Gamma

8) The regression model $Y = \beta_0 + \beta_1 X + \beta_2 X^2$ is called _____.
   a) Linear b) Non-linear
   c) Polynomial d) None of these

B) Fill in the blanks. **04**

1) The sum of residuals weighted by corresponding fitted values is always _____.
2) The variance stabilizing transformation $\sqrt{Y}$ is used when distribution of Y is _____.
3) The joint points of pieces in polynomial fitting are usually Called _____.
4) Any model that is not in linear in the unknown parameters is a _____ regression model.

Q.2 Answer the following. (Any Six) **12**

a) Prove that hat matrix H and I-H are idempotent.
b) Define condition number and condition indices of $(X'X)$ matrix.
c) Define standardized residuals, and explain its role in regression analysis.
d) Derive the relationship between coefficient of determination R^2 and adj. R^2.
e) With usual notations, show that $Var(\hat{Y}) = H\sigma^2$.
f) Discuss Box-Cox power transformation.
g) Explain the problem of autocorrelation.
h) Discuss the logit transformation in the context of logistic regression model.

Q.3 Answer the following. (Any Three) **12**

a) Describe the ANOVA approach to test the significance of regression in a simple linear regression model.
b) Discuss the variance inflation factor (VIF) method for detection of multicollinearity.
c) In multiple linear regression model with usual notations, show that an unbiased estimator of σ^2 is MS_{Res}.
d) Discuss variance stabilizing transformation and its use.

Q.4 Answer the following. (Any Two) **12**

a) Define multiple linear regression model. Find unbiased estimators of the parameters of this model.
b) Derive Mallows' C_p statistics. Explain the procedure of variable selection in linear regression using C_p.
c) Explain the problem of autocorrelation. Discuss Cochrane-Orcutt method of parameter estimation.

Q.5 Answer the following. (Any Two) **12**

a) Define k^{th} order polynomial regression model in one variable. Describe orthogonal polynomial to fit the polynomial model in one variable.

b) Obtain the weighted least squares estimator of the parameters involved in generalized linear model.

c) Explain the residual plots. Outline the procedure of construction of normal probability plot and procedure for checking normality assumption.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - III) (New) (NEP CBCS) Examination: October/November – 2025
Design and Analysis of Experiments (2329306)

Day & Date: Monday, 03-11-2025 Max. Marks: 60
Time: 11:00 AM To 01:30 PM

Instructions: 1) All questions are compulsory.
2) Figures to right indicate full marks.

Q.1 A) Fill in the blanks by choosing correct alternatives given below. 08

1) The main assumption of ANOVA is that the groups being compared have _____.
   a) Equal variances b) Unequal variances
   c) Unequal sample sizes d) dependent observations

2) In a 2^3 factorial experiment with, the contrast due to main effect A is _____.
   a) $[(a)+(ab)+(ac)+(abc)-(1)-(b)-(c)-(bc)]$
   b) $[(abc)-(ac)+(bc)-(c)+(ab)-(a)+(b)-(1)]$
   c) $[(abc)+(ac)-(bc)-(c)+(ab)-(a)-(b)+(1)]$
   d) $[(bc)+(ab)-(abc)-(c)+(ac)-(a)-(b)+(1)]$

3) The error degrees of freedom in two-way classification model $Y_{ij}=\mu+\alpha_i+\beta_j++\epsilon_{ij}; 1,2,\dots,I, j=1,2,\dots J$ and assumptions on errors are followed; are _____?
   a) $(J-1)$ b) $(I-1)$
   c) $(I-1)(J-1)$ d) $(IJ-1)$

4) Covariance between vector of adjusted treatment totals and vector of block total $\big(Cov(Q,B)\big)$ is _____.
   a) 0 b) 1
   c) 3 d) 2

5) In a two-way ANOVA with no interaction, which of the following is true?
   a) Only the main effects of factors A and B are considered
   b) The interaction between factors A and B is significant
   c) Both main effects and interaction are tested
   d) There are no main effects

6) For Balanced incomplete block design (b,k,v,r,λ)
   a) $bv=kr$ b) $bk=vr$
   c) $r(v-1)=\lambda(k-1)$ d) $v(r-1)=k(\lambda-1)$

7) Which of the following is one-way model?
   a) $Y_i = \mu + \alpha_i + \epsilon_i$; for all $i = 1,2, \dots p, j = 1,2 \dots n_i$
   b) $Y_{ijk} = \mu + \alpha_i + \beta(\bar{X}_{ijk} - \bar{X}..) + \epsilon_{ijk}$; for all $i = 1,2, \dots p, j = 1,2 \dots n_i$
   c) $Y_i = \mu - \alpha_i + \epsilon_i$; for all $i = 1,2, \dots p, j = 1,2 \dots n_i$
   d) $Y_i = \mu + \alpha_i - \beta_j + \epsilon_i$; for all $i = 1,2, \dots p, j = 1,2 \dots n_i$

8) Smaller the experimental error _____ efficient the design.
   a) Less b) More
   c) Equally d) None of these

B) Fill in the blanks. **04**

1) The degrees of freedom for the error term in a one-way ANOVA are calculated as _____ (where N is the total number of observations and k is the number of groups or treatments).
2) In a one-way ANOVA, the null hypothesis states that the means of all the groups are ____.
3) BLUE of estimable treatments contrast can be expressed in _____ function of Q.
4) In a 2^6 factorial experiments, there are _____ three factor effects.

Q.2 Answer the following. (Any Six) **12**

a) Explain in brief analysis of co-variance.
b) What is complete block design?
c) What is intrablock analysis?
d) In a 2^3 factorial experiment, the eight treatment combinations are grouped into two blocks of size 4 as follows:

Block 1	(c), (b),(a),(abc)
Block 2	(1), (bc),(ac),(ab)

Determine the confounded factorial effect.
e) Explain in brief balanced design.
f) Check whether the given design is connected or not.

Block	Treatments
I	1,2,3
II	5,6
III	7,8

g) Explain in brief analysis of variance.
h) Define resolution IV design

Q.3 Answer the following. (Any Three) **12**

a) Define two-way ANOVA model with interaction. Obtain the least square estimates of parameters of the same model.
b) Define BIBD. Show that in a BIBD (b, k, v, r, λ) $\lambda(v-1) = r(k-1)$
c) Explain partial confounding with example.
d) Show that $c = R^{\delta} - Nk^{-\delta}N'$

Q.4 Answer the following. (Any Two) **12**

a) Derive the test for testing treatment in one-way ANCOVA model.

b) Describe the analysis of 2^2 factorial experiments.

c) What is confounding? Explain in detail types of confounding.

Q.5 Answer the following. (Any Two) **12**

a) Derive the test for testing hypothesis of equality of all treatment effects in two-way classification model without interaction.

b) Describe the analysis of 3^2 factorial experiments.

c) Explain the structure of suit fraction in fraction factorial experiment.

Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - IV) (New) (NEP CBCS) Examination: October/November - 2025
Reliability and Survival Analysis (2329402)

Day & Date: Tuesday, 28-10-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose correct alternative. (MCQ) **08**

1) Which of the following is not true?
a) K-M estimator is parametric in nature.
b) K-M estimator is generalized maximum likelihood estimator.
c) K-M estimator is consistent.
d) K-M estimator is also known as product limit estimator.

2) Actuarial method of estimation of survival function is used when data consists of _____.
a) only censored observations
b) only uncensored observations
c) complete data
d) All the above

3) In type I censoring, _____.
a) the number of failures is fixed
b) duration of an experiment is fixed
c) both time and number of failures is fixed
d) None of these

4) Kaplan-Meier estimator is also called as _____.
a) estimator of product
b) product limit estimator
c) maximum likelihood estimator
d) moment estimator

5) In right censoring, _____.
a) the number of complete observations is random variable
b) the time of an experiment is random variable.
c) both a and b
d) the time of an experiment is fixed.

6) In series system of five components, the entire system will fail if _____.
a) any two components fail
b) any three components fail
c) any one of the components fail
d) any four components fail

7) A series system is a special case of $k - out - of - n$ system when _____.
a) $k = 1$ b) $k = n - 1$
c) $k = 2$ d) $k = n$

8) Which of the following distribution has no ageing property?
a) Lognormal b) Gamma
c) Exponential d) None of these

B) Fill in the blanks. **04**

a) MTTF stands for _____.
b) Parallel system of n components has _____ minimal path sets.
c) The number of failures is fixed in _____ censoring.
d) For a series system of two independent components each having reliability 0.7 then the reliability of system is _____.

Q.2 Answer the following. (Any Six) **12**

a) Define series system.
b) Obtain structure function for series system.
c) Define minimal path vector.
d) Define DFR distribution.
e) Define NWU class.
f) Define k-out-of-n system of components.
g) Define NBUE class.
h) What is meant by path vector?

Q.3 Answer the following. (Any Three) **12**

a) Define reliability of a system. Obtain the reliability of parallel system of n independent components.
b) Show that empirical distribution function is unbiased estimate of population distribution function.
c) Discuss censoring.
d) Define hazard rate of a distribution.

Q.4 Answer the following. (Any Two) **12**

a) Define IFR and IFRA classes of distributions. Prove that $IFR \subset IFRA$.
b) Describe type II censoring. Obtain the likelihood under this censoring.
c) Discuss maximum likelihood estimation of parameters of a Weibull distribution based on complete data.

Q.5 Answer the following. (Any Two) **12**

a) Describe Type-I censoring. Obtain MLE of mean of exponential distribution under Type I censoring.

b) Describe actuarial method of estimation of survival function, with suitable illustration.

c) Obtain Greenwood's formula for the variance of the Actuarial estimator.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - IV) (New) (NEP CBCS)
Examination: October/November - 2025
Industrial Statistics (2329401)

Day & Date: Thursday, 30-10-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose correct alternative. (MCQ) **08**

1) In the development of $\bar{X}$ and S charts, the distribution of quality characteristic X is assumed to be _____.
a) Normal b) Bernoulli
c) Geometric d) Poisson

2) Which of the following is useful in data collection activity?
a) Control chart b) Check sheet
c) Histogram d) Pareto chart

3) Which of the following is on-line monitoring tool?
a) Histogram b) Pareto chart
c) Control chart d) Scatter diagram

4) The capacity index C_{pk} involves _____ parameter(s) to be estimated.
a) both μ and σ b) only μ
c) only σ d) None of the above

5) For a variable sampling plan, the distribution of quality characteristic is assumed to be _____.
a) exponential b) geometric
c) normal d) lognormal

6) In demerit system, the unit will not fail in service but has minor defects in finish or appearance is classified as _____ defects.
a) class A b) class B
c) class C d) class D

7) Type I error occurs when _____.
a) a good lot is rejected
b) a bad lot is accepted
c) the number of defectives is very large
d) the population is worse than the AQL

8) The p chart is used to control _____.
a) nonconformities
b) nonconformities per unit
c) fraction nonconforming
d) variability

B) Fill in the blanks. **04**

1) In demerit system, the occurrence of defects in each class is modelled by _____ distribution.
2) To determine location of a defect, _____ SPC tool is used.
3) The statistical process _____ chart used to control number of nonconformities of output.
4) An index C_p was introduced by _____.

Q.2 Answer the following. (Any Six) **12**

a) Define ARL of control chart. State the distribution of ARL.
b) Distinguish between process control and product control.
c) Describe a single sampling plan for attributes.
d) Define type I and type II errors relative to control charts.
e) What is principle behind Pareto Analysis?
f) Define Index C_P stating the assumptions clearly.
g) What is the difference between p and np charts?
h) Define Hotelling T^2 statistic.

Q.3 Answer the following. (Any Three) **12**

a) Describe double sampling plan for attributes. Obtain AOQ for the same.
b) Discuss relationship between a control chart and statistical hypothesis testing.
c) Describe DMAIC with reference to six-sigma.
d) What is moving average chart? Explain its construction and operation.

Q.4 Answer the following. (Any Two) **12**

a) Describe the development and operation of Hotelling's T^2 chart to monitor process mean vector.
b) Explain the process capability index C_{PK}. With usual notations, prove that $\Phi(-3C_{PK}) \leq p \leq 2\Phi(-3C_{PK})$
c) List seven quality control tools and explain in detail any two of them.

Q.5 Answer the following. (Any Two) **12**

a) Define control chart. Discuss $\bar{X}$ and R charts for controlling the quality of product.
b) Explain variable sampling plan when lower specification is given and σ is known.
c) What is CUSUM chart? Explain its construction and operation.

Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - IV) (New) (NEP CBCS) Examination: October/November – 2025
Data Mining (2329405)

Day & Date: Saturday, 01-11-2025 Max. Marks: 60
Time: 03:00 PM To 05:30 PM

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

Q.1 A) Choose correct alternative. (MCQ) **08**

1) Which one is example of case based learning?
   a) Decision Tree b) K-Nearest neighbor
   c) Genetic algorithm d) Neural networks

2) A signal is received at _____ of a neuron.
   a) Axon b) Synapse
   c) Nucleus d) Dendrites

3) Support vector machine is _____.
   a) unsupervised learning
   b) supervised learning
   c) reinforcement learning
   d) genetic algorithm

4) Which of the following refers to the problem of finding abstracted patterns (or structures) in the unlabeled data?
   a) supervised learning b) unsupervised learning
   c) mixed learning d) all of these

5) _____ maps data into predefined groups.
   a) Regression b) Time series analysis
   c) Prediction d) Classification

6) An agglomerative hierarchical clustering method uses a _____ strategy.
   a) Top-down b) Bottom-up
   c) Random d) None of these

7) Which one is non-hierarchical clustering algorithm?
   a) Agglomerative clustering
   b) Divisive clustering
   c) k-means clustering
   d) All of these

8) k-nearest neighbor method can be used _____.
   a) only when class labels are qualitative
   b) only when class labels are quatitative
   c) Both (a) and (b)
   d) None of these

B) Fill in the blanks. **04**

1) In kNN classifier, k stands for _____.
2) Looking for combinations of items purchased together is Called _____.
3) In data mining, ANN stands for _____.
4) The _____ algorithm of supervised learning is known as 'Lazy learning algorithms'.

Q.2 Answer the following. (Any Six) **12**

a) Define supervised learning.
b) Define recall of a classifier.
c) Define confusion matrix.
d) Explain the need of data cleaning.
e) What is meant by True Negatives?
f) List any four classifiers based on supervised learning.
g) What is meant by False Negatives?
h) Why Bayes' Classifier is called as Naive classifier?

Q.3 Answer the following. (Any Three) **12**

a) Discuss kNN classifier.
b) Discuss the problem of imbalanced data.
c) What are the advantages of unsupervised learning.
d) Discuss accuracy and precision of a classifier.

Q.4 Answer the following. (Any Two) **12**

a) Discuss:
   i) Support Vectors.
   ii) Hyperplane.
b) Describe decision tree classifier in detail.
c) Describe-
   i) Sensitivity of a model.
   ii) Specificity of a model.
   Illustrate with the help of example.

Q.5 Answer the following. (Any Two) **12**

a) Explain in detail, market basket analysis.
b) Explain artificial neural network in detail.
c) Explain logistic regression as a classifier.

Seat No.	

Set **P**

M.Sc. (Statistics) (Semester - IV) (New/Old) (CBCS) Examination: October/November - 2025
Data Mining (MSC16401)

Day & Date: Tuesday, 28-10-2025 Max. Marks: 80
Time: 03:00 PM To 06:00 PM

Instructions: 1) Questions no. 1 & 2 are compulsory.
2) Attempt any three questions from Q No.3 to Q No.7
3) Figures to the right indicate full marks.

Q.1 A) Choose correct alternative. **10**

1) Looking for combinations of items purchased together is called _____.
a) market data analysis
b) market basket analysis
c) marketing data analysis
d) Combo analysis

2) Market-basket problem was formulated by _____.
a) Agrawal et al. b) Toda et al.
c) Steve et al. d) Simon et. Al

3) Support vector machine is _____.
a) unsupervised learning
b) supervised learning
c) reinforcement learning
d) genetic algorithm

4) Which one is non-hierarchical clustering algorithm?
a) Agglomerative clustering
b) Divisive clustering
c) k-means clustering
d) All of these

5) Naive Bayesian classifier uses _____ tool.
a) Information gain b) Probability
c) Both (a) and (b) d) None of these

6) Task of inferring a model from unlabeled training data is called _____.
a) supervised learning
b) unsupervised learning
c) Reinforcement learning
d) None of these

7) The final output of data mining is _____.
a) Data b) Clean data
c) Information d) All of these

8) Each neuron is made up of a number of nerve fibres called _____.
a) Molecules b) Dendrites
c) Atoms d) Sigmoid

9) Which of the following is the not a type of clustering?
a) k-means b) Hierarchical
c) Non-hierarchical d) Splitting

10) _____ data are noisy and have many missing attribute values.
a) Discretized b) Real-world
c) Cleaned d) Transformed

B) Fill in the blanks. **06**

1) KNN is an example of _____ Learning method.
2) In _____ learning, class labels are provided.
3) The part of the entire data, which is used for building the model is called as _____.
4) In data mining, ANN stands for _____.
5) The _____ algorithm of supervised learning is known as 'Lazy learning algorithms'.
6) _____ is the type of machine learning in which machines are trained using well "labelled" training data.

Q.2 Answer the following. **16**

a) Differentiate between training data and testing data.
b) What is meant by imbalanced data?
c) What are the advantages of unsupervised learning?
d) Why Bayes' classifier is called Naive classifier?

Q.3 Answer the following.

a) Discuss the different metrics for Evaluating Classifier Performance. **08**
b) Describe decision tree classifier in detail. **08**

Q.4 Answer the following.

a) Discuss logistic regression classifier in detail. **08**
b) Discuss in detail about how the order of features is considered in decision tree with respect to information gain. **08**

Q.5 Answer the following.

a) Describe unsupervised learning. **08**
b) Discuss the Bayesian classifier. **08**

Q.6 Answer the following.

a) Discuss density based methods for unsupervised learning. **08**

b) Explain in detail, market basket analysis. **08**

Q.7 Answer the following.

a) Discuss k-nearest neighbor classifier in detail. **08**

b) Describe supervised learning method. Also explain SVM classifier. **08**

Seat No.	

Set P

M.Sc. (Statistics) (Semester - IV) (New/Old) (CBCS) Examination: October/November - 2025
Industrial Statistics (MSC16402)

Day & Date: Thursday, 30-10-2025 Max. Marks: 80
Time: 03:00 PM To 06:00 PM

Instructions: 1) Questions no. 1 & 2 are compulsory.
2) Attempt any Three Question from Q No.3 to Q No.7
3) Figures to the right indicate full marks.

Q.1 A) Choose correct alternative. **10**

1) The performance measure of c and u charts is based on the assumption that the occurrence of nonconformities follows _____ distribution.
a) Geometric b) Binomial
c) Poisson d) Normal

2) Which of the following is useful in data collection activity?
a) Check sheet b) Control chart
c) Histogram d) Pareto chart

3) Normality assumption of population data values is made for _____ index.
a) C_p b) C_{pk}
c) C_{pm} d) All the above

4) The curve showing the probability of acceptance of a lot of quality p is known as _____.
a) AOQ curve b) ASN curve
c) OC curve d) ARL curve

5) Producer's risk is the probability of _____.
a) accepting a good lot b) rejecting a good lot
c) rejecting a bad lot d) accepting a bad lot

6) An index C_p was introduced by _____.
a) Kane b) Juron
c) Chan d) Pearn

7) When acceptance quality levels are very small, the sample size required by attribute sampling plan is _____.
a) very small b) very large
c) moderate to large d) None of the above

8) When the upper and lower natural tolerance limits are equal to the upper and lower specification limits, the process capability index, C_p is _____.

a) Greater than 1 b) 0
c) Less than 1 d) Equal to 1

9) The capacity index C_{pk} involves _____ parameter(s) to be estimated.

a) Only μ b) Only σ
c) Both μ and σ d) None of the above

10) The ASN of a double sampling plan reduces to that of a single sampling plan if probability of making a decision on the basis of first sample is _____.

a) 0 b) 0.5
c) 0.75 d) 1

B) Fill in the blanks. **06**

1) Quality is inversely proportional to _____.
2) In Shewhart type control chart, an appropriate distribution of run length is _____.
3) In demerit system, the occurrence of defects in each class is modelled by _____ distribution.
4) For a variable sampling plan, the distribution of quality characteristic is assumed to be _____.
5) If the LSL of the quality characteristic increases, the process capability index C_p will _____.
6) In 'DMAIC', M stands for _____.

Q.2 Answer the following. **16**

a) Distinguish between process control and product control. What are the statistical techniques to achieve these?

b) Define
   i) Consumer's risk and
   ii) Producer's risk.

c) Explain the procedure of estimating the process capability index C_p.

d) Define type I and type II errors relative to control charts.

Q.3 Answer the following.

a) Define OC function and ARL of a control chart. Obtain the same for $\bar{X}$ chart assuming normality of process with known standards. **08**

b) Explain in detail the development and implementation of Hotelling's T^2 chart. **08**

Q.4 Answer the following.

a) What is CUSUM chart? Explain tabular CUSUM procedure for monitoring process mean. **08**

b) Discuss various definitions of 'Quality' and various dimensions of quality. **08**

Q.5 Answer the following.

a) Define process capability index C_p. Stating the underlying assumption clearly, establish relationship between C_p and probability of nonconforming item. **08**

b) Discuss various steps involved in the construction of p chart with fixed sample size and variable sample size. **08**

Q.6 Answer the following.

a) Describe the single sampling plan and obtain its OC function. **08**

b) Distinguish between defect and defective. Give some examples of defects for which c chart is applicable. How do you calculate control limits for c chart? Stare assumptions made. **08**

Q.7 Answer the following.

a) Explain the basic concepts of six-sigma methodology. Also explain the benefits of implementing the same. **08**

b) Explain the variable sampling plan when lower specification is given and standard deviation is known. **08**

Seat No.		Set	P

M.Sc. (Statistics) (Semester - IV) (New/Old) (CBCS)
Examination: October/November – 2025
Reliability and Survival Analysis (MSC16403)

Day & Date: Saturday, 01-11-2025 Max. Marks: 80
Time: 03:00 PM To 06:00 PM

Instructions: 1) Questions no. 1 & 2 are compulsory.
2) Attempt any Three Question from Q No.3 to Q No.7
3) Figures to the right indicate full marks.

Q.1 A) Choose the correct alternative. **10**

1) For Weibull distribution _____ parameter decides whether distribution belongs to IFR or DFR class.
   a) Location b) Shape
   c) Scale d) all the above

2) Which of the following rate function corresponds to IFR distribution?
   a) $h(t) = t$ b) $h(t) = e^t$
   c) $h(t) = t\, e^t$ d) All the above

3) A vector $\underline{X}$ is called path vector if _____.
   a) $0 \leq \phi(\underline{X}) \leq 1$ b) $\phi(\underline{X}) = 1$
   c) $\phi(\underline{X}) = 0$ d) $\phi(\underline{X}) = 0.5$

4) Let p_i is the reliability of i^{th} component then reliability of parallel system of n independent components is _____.
   a) $1 - \coprod_{i=1}^{n} p_i$ b) $1 - \prod_{i=1}^{n} p_i$
   c) $1 - \prod_{i=1}^{n} (1 - p_i)$ d) $1 - \sum_{i=1}^{n} (1 - p_i)$

5) A life time distribution F having finite mean is said to be NBUE for $t \geq 0$, if _____.
   a) $\mu_t \leq \mu_0$ b) $\mu_t \geq \mu_0$
   c) $\mu_t = \mu_0$ d) None of the above

6) IFRA property is preserved under _____.
   a) Mixture b) Convolution
   c) Coherent d) all the above

7) In survival analysis, the outcome variable is _____.
   a) Continuous b) Discrete
   c) Dichotomous d) none of the above

8) Study period is fixed in _____ censoring.
   a) random censoring b) Type I
   c) Type II d) all the above

9) The cumulative distribution function of survival time random variable T is defined as _____.
   a) $P(T = t), \forall t \geq 0$ b) $P(T \geq t), \forall t \geq 0$
   c) $P(T > t), \forall t \geq 0$ d) $P(T \leq t), \forall t \geq 0$

10) Let X follows Weibull distribution with parameters λ and α. Then $Y = X^{\alpha}$has _____ distribution.
   a) Exponential with parameter α
   b) Exponential with parameter λ
   c) Standard exponential
   d) $U(0,1)$

B) Fill in the blanks. **06**

1) As the number of components n increases, the reliability of series system _____.
2) Let p_i is the reliability of i^{th} component then reliability of series system of n independent components is _____.
3) Parallel system of n components has _____ minimal path sets.
4) In type I censoring, the number of uncensored observations has _____ distribution.
5) Product limit estimator of survival function is developed by _____ and _____.
6) The hazard function ranges between _____ and _____.

Q.2 Answer the following. **16**

a) Define minimal path sets and minimal cut sets. Illustrate the same by example.

b) Define mean residual life function and obtain the same for exponential distribution.

c) Define Poly function of order 2(PF_2) Prove that if $f \in PF_2$ then $F \in IFR$.

d) Describe Type-I and Type-II censoring with suitable examples.

Q.3 Answer the following.

a) Define k out of n system. Obtain the reliability function of this system. **08**

b) If $X_1, X_2, \ldots, X_n$ are associated state variables of coherent system then prove that **08**

$$\coprod_{i=1}^{n} P(X_i = 1) \leq P(\phi(X) = 1) \leq \coprod_{i=1}^{n} P(X_i = 1).$$

Q.4 Answer the following.

a) Define IFR and IFRA class of distributions. If $F \in IFR$ then show that $F \in IFRA$. **08**

b) If failure time of an item has the distribution **08**
$f(t) = \frac{\lambda^{\alpha}}{\Gamma\alpha} t^{\alpha-1} e^{-\lambda t}, t > 0, \lambda, \alpha > 0.$
Examine whether it belongs to IFR or DFR.

Q.5 Answer the following.

a) Define star shaped function. Prove that $F \in IFRA$ if and only $if - logR(t)$ is star shaped. **08**

b) Discuss maximum likelihood estimation of parameters of a Weibull distribution based on complete data. **08**

Q.6 Answer the following.

a) Describe Kaplan-Meier estimator and derive an expression for the same. **08**

b) Obtain maximum likelihood estimate of mean of the exponential distribution under, type II, censoring. **08**

Q.7 Answer the following.

a) Obtain the actuarial estimator of the survival function. Clearly state the assumption that you need to make. State Greenwood's formula for the variance of the estimator. **08**

b) Describe Gehan's test for two sample testing problem in presence of censoring. **08**

Set No.		Set	P

M.Sc. (Statistics) (Semester - IV) (New/Old) (CBCS) Examination: October/November - 2025
Optimization Techniques (MSC16404)

Day & Date: Tuesday, 04-11-2025 Max. Marks: 80
Time: 03:00 PM To 06:00 PM

Instructions: 1) Questions no. 1 & 2 are compulsory.
2) Attempt any Three Questions from Q No.3 to Q No.7
3) Figures to the right indicate full marks.

Q.1 A) Choose correct alternative. (MCQ) **10**

1) Which of the following is not associated with an LPP?
a) Additivity b) Uncertainty
c) Proportionality d) Divisibility

2) What is the optimum BFS of following LPP?
Max $Z = x_1 + X_2$,
subject to, $x_1 + 2x_2 \leq 4, 3x_1 + 2x_2 \leq 10, x_1 \geq 0, x_2 \geq 0$

a) $x_1 = 0, x_2 = 2$ b) $x_1 = 2, x_2 = 1$
c) $x_1 = \frac{10}{3}, x_2 = 2$ d) $x_1 = 2, x_2 = 2$

3) A two-person zero sum game is said to be zero sum game, if _____.
a) Gain of one player is exactly matched by a loss to the other so that their sum is equal to zero
b) Both players must have exact number of strategies
c) Diagonal entries of pay- off matrix are zero
d) Gain of one player does not match to the loss of other player

4) Given a system of m simultaneous linear equations with n unknowns $(m < n)$. The number of basic variables will be _____**.**
a) n b) m
c) $n - m$ d) none of the above

5) If $X'QX$ is positive semi-definite then it is _____.
a) Strictly convex b) Strictly concave
c) Convex d) Concave

6) Post optimal analysis is technique to _____.
a) Determine how optimum solution to an LPP changes in response to problem inputs
b) Allocate resources optimally
c) Minimize cost operations
d) Spell out the relation between dual and its primal

7) A two person zero sum game is said to be fair if _____.
   a) The upper value and lower value of the game are not equal
   b) The upper value is more than lower value of the game
   c) The upper value and lower value of the game are same and equal to zero
   d) None of the above

8) A minimax and maximin values of the game are same, then _____.
   a) There is saddle point b) Solution does not exist
   c) Strategies are mixed d) None of the above

9) Dynamic programming deals with the _____.
   a) Multistage decision-making problems
   b) Single stage decision making problems
   c) Time dependent decision-making problems
   d) Problem which fix the levels of different so as to maximize profit or minimize cost.

10) Basic feasible solution of an LPP corresponds to _____.
   a) Extreme points of feasible region
   b) Any point in the feasible region
   c) Any point on the edge of feasible region
   d) Optimum feasible solution

B) Fill in the blanks. **06**

1) Dual variables corresponding to standard primal problem are to be _____ sign.
2) A linear programming problem in which all or some of the decision variable are restricted to be an integer are called _____.
3) The general form for Quadratic programming problem is _____.
4) The dual simplex method starts with _____ solution.
5) Beale's method is used to solve _____ programming problem.
6) Pure strategy with respect to game theory means _____.

Q.2 Answer the following. **16**

a) Describe in detail assumptions in LPP.
b) Discuss graphical procedure to solve LPP.
c) Explain the Maxmin and minimax principal in game theory.
d) Write a note on characteristics of dynamic programing.

Q.3 Answer the following. **16**

a) Define the following terms with respect to LPP
   i) Feasible solution
   ii) Basic solution
   iii) Basic feasible solution
   iv) Optimal basic feasible solution
   v) Unbounded solution
   vi) Infeasible solution

b) Use simplex method to solve following LPP
Maximize $Z = 12x_1 + 20x_2$,
subject to:

$$6x_1 + 8x_2 \leq 100,$$
$$7x_1 + 12x_2 \leq 120,$$
$$X_1, X_2 \geq 0$$

Q.4 Answer the following. **16**

a) Solve the following IPP using cutting plane algorithm.
$MaxZ = 2X_1 + 20X_2 - 10X_3$
Subject to:

$$2X_1 + 20X_2 + 4X_3 \leq 15,$$
$$6X_1 + 20X_2 + 4X_3 = 20,$$
$X_1, X_2, X_3 \geq 0$ and are integers.

b) Define quadratic programming problem and write down Wolfe's algorithm to solve quadratic programming problem.

Q.5 Answer the following. **16**

a) State and prove basic duality theorem.

b) Obtain the range of change in C_j values to maintain feasibility of the optimal solution.

Q.6 Answer the following. **16**

a) Consider the payoff matrix of player A use graphical procedure and obtain optimum strategies of both players. Also obtain value of the game.

Player $A \begin{bmatrix} 2 & 7 \\ 3 & 5 \\ 11 & 2 \end{bmatrix}$

b) Describe recursive approach to solve dynamic programming problem.

Q.7 Answer the following. **16**

a) Discuss the importance of artificial variable while solving LPP using simplex method. Also write an algorithm of Big-M method to solve LPP.

b) Describe the Gomory's fractional cut method for solving all integer programming problem.

Seat No.	

Set P

M.Sc. (Statistics) (Semester - IV) (New/Old) (CBCS) Examination: October/November - 2025
Time Series Analysis (MSC16407)

Day & Date: Friday, 07-11-2025 Max. Marks: 80
Time: 03:00 PM To 06:00 PM

Instructions: 1) Questions no. 1 & 2 are compulsory.
2) Attempt any Three Question from Q No.3 to Q No.7
3) Figures to the right indicate full marks.

Q.1 A) Choose the correct alternative. (MCQ) **10**

1) Random walk is not a weakly stationary series, because _____.
   a) Its mean function depends on t
   b) Its covariance function depends on t
   c) Both mean and covariance function depends on t
   d) None of these

2) Moving average method is used for measurement of trend when:
   a) Trend is linear b) Trend is curvilinear
   c) Trend is non linear d) None of them

3) A sequence of independent random variables with same mean and variance is called as _____.
   a) Random Walk b) MA (1)
   c) IID noise d) White noise

4) The term SARIMA with respect to time series analysis stands for _____.
   a) Seasonal and recyclable invertible moving averages
   b) Seasonal and recyclable invertible median averages
   c) Seasonal autoregressive integrated moving averages
   d) Seasonal autoregressive invertible median averages

5) In time series analysis, the term ARCH means _____.
   a) Autoregressive conditionally hetero-scedastic
   b) Autoregressive conditionally hypergeometric
   c) Autoregressive clearly hetero-scedaustic
   d) None of these

6) A causal process is _____.
   a) Always invertible
   b) Can't be invertible
   c) May or may not be invertible
   d) None of these

7) The purpose of smoothing a time series is _____.
   a) To estimate noise components
   b) To estimate and remove noise components
   c) To magnify noise components
   d) None of these

8) For large n, the sample autocorrelations of an iid sequence $Y_1 \ldots, Y_n$ with finite variance are _____.
   a) approximately iid with $N(0,1/n)$ distribution
   b) approximately iid with chi-square distribution
   c) Not identical
   d) None of these

9) An iid sequence is _____.
   a) 0-dependent b) 1-dependent
   c) 2-dependent d) 3-dependent

10) A collection of time dependent variables is called as _____.
   a) A time data b) A time series
   c) Variable-time d) None of these

B) Fill in the blanks. **06**

1) In the classical decomposition model, the slowly changing function of t is called as _____ component.
2) The time series data obtained after removing seasonal component is called as _____ data.
3) If mean and covariance functions are both independent of time t, then the process is called as _____ stationary.
4) MA process stands for _____.
5) The IID noise process is _____ stationary process.
6) There are _____ components in the time series model.

Q.2 Answer the following. **16**

a) Describe $AR(1)$ model.
b) Discuss $MA(1)$ model.
c) Define an invertible process. What is necessary and sufficient condition for invertibility?
d) Define autocovariance function (ACVF). State the characterizing properties of ACVF.

Q.3 Answer the following.

a) Discuss ARMA(p,q) model in details. **08**
b) Explain exponential smoothing of a time series. **08**

Q.4 Answer the following.

a) Explain moving average smoothing in the absence of seasonality. **08**
b) What are the ARCH and GARCH models? Explain in detail. **08**

Q.5 Answer the following.

a) Define: **08**

i) Weakly stationary time series
ii) Strictly stationary time series
iii) IID Noise
iv) White Noise

b) Discuss ACF and PACF in details. **08**

Q.6 Answer the following.

a) Write the procedure of obtaining Ψj weights and hence to obtain the autocovariance function. **08**

b) Describe Yule-Walker method of estimating the parameters of an $AR(p)$ process. Obtain the same for $AR(2)$ process. **08**

Q.7 Answer the following.

a) Explain any two tests for testing the independence in estimated noise sequence. **08**

b) What are the different methods of diagnostic checking in time series? Explain the role of residual analysis in model checking. **08**