C 22726 (Pages : 2) Name.......................................

Reg. No..................................

FOURTH SEMESTER M.Sc. DEGREE (SUPPLEMENTARY) EXAMINATION APRIL 2022

(CUCSS)

Statistics

ST 4E 06—TIME SERIES ANALYSIS

(2013 to 2018 Admissions)

Time : Three Hours Maximum : 36 Weightage

Part A

Answer **all** *questions.*
Weightage 1 *for each question.*

1. What is meant by stationary time series ?
2. Define auto-covariance function.
3. What is trend in time series analysis ?
4. Define autoregressive process.
5. Which operator is used to convert the nonstationarity data into stationarity data ?
6. What is the difference between ARMA and ARIMA model ?
7. State the purpose of diagnostic checking.
8. What do you understand by subjectivity in forecasting ?
9. What are the commonly used estimation procedures in forecasting ?
10. Define spectral density.
11. What is the significance of autocorrelation plot ?
12. What is the use of periodogram ?

(12 × 1 = 12 weightage)

Part B

Answer any **eight** *questions.*
Weightage 2 *for each question.*

13. State the properties of auto-covariance function.
14. What is exponential smoothing ?
15. What is spectral plot ? State its use.

16. Show that MA(q) process is covariance stationary.

17. Indicate how the initial estimates of an MA(q) process would be obtained.

18. What is meant by invertibility of a time series model ?

19. Draw the limitations of forecasting.

20. What is meant by residual analysis ?

21. How would you check the adequacy of a model ?

22. What are ARCH and GARCH models ?

23. Define correlogram and periodogram.

24. Describe generalized autoregressive conditional heteroscedasticity (GARCH) model.

(8 × 2 = 16 weightage)

Part - C

Answer any **two** *questions.*
Weightage 4 *for each question.*

25. Describe the graphical procedures of detecting seasonality in time series data.

26. Discuss the stationarity and invertibility properties of an MA(q) process.

27. Derive the least square estimate of the parameter involved in AR(1) model.

28. Explain the method of estimating the lag length of a ARCH (q) model.

(2 × 4 = 8 weightage)

C 22725 (Pages : 2) Name.......................................

Reg. No..................................

FOURTH SEMESTER M.Sc. DEGREE (SUPPLEMENTARY) EXAMINATION APRIL 2022

(CUCSS)

Statistics

ST 4C 13—MULTIVARIATE ANALYSIS

(2013 to 2018 Admissions)

Time : Three Hours Maximum : 36 Weightage

Part A

Answer **all** *questions.*
Weightage 1 for each question.

1. Define multivariate non-singular normal distribution. Identify its parameters.
2. What do you mean by spherical normal distribution ?
3. Let $X^{(1)}$ and $X^{(2)}$ are two sub-vectors of X and X has a *p*-variate normal distribution, give the conditional distribution of $X^{(1)}$ and $X^{(2)} = x^{(2)}$.
4. Define multiple correlation coefficient in the multivariate case.
5. Define Wishart distribution.
6. Give the maximum likelihood estimator of the parameters of $N_p(\mu, \Sigma)$ and state whether they are unbiased.
7. What do you understand by generalized variance ?
8. Define Hotelling's T^2 statistic.
9. Give the asymptotic distribution of likelihood criterion.
10. What is multivariate Fisher-Behren problem ?
11. Define principal components.
12. What is the problem discussed in discriminant analysis ?

($12 \times 1 = 12$ weightage)

Part B

Answer any **eight** *questions.*
Weightage 2 for each question.

13. If $X \sim N_p(\mu, \Sigma)$, derive the distribution of $Y = CX$, where C is a non-singular matrix.
14. Derive the characteristic function of X, if $X \sim N_p(\mu, \Sigma)$.

Turn over

15. Find the characteristic function of a Wishart distribution.

16. Obtain the unbiased estimator of μ in $N_p(\mu, \Sigma)$.

17. Explain Canonical variables and canonical correlation.

18. If $X \sim N_p(0, I_p)$, obtain a necessary and sufficient condition for the independence of the quadratic forms $X'AX$ and $X'BX$.

19. If $X \sim N_p(\mu, \Sigma)$, Explain the procedure for testing $H_0 : \mu = \mu_0$.

20. Describe the application of Mahalanobis D^2 in testing problem.

21. Explain the sphericity test.

22. Describe the problem of classification.

23. What is factor analysis?

24. Explain the uses of principal components.

(8 × 2 = 16 weightage)

Part C

Answer any **two** *questions.*
Weightage 4 for each question.

25. Let $X \sim N_p(\mu, \Sigma), \Sigma$ is positive definite. Show that $Q = (X-\mu)'\Sigma^{-1}(X-\mu)$ is a χ^2 variate with p degrees of freedom.

26. A Wishart matrix W is partitioned into $W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$. Find the distribution of $W_{11.2} = W_{11} - W_{12} W_{22}^{-1} W_{21}$.

27. Derive the likelihood ratio test for testing the equality of mean vectors of two multivariate normal population having the same covariance matrix (unknown).

28. Explain the procedure to construct principal components when the population dispersion matrix is known.

(2 × 4 = 8 weightage)

C 22614 (Pages : 2) Name..................................

Reg. No..............................

FOURTH SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2022

(CBCSS)

Statistics

MST 4E 23—STATISTICAL MODELLING AND DATA MINING TECHNIQUES

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
2 weightages each.

1. Define knowledge based system.
2. Identify the importance of evolution analysis in data mining.
3. List the significance of data discretization in data mining tasks.
4. Identify the need for model adequacy checking.
5. Write the type of tasks performed in OLTP.
6. Interpret the significance of regression parameters in linear regression.
7. What is the purpose of association analysis ?

(4 × 2 = 8 weightage)

Turn over

Part B

Answer any **four** *questions.*
3 *weightages each.*

8. Compare classification and clustering with examples.
9. Illustrate how concept hierarchy is constructed over the attribute "location".
10. How are organizations using the information from data warehouses ?
11. Compare support and confidence with examples.
12. Discuss the purpose of statistical modeling.
13. How can you fit the model to data in regression analysis ?
14. Write the significance of data cube in dataware house.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
5 *weightages each.*

15. Explain the different data transformation techniques.
16. Compare OLAP and OLTP.
17. Explain Ridge regression and robust regression with equations.
18. Describe the classification of data mining systems.

(2 × 5 = 10 weightage)

C 22613 (Pages : 2) Name.......................................

Reg. No..................................

FOURTH SEMESTER M.Sc. DEGREE [REGULAR/SUPPLEMENTARY] EXAMINATION, APRIL 2022

(CBCSS)

Statistics

MST4E 18—DATA MINING TECHNIQUES

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each carries 2 weightage.

1. Distinguish between clustering and classification.
2. What are decision trees ?
3. What is meant by training set ?
4. What are principal components ? Mention their use data mining.
5. What is meant by back propagation in artificial neural networks ?
6. Mention any two measures used in developing decision trees.
7. Give an example for a transactional data set.

(4 × 2 = 8 weightage)

Turn over

Part B

Answer any **four** *questions.*
Each carries 3 *weightage.*

8. Explain the steps involved in developing decision trees.
9. Explain any two approaches used in clustering.
10. List the steps involved in k-nearest neighbourhood classification.
11. Explain various components of an artificial neural network.
12. Write a descriptive note on relational databases.
13. Explain the following terms : Support of a rule, Confidence of a rule with examples.
14. How many rules one can form using a k-item data set ? List all possible rules based on the item set $\{b, e, g, f\}$

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each carries 5 *weightage.*

15. Explain any two classification method based on multivariate analysis.
16. Explain naïve-Bayesian classification with the help of an example training data set
17. (a) List any two activation functions used in artificial neural networks and give their applications.

    (b) Explain how a regression model can be implemented using ANN
18. (a) Distinguish between data ware housing and datamining.

    (b) Write a note on online analytical data processing.

(2 × 5 = 10 weightage)

C 22612 (Pages : 2) Name.................................

Reg. No..............................

FOURTH SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2022

(CBCSS)

Statistics

MST 4E 13—BIOSTATISTICS

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions*
Each question carries weightage of 2.

1. Write down the types of Biological Data.
2. List the applications of gamma distribution in survival analysis.
3. Explain type I and type II random censoring with the help of an example.
4. What are the uses of actuarial estimator in survival analysis ?
5. Give a situation where logistic regression can be used.
6. Define mutation with an example.
7. What is the importance of randomization ?

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions*
Each question carries weightage of 3.

8. Describe L.R. tests.
9. Write down the applications of Weibull distribution in survival analysis.
10. Derive the Kaplan-Meier estimator using a suitable example.
11. Explain how will you estimate the mean survival time for type I and type II censored data.
12. Wrife a short note on competing risks in survival data.
13. Explain Mantel's Law of genetics using a suitable example.
14. Describe the method for sample size determination of fixed sample designs.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries weightage of 5.

15. Explain the Cox's F-test for comparing two exponential survival distributions and describe the test procedure.
16. What are the non-parametric methods for estimating survival functions ?
17. Estimate the probabilities of death under risks by ML method.
18. Write notes on design and planning of sequential, comparative and randomized control trials.

(2 × 5 = 10 weightage)

C 22611 (Pages : 3) Name..................................

Reg. No...............................

FOURTH SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2022

Statistics

MST 4E 10—STATISTICAL QUALITY CONTROL

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries weightage 2.

1. Distinguish between chance causes and assignable causes in SPC.
2. State true or false and give reasons. "If 8 points fall consecutively on one side of the central line, the process can be judged out of control".
3. Justify the use of three-sigma control charts.
4. State in what way OC and ARL functions of a control chart are useful.
5. A p-chart has central line = 0.04, LCL = 0.005 and UCL = 0.075, n = 100. Find ARL if process proportion defective shifts to 0.06.
6. In a cloth manufacturing industry, the number of imperfect spots across eveiy 50 square metre is to be controlled. The measuring practice shows that the area inspected varies from 40 to 60 square metre. What kind of control chart is appropriate for this process ? Obtain the control limits by clearly giving the details.

Turn over

7. Illustrate the concept of shift in the process average. What type of control charts are useful in identifying shifts of small magnitude in process average ?

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions.*
Each question carries weightage 3.

8. Describe the main features of double sampling plan by attributes. Define O.C. and A.S.N. functions of double sampling plan.

9. How does a sequential sampling plan differ from other types of sampling plans? Bring out the advantages of the sequential sampling plan.

10. An industrial manufacturing process is known to be working to a standard p_1 fraction defective. A new process is anticipated to reduce this to p_2. Derive a suitable sequential plan with instructions for use, given that α and β are the risks of Type I and Type II error respectively.

11. Write short notes on continuous sampling plans I, II and II.

12. Explain the construction and use of EWMA charts.

13. Distinguish between control chart for defectives and control chart for defects. Explain the methods of construction and analysis used in each of the above charts.

14. EMI devices Inc. Manufactures a relay, which incorporates a critical part, an electromagnetic coil. The specified resistance is 7.5 ohms ± 0.5 ohms. From the past experience, it is well known that the resistance varies according to a normal distribution. The process is operating with a mean of 7.4 ohms and standard deviation of 0.22 ohms.

    (a) Obtain process capability index for this process. Is the process capable ?

    (b) Which of the following measures you will choose to improve the process ? Why ?

        (1) Work on improving the process average.

        (2) Work on improving the process spread.

    (c) Calculate the percentage of defectives of the current process.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries weightage 5.

15. Derive the expressions for control chart constants used for $\bar{X}$ and R charts based on 3 sigma limits and 0.002 probability limits when :

    (a) Standard values are given ; and

    (b) standard values are not given.

16. (a) Explain the construction and uses of the C-chart. Is the use of 3σ limits justified for C-charts ?

    (b) Explain ASN and ATI curves of sampling plan by attributes. Discuss how these curves can be used in selecting inspection plan.

17. Determine the equations of rejection and acceptance lines for an item by item sequential plan in which AQL = 0.10, LTPD = 0.20, Producer's risk = 0.05 and consumer's risk = 0.10.

18. Describe the various ways in which a control chart may be modified to meet special situations. Explain how you can find out lack of control even if all points are within the control limits.

(2 × 5 = 10 weightage)

C 22610 (Pages : 3) Name.......................................

Reg. No..................................

FOURTH SEMESTER M.Sc. DEGREE [REGULAR/SUPPLEMENTARY] EXAMINATION, APRIL 2022

(CBCSS)

Statistics

MST4E 08—RELIABILITY MODELLING

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries a weightage 2.

1. Obtain the structure function of a parallel system of n components.
2. Define a Coherent system.
3. Show that DMRL $\Rightarrow$ NBUE.
4. Define positive ageing in Reliability.
5. Explain censored samples in Reliability estimation.
6. Define the reliability function of a bivariate random vector.
7. What is corrective maintenance ?

(4 × 2 = 8 weightage)

Turn over

Part B

Answer any **four** *questions.*
Each question carries a weightage 3.

8. Prove that redundancy at component level is more effective than redundancy at system level.

9. Find the minimal paths and minimal cuts for the following structure function configuration. $\varnothing(\underline{x}) = \max(x_1, x_3)\, x_5 \max(x_2, x_4)$. Draw the structure diagram.

10. Prove or disprove IFRA $\Rightarrow$ DMRL.

11. Explain the role of exponential distribution in Reliability analysis.

12. State and prove the IFRA closure property of a Coherent system of n independent components.

13. Let the joint survival function $\bar{F}(x, y)$ of a bivariate random vector (X, Y) satisfies the BVLMP, then show that $\bar{F}(x, y) = \begin{cases} e^{-\theta y}\, \bar{F}_1(x - y), \text{ for } x > y \\ e^{-\theta x}\, \bar{F}_2(y - x), \text{ for } y > x, \end{cases}$ where $\bar{F}_1(\cdot)$ and $\bar{F}_2(\cdot)$ are the marginal survival functions.

14. What are the assumptions made in developing a model for the availability analysis of a system ?

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries a weightage 5.

15. Discuss the structural importance of components : (i) Let $\varnothing$ be the structure function of a 2-out-of-3 system. Then find the critical path vectors for the components and order them according to their structural importance ; and (ii) Let $\varnothing(\underline{x}) = x_1 (x_2 \amalg x_3)$, then find the importance of components and order them.

16. Derive the inter relations between failure(hazard) rate, mean residual life function and the reliability function.

17. Sixty items are put to test and the test is continued until 10 items failed. The failure times in hours are recorded as 85, 151, 280, 376, 492, 520, 623, 715, 820, 914. Assuming the failure times to be exponentially distributed, estimate the MLE and UMVUE of the parameter and the reliability function at t = 600 hours, if the failed items are :

(i) Not replaced ; and (ii) Replaced.

18. Discuss the test for Homogeneous Poisson Process against the Non-Homogeneous Poisson Process for repairable systems.

(2 × 5 = 10 weightage)

C 22609 (Pages : 3) Name.......................................

Reg. No...................................

FOURTH SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2022

(CBCSS)

Statistics

MST 4E 07—STATISTICAL DECISION THEORY

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries 2 *weightage.*

1. Define decision function with an example. What is the difference between decision making under risk and decision making under uncertainty ?
2. Define loss function. Give different types of loss functions.
3. Suppose there exists a prior distribution π for 0 for which expected risk of Bayes decision rule is constant. Show that π is a least favorable prior distribution.
4. What is admissibility of estimators ? Show that a Bayes estimator if it is unique is admissible.
5. Explain (i) Two person zero sum game, and (ii) pay-off matrix with an example.
6. Describe maximin-minimax principle in Game theory.
7. Define Saddle point. What are the rules to determine saddle point in Game theory.

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions.*
Each question carries 3 *weightage.*

8. Define (i) Risk function, (ii) Bayes risk and (iii) Bayes estimator. Show that Bayes estimator is the mean of posterior distribution under squared error loss function.

9. Define (i) complete and (ii) minimal complete decision rules. Prove that if a class of admissible decision rules is complete, it must be minimal complete.

10. Explain prior and posterior distributions. Find posterior distribution of parameter θ of Uniform $(0,\theta)$ distribution based on a random sample of size n, assume that prior distribution of θ is uniformly distributed over (0, 1).

11. Write short notes on the following :

    (i) Non-informative priors ; (ii) Maximum entropy priors ; and (iii) Bayesian robustness.

12. Explain (i) pure and mixed strategies, (ii) games without saddle point and (iii) Principles of dominance.

13. Two breakfast food manufactures ABC and XYZ competing for an increased market share. The pay-off matrix, shown in the following table, describes the increase in market share of ABC and decrease in market share of XYZ :

ABC	XYZ			
	Give coupons	Decrease in price	Maintain present strategy	Increase advertizing
Give coupons	2	–2	4	1
Decrease in price	6	1	12	3
Maintain present strategy	–3	2	0	6
Increase advertizing	2	–3	7	1

Determine the optimum strategies for both manufacturers and value of the game.

14. Explain the role of statistical decision theory to solve real problems.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries 5 weightage.

15. Explain Bayes and minimax decision rules. Find the minimax estimator of the parameter p in Bernoulli distribution under squared error loss function, assuming beta prior distribution.

16. (i) Explain conjugate prior distributions. How to determine conjugate prior distributions of Bernoulli, Poisson and normal distributions ?

    (ii) Prove that a Bayes estimator with constant risk is minimax.

17. Describe the graphical method of solving $2 \times n$ game. Use Graphical method to solve the following game :

	Player B		
Player A	B_1	B_2	B_3
A_1	1	3	11
A_2	8	5	2

18. (i) Let $X_1, X_2, ..., X_n$ be a random sample from $N(\mu,1)$. Assuming the prior distribution of μ is N (0,1), find the Bayes estimator of μ using squared error loss function.

    (ii) Explain empirical Bayes analysis.

(2 × 5 = 10 weightage)

C 22607 (Pages : 3) Name.....................................

Reg. No...................................

FOURTH SEMESTER M.Sc. (CBCSS) DEGREE [REGULAR/ SUPPLEMENTARY] EXAMINATION, APRIL 2022

Statistics

MST 4C 14—MULTIVARIATE ANALYSIS

(2019 Admission onwards)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *The instruction if any, to attend a minimum number of questions from each sub section / sub part / sub division may be ignored.*
4. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries a weightage 2.

1. If $Y = tr(X)$, where X is a square matrix , then find $\frac{dY}{dX}$.

2. Find the distribution of the quadratic form $(X-\mu)' \Sigma^{-1} (X-\mu)$, if $X \sim N_p(\mu, \Sigma)$ and Σ is known.

3. Let $X \sim N_p(\mu, \Sigma)$. Find the distribution of $l'X$, where l is a real $px1$ vector.

4. Define a bivariate Normal random vector. Identify the parameters.

5. What is the relationship between Hotelling's T^2 and Mahalanobis D^2 statistic ?

6. What is the purpose of a scree-plot ?

7. Distinguish between Confirmatory Factor Analysis and Exploratory factor analysis.

(4 × 2 = 8 weightage)

Turn over

Part B

Answer any **four** *questions.*
Each question carries a weightage 3.

8. Let $X \sim N_p(0, \Sigma)$, then write the necessary and sufficient condition for the independence of the quadratic forms $X'AX$ and $X'BX$ where A and B are real symmetric matrices.

9. Show that $\frac{1}{N}\sum_{\alpha=1}^{N}(X_\alpha - \bar{X})(X_\alpha - \bar{X})'$ is a biased estimate of the population dispersion matrix Σ. Obtain an unbiased estimate of it.

10. If (X, Y) follows bivariate Normal distribution then show that X + Y and X – Y are independent if and only if $\sigma_x^2 = \sigma_y^2$. Use this result to prove the independence of $\bar{X}$ and s^2 when a sample of size two is taken from a univariate Normal population.

11. Explain T^2 statistic as a likelihood ratio criterion.

12. Show that every principal submatrix of a Wishart Matrix is again Wishart.

13. Explain the Sphericity test.

14. Describe a method of estimating the principal components while sampling from a multivariate normal population.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries a weightage 5.

15. Suppose $X = (X_1, X_2, X_3)' \sim N_3(\mu, \Sigma)$, where $\mu = \begin{bmatrix} 3 \\ 10 \\ 8 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 16 & 2 \\ 1 & 2 & 4 \end{bmatrix}$. Find :

(i) The distribution of X_1 ; (ii) The distribution of (X_1, X_2) ; (iii) The distribution of $X_1 + X_2 + X_3$; and (iv) The conditional distribution of X_3 given $X_1 = 5$ and $X_2 = 3$.

16. Obtain the distribution of the simple correlation coefficient in the null case.

17. Derive the M.L.E. of the mean vector and the dispersion matrix of a multivariate normal distribution.

18. What do you mean by principal Component Analysis ? Establish the relationship between principal components and eigen-value, eigen-vector structure of the variance-covariance matrix.

($2 \times 5 = 10$ weightage)

C 21088 (Pages : 3) Name..............................

Reg. No...........................

FOURTH SEMESTER P.G. DEGREE EXAMINATION, APRIL 2022

(CCSS)

Statistics

STA 4E 04—OPERATION RESEARCH—II

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Use of Calculator is permitted.

Part I

Answer any **four** *questions.*
Each question carries 4 marks.

I. (a) Define general non-linear programming problem and state the necessary conditions for optimum.

(b) What are the basic features which characterize a dynamic programming problem ?

(c) Distinguish between continuous review and periodic review inventory systems.

(d) Explain group replacement policy.

(e) Describe the importance of demand and lead time in inventory management system.

(f) What is replacement problem ? Describe briefly the situations when the replacement of certain items needs to be done.

(g) Solve the following linear programming problem using the method of Lagrangian's multipliers :

Maximize $f(x_1, x_2) = 3x_1^2 + 6x_1 + 2x_1x_2 + 2x_2 + x_2^2$

subject to the constraints : $2x_1 - x_2 = 4$ and $x_1, x_2 \geq 0$.

(h) What is simulation ? Explain event type simulation.

(4 × 4 = 16 marks)

Part II

Answer **either** *Part A* **or** *Part B of all questions.*
Each question carries 16 marks.

II. A (i) Derive the Kuhn-Tucker conditons for the following non-linear programming problem :

Optmize $f(x_1, x_2, \ldots, x_n)$ subject to the constraints

$g(x_1, x_2, \ldots, x_n) \leq C$, where C is a constant and $x_1, x_2, \ldots, x_n \geq 0$.

Also, establish the sufficiency of Kuhn-Tucker conditions.

Turn over

(ii) Define Saddle point. Write short notes on saddle point criteria of non-linear programming problems.

(12 + 4 = 16 marks)

Or

B (i) Define quadratic programming problem. What are its basic characteristics ? How does a quadratic programming problem differ from a linear programming problem ?

(ii) Solve the following quadratic programming problem by Wolfe's method :

Maximize $Z = 4x_1 - 2x_1^2 + 6x_2 - 2x_2^2 - 2x_1x_2$

subject to the constraints :

$x_1 + 2x_2 \leq 2$ and $x_1, x_2 \geq 0$.

(5 + 11 = 16 marks)

III. A (i) Define dynamic programming problem and state Bellman's principle of optimality.

(ii) Explain the recursive equation approach to solve dynamic programming problem.

(iii) Use dynamic programming to solve the following problem :

Minimize $Z = x_1^2 + x_2^2 + x_3^2$

subject to the constraints :

$x_1 + x_2 + x_2 \geq 15$ and $x_1, x_2, x_3 \geq 0$.

(4 + 5 + 7 = 16 marks)

Or

B (i) Define geometric programming problem. Explain the technique of solving unconstrained geometric programming problem.

(ii) Using geometric programming, solve the following problem :

Minimize $f(x_1, x_2) = \dfrac{10}{x_1x_2} + 2x_1 + 4x_2$ and $x_1, x_2 > 0$.

(9 + 7 = 16 marks)

IV. A (i) Explain inventory system. Describe briefly the costs associate with inventory management problems.

(ii) Develop a deterministic EOQ model where shortages are not allowed. Find also its optimum size.

(iii) A manufacturer has to supply his customer with 600 units of his product per year. Shortages are not allowed and the storage cost amount to Rs. 0.60 per unit per year. The set up cost per run is Rs. 80. Find the optimum run size and minimum average yearly cost.

(6 + 5 + 5 = 16 marks)

Or

B (i) Explain newsboy problem.

(ii) A newspaper boy buys papers for Rs. 6.00 and sells them for Rs. 7.50 each. He cannot return unsold newspapers. Daily demand has the following probability distribution :

No. of customers	150	151	152	153	154	155	156	157	158	159	160
Probability	0.02	0.04	0.07	0.12	0.12	0.15	0.16	0.14	0.10	0.06	0.02

If each day is independent of the previous day, how many papers he should order each day ?

(iii) Obtain expression for expected total cost function for an inventory problem where demand is uncertain and unfilled demand during lead time is backlogged.

(5 + 5 + 6 = 16 marks)

V. A (i) Develop a model for the replacement of equipment whose maintenance costs increase with time and value of money does not change with time.

(ii) A transport manager finds from his part records that the costs per year of running a truck whose purchase price is Rs. 6,000 are as given below :

Year	1	2	3	4	5	6	7	8
Running cost	1000	1200	1400	1800	2300	2800	3400	4000
Resale value	3000	1500	750	375	200	200	200	200

Determine at what age replacement is due.

(6 + 10 = 16 marks)

Or

B Write notes on the following :—

(i) Monte Carlo simulation technique.

(ii) Simulation of a single server queuing model.

(iii) Generation of random samples from Erland distribution using convolution method.

(iv) Production lot size inventory model.

(4 × 4 = 16 marks)

[4 × 16 = 64 marks]

C 21087 (Pages : 2) Name..............................

Reg. No...........................

FOURTH SEMESTER P.G. DEGREE EXAMINATION, APRIL 2022

(CCSS)

Statistics

STA 4E 03—LIFE TIME DATA ANALYSIS

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (a) Explain progressive type II censoring.

(b) Write a short note on Inverse Gaussian distribution.

(c) Derive Greenwood's formula for variance of Kaplan-Meier estimate.

(d) Briefly describe Quantile-Quantile plot.

(e) Explain briefly the general methods for comparing two location-scale distributions.

(f) Obtain an exact confidence interval for the parameter when the lifetimes follow exponential distribution.

(g) Distinguish between marginal, conditional and partial likelihoods.

(h) Explain log-rank test.

($4 \times 4 = 16$ marks)

Section B

Answer either Part A **or** *Part B of all questions.*
Each question carries 16 *marks.*

II. A (a) Discuss the monotonicity of hazard function of Gamma distribution.

(b) Briefly describe discrete and continuous mixture models.

($8 + 8 = 16$ marks)

Or

B (a) Define hazard rate in the discrete set up. Also show that it uniquely determines the survivor function.

(b) Describe log-logistic distribution in the context of life time data analysis.

($8 + 8 = 16$ marks)

Turn over

III. A (a) What do you mean by interval censored data ? How will you estimate survivor function for interval censored data ?

(b) Explain standard life table methodology.

(8 + 8 = 16 marks)

Or

B (a) Show that Product-Limit estimate of survivor function can be derived as nonparametric MLE.

(b) Write a short note on interval estimation of survival probabilities.

(8 + 8 = 16 marks)

IV. A (a) Explain inference procedures based on large sample theory for exponential distribution.

(b) The following data concerns the lifetime of 10 pieces of equipment. Assuming exponential with p.d.f. $f(t) = \frac{1}{\theta} e^{\frac{-1}{\theta}}$ find the MLE of θ and obtain 95 % confidence interval for θ.

t_i	2	72	51	60	33	27	14	24	4	21
δ_i	1	0	1	0	1	1	1	1	1	0

(8 + 8 = 16 marks)

Or

B (a) Explain likelihood based inference procedures for log location scale distributions.

(b) Define two parameter exponential distribution. Briefly describe the likelihood based inference procedures for two parameter exponential distribution.

(8 + 8 = 16 marks)

V. A (a) Define Cox proportional hazard model. Derive Cox likelihood as a marginal likelihood.

(b) Describe the generalized Wilcoxon test with the censored data.

(10 + 6 = 16 marks)

Or

B (a) Discuss briefly the linear rank tests for the m-sample problem.

(b) Briefly describe the inference procedures for accelerated failure time models.

(8 + 8 = 16 marks)

[4 × 16 = 64 marks]

C 21087 (Pages : 2) Name.............................

Reg. No...........................

FOURTH SEMESTER P.G. DEGREE EXAMINATION, APRIL 2022

(CCSS)

Statistics

STA 4E 03—LIFE TIME DATA ANALYSIS

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (a) Explain progressive type II censoring.

(b) Write a short note on Inverse Gaussian distribution.

(c) Derive Greenwood's formula for variance of Kaplan-Meier estimate.

(d) Briefly describe Quantile-Quantile plot.

(e) Explain briefly the general methods for comparing two location-scale distributions.

(f) Obtain an exact confidence interval for the parameter when the lifetimes follow exponential distribution.

(g) Distinguish between marginal, conditional and partial likelihoods.

(h) Explain log-rank test.

($4 \times 4 = 16$ marks)

Section B

Answer either Part A **or** *Part B of all questions.*
Each question carries 16 *marks.*

II. A (a) Discuss the monotonicity of hazard function of Gamma distribution.

(b) Briefly describe discrete and continuous mixture models.

($8 + 8 = 16$ marks)

Or

B (a) Define hazard rate in the discrete set up. Also show that it uniquely determines the survivor function.

(b) Describe log-logistic distribution in the context of life time data analysis.

($8 + 8 = 16$ marks)

Turn over

III. A (a) What do you mean by interval censored data ? How will you estimate survivor function for interval censored data ?

(b) Explain standard life table methodology.

(8 + 8 = 16 marks)

Or

B (a) Show that Product-Limit estimate of survivor function can be derived as nonparametric MLE.

(b) Write a short note on interval estimation of survival probabilities.

(8 + 8 = 16 marks)

IV. A (a) Explain inference procedures based on large sample theory for exponential distribution.

(b) The following data concerns the lifetime of 10 pieces of equipment. Assuming exponential with p.d.f. $f(t) = \frac{1}{\theta} e^{\frac{-1}{\theta}}$ find the MLE of θ and obtain 95 % confidence interval for θ.

t_i	2	72	51	60	33	27	14	24	4	21
δ_i	1	0	1	0	1	1	1	1	1	0

(8 + 8 = 16 marks)

Or

B (a) Explain likelihood based inference procedures for log location scale distributions.

(b) Define two parameter exponential distribution. Briefly describe the likelihood based inference procedures for two parameter exponential distribution.

(8 + 8 = 16 marks)

V. A (a) Define Cox proportional hazard model. Derive Cox likelihood as a marginal likelihood.

(b) Describe the generalized Wilcoxon test with the censored data.

(10 + 6 = 16 marks)

Or

B (a) Discuss briefly the linear rank tests for the m-sample problem.

(b) Briefly describe the inference procedures for accelerated failure time models.

(8 + 8 = 16 marks)

[4 × 16 = 64 marks]

C. 21086 (Pages : 3) Name..............................

Reg. No..........................

FOURTH SEMESTER P.G. DEGREE EXAMINATION, APRIL 2022

(CCSS)

Statistics

STA 4E 04—OPERATION RESEARCH—II

(2010 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Use of calculator is permitted.

Part I

Answer any **four** *questions.*
Each question carries 4 *marks.*

1. (a) Define (i) Non-linear programming problem ; and (ii) Quadratic programming problem.

   (b) What is inventory system ? Explain continuous review inventory system.

   (c) Define stage, slack variable, decision variable, return function and optimum return in dynamic programming problem.

   (d) What is simulation ? Explain inverse method of generating random observations in simulation techniques.

   (e) Distinguish between individual and group replacement policies.

   (f) Describe briefly the general procedure for obtaining optimum solution using the dynamic programming approach.

   (g) Explain the importance of demand and lead time in inventory management problem.

   (h) Solve the following non-linear programming problem using the method of Lagrangian's multipliers :

   Maximize $f(x_1, x_2) = 3x_1^2 + 6x_1 + 2x_1x_2 + 2x_2 + x_2^2$

   Subject to the constraints $2x_1 - x_2 = 4$ and $x_1, x_2 \geq 0$.

(4 × 4 = 16 marks)

Part II

Answer either Part A **or** *B of all questions.*
Each question carries 16 *marks.*

II. A Derive the Kuhn-Tucker conditions for unconstrained and constrained non-linear programming problems. Also, establish the sufficiency of Kuhn-Tucker conditions.

Or

Turn over

B (i) Explain Wolfe's method for an optimum solution to QPP.

(ii) Use the Wolfe's method to solve the following QPP :

Maximize $Z = 4x + 6y - x^2 - 3y^2$

Subject to $x + 2y \leq 4$ and $x, y \geq 0$.

(8 + 8 = 16 marks)

III. A (i) Explain geometry programming problem. Describe the technique of solving unconstrained geometric programming problem.

(ii) Using geometric programming, solve the following optimization problem :

$$\text{Minimize } f(x_1, x_2) = \frac{10}{x_1 x_2} + 2x_1 + 4x_2 \text{ and } x_1, x_2 > 0.$$

(9 + 7 =- 16 marks)

Or

B (i) Derive dynamic programming problem and state Bellman's principle of optimality. Explain the recursive equation approah to solve dynamic programming problem.

(ii) Solve the following problem by dynamic programming method :

Minimize $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$

Subject to $x_1 + x_2 + x_3 = 10$ and $x_1, x_2, x_3 \geq 0$.

(10 + 6 = 16 marks)

IV. A (i) Explain various costs and revenues associated with an inventory management system.

(ii) Stating the assumptions, derive the classic EOQ formula.

(iii) The annual consumption of material is 3600 units, ordering costs are Rs. 4 per order. The unit price of material is 0.64 rupees and storage costs are 50 % per annum of stock value. Find the economic order quantity.

(6 + 6 + 4 = 16 marks)

Or

B (i) Develop a news boy model and find its optimum solution.

(ii) The demand distribution of daily sales of a newspaper is as follows :

Daily sales	:	1000	2000	3000	4000	5000	6000	7000
Probability	:	0.01	0.06	0.25	0.35	0.20	0.03	0.10

The cost of carrying inventory is Rs. 30 per unit per day and cost of unit shortage is Rs. 70 per day. Determine the optimum stock level which minimizes the total expected cost.

(iii) Explain backorders and lost sales cases in inventory management problems.

(7 + 5 + 4 = 16 marks)

V. A (i) Explain different types of replacement problems in real situations.

(ii) A firm is considering replacement of equipment whose first cost is Rs. 1,750 and the scrap value is negligible at any year. Based on experience, it is found that maintenance cost is zero during the first year and it increases by Rs. 100 every year thereafter. When should be the equipment replaced if the annual interest rate is 12 % ?

(iii) Write short notes on staff replacement problem.

(5 + 8 + 3 = 16 marks)

Or

B (i) Explain why simulation is needed ? What is event type simulation ?

(ii) Describe Monte-Carlo simulation technique.

(iii) Write short notes on the following :—

(a) Convolution method.

(b) Acceptance Rejection method.

(4 + 7 + 5 = 16 marks)

[4 × 16 = 64 marks]