M.Sc. Examination, 2024
Semester-III
Statistics
Course: MSC-31
Stochastic Process

Time: 3 hrs Full Marks:40

Answer **all** questions.

1. Write TRUE/FALSE by the following statements.
   (a) A state i is recurrent if $lim_{n\to\infty} p_{ii}{}^{(n)} = 0$.
   (b) For an infinite irreducible Markov chain all states are positive recurrent.
   (c) For a finite Markov chain, unique stationary distribution exists.
   (d) Consider a Markov Chain $X_0, X_1, \cdots$ with state space S . Suppose $i, j \in S$ are two states which communicate with each other. Then $lim_{n\to\infty} P(X_n = j/X_0 = i) = lim_{n\to\infty} P(X_n = j/X_0 = j)$.
   (e) $X_t = X_{t-1} + \epsilon_t$ where $\epsilon \sim N(0, 1)$. This model is covariance stationary.
   (f) For a Wiener process, measure of drift is function of time but measure of spread is function of time square.
   (g) Consider two independent series of events where both of the event follow Poisson law. Probability distribution that first event of the first process will happen before the second event of the second process is a F distribution with degree of freedom 2 and 2.
   (h) For an irreducible Markov chain, transition matrix is regular.

   8

2. (a) What is limiting distribution of a Markov chain? Give an example of a process which possesses limiting distribution but no stationary distribution.
   (b) Consider the Markov chain will state space $\{0, 1, 2, \cdots, k\}$ and transition probability is $p(m, m+1) = \frac{1}{2}(1 - \frac{1}{m+2})$ for $m \geq 0$, $p(m, m-1) = \frac{1}{2}(1 + \frac{1}{m+2})$ for $m \geq 1$ and $p(0, 0) = \frac{3}{4} = 1 - p(0, 1)$. Find the stationary distribution of the process.

   4+4

3. (a) Ron, Sue, and Ted arrive at the beginning of a professor's office hours. The amount of time they will stay is exponentially distributed with means of 1, $\frac{1}{2}$, and $\frac{1}{3}$ hour. What is the expected time until only one student remains? What is the expected time until all three students are gone?
   (b) Suppose for a continuous time Markov chain inter-arrival time between two successive occurrences of events follow exponential distribution with mean 50 minute. Given that two events occurred within 12 noon, what is the probability that next two events will take place after 2:30 p.m.? What is the probability that 2 events have occurred within 11.30 a.m. given that 5 events occur within 1:15 p.m.?

   4+4

4. (a) Prove that for a closed set, probability of moving out of the set is zero.
   (b) Define renewal density. Show that renewal density is the derivative of renewal function.

   4+4

5. (a) Define a Martingale of order r. Let $U_1, U_2, \cdots$ be independent random variables each having uniform distribution $(0, 1)$. Let $X_n = 2^n U_1 \cdots U_n$. Show that $\{X_n\}$ is a martingale.
   (b) Let $W(t)$ be a standard Brownian motion. Define $X(t) = exp(W(t))$, for all $t \in [0, \infty)$. Find $V(X(t))$.
   Let $0 < s < t$. Find $Cov(X(s), X(t))$.

   4+4

M.Sc. Examination 2024
Semester III
Subject: Statistics
Paper: MSC-32 [Advanced Data Analysis Techniques]

Full Marks: 40 Time: 3 Hrs.

Answer any four questions (*Symbols have their usual meaning*)

1. a) For a given logistic regression model, if the odds ratio is 1, what does it indicate about the effect of the predictor X on the outcome?
   b) Compare Newton-Raphson, Gradient Descent and Iterative Reweighted Least Squares (IRLS) in the context of maximize the log-likelihood function.
   c) Write down the detailed steps involved in fitting a poisson regression model. 1+3+6

2. a) What is logistic regression, and how is it different from linear regression? Define the logit function and explain its role in logistic regression. Why is the logistic function used in modeling binary outcomes? What are the key assumptions of logistic regression? Explain the meaning of the regression coefficients in a logistic regression model.
   b) If a probit model yields a probability $P(Y = 1) = 0.6$, how would you calculate the corresponding z-score? In what situations would you prefer a logit model over a probit model? (2+1+1+2+2)+(1+1)

3. a) Briefly discuss the idea of Metropolis-Hestings Algorithm.
   b) Show that the choice of g which minimizes the variance of the estimator $\frac{1}{n}\sum_{i=1}^{n}\left(\frac{f(x_i)}{g(x_i)}\right)h(x_i)$ is given by $g^*(x) = \frac{|h(x)|f(x)}{\int_{\mathbb{X}}|h(x)|f(x)\,dx}$.
   c) Explain how to generate samples from $N(0, \sigma = 3)$ using the accept-reject method, with a double exponential as the proposal distribution. 3+3+4

4. a) If $s(x) = \frac{1}{n}\sum(x_i - \bar{x})^2$ then show that $Bias(\hat{F}) = -\frac{1}{n^2}\sum(x_i - \bar{x})^2$. Here $s(x)$ is an estimate of θ on the basis of sample and Bootstrap estimate of bias i.e. $Bias(\hat{F}) = E_{\hat{F}}[s(x^*)] - t(\hat{F})$.
   b) What is the Jackknife method, and how does it differ from the Bootstrap method? Explain the steps involved in performing the Jackknife procedure for a given statistic. 4+(3+3)

5. a) How can we generate Gibbs samples in the context of a bivariate probability distribution? Discuss the process in detail, including an example.
   b) Suppose, we are studying the distribution of the number of defectives X in the daily production. Consider the model $X \mid Y \sim Bin(Y, \theta)$ and $Y \sim Poisson(\lambda)$. Consider a $Beta(a, b)$ prior for θ and show how Gibbs Sampling can be implemented to sample from the posterior distribution of $\theta \mid X$.
   5+5

6. a) Let $Y_i \sim i.i.d.N(\theta, 1),\ i = 1(1)n$. Here, $Y_1, Y_2, \cdots, Y_m$ are the observed values and true values $Z_{m+1}, Z_{m+2}, \cdots, Z_n$ are unobserved. We observe $Y_i = a$ if $Y_i > a$ for $i = \overline{m+1}(1)n$. Using the EM (Expectation- Maximization) algorithm, derive the recurrence relation to estimate θ.
   b) Write a brief note on the convergence of the EM algorithm. 6+4

M.Sc. Examination, 2024
Semester-III
Statistics
Course: MSC-33(MSE-1)
(Operations Research and Optimization Techniques)
Time: Three Hours Full Marks: 40

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **any five** questions

1. What is a transportation problem? Is it considered to be a Linear Programming Problem? Show that a balanced transportation problem always has a feasible solution. 2+3+3

2. Briefly state the role of modeling in Operations Research. Mention different types of models and their solutions. 2+6

3. For the M/M/1 queuing system find the expected number of customers in the system in the steady state and also the expected queue length. Find the cumulative distribution function for the waiting time of a customer who has to wait in an M/M/1 queuing system. 4+4

4. What is two-person zero-sum game? Transform this game to a Linear Programming Problem. Prove that if mixed strategies be allowed, then there always exists a value of the game. 2+2+4

5. (a) Define an inventory. What are the advantages and disadvantages of having inventories?
(b) Suppose that Q^* is the optimal order quantity and K^* is the corresponding minimum annual variable cost. Show that if a value of $Q = (1+\alpha)Q^*$ is used, $\frac{K}{K^*} = 1 + \frac{\alpha^2}{2(1+\alpha)}$, where K is the annual variable cost corresponding to an order quantity Q. 4+4

6. (a) What is replacement Problem? Give some illustrations. Discuss replacement policy of equipments that deteriorates gradually with change in time value of money.
(b) What is preventive replacement? Find out criterion for optimal replacement time in such situation. 4+4

7. (a) Distinguish between deterministic and probabilistic models of inventory.
(b) For an inventory model, if P(r) denotes the probability of requiring r units, where r is a discrete variable, C_1 is the inventory holding cost per unit of time, C_2 is the shortage cost per unit per unit of time, then show that the stock level which minimizes the total expected cost is that value of S which satisfies the conditions:

$$\sum_{r=0}^{S-1} P(r) < \frac{C_2}{C_1 + C_2} < \sum_{r=0}^{S} P(r).$$ 4+4

8. Write short notes on any two of the following: 4+4

(a) (s, S) inventory policy
(b) Duality problem in LPP
(c) Saddle point in game theory
(d) Congestion factor in Queuing model

MSc in Statistics Sem-III Examination 2024

Subject: Statistics

Paper: MSC-34/MSS-3
(Time Series Analysis)

Full Marks: 40 | Time: **Three** Hours

Answer any **four** from the following seven questions of equal marks. Symbols are of usual meanings.

1. (a) Derive the Yule-Walker equations for an AR(p) process. Clearly show how these equations relate to the autocorrelation function (ACF) for the model parameters.

   (b) Explain how one would use the sample autocorrelation function to estimate the parameters in an AR(p) process via the Yule-Walker approach. Discuss any limitations of this method.

   $5 + 5 = 10$

2. (a) Show that an AR(1) process can be expressed as an infinite MA series. Specify the conditions under which this infinite series converges.

   (b) Derive the partial autocorrelation function (PACF) for an AR(2) process. Illustrate how the PACF values for lags 1 and 2 depend on the AR parameters.

   $5 + 5 = 10$

3. Consider the following MA(3) process

$$Y_t = \mu + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \theta_3 e_{t-3}$$

   where e_t is a zero mean white noise process with variance σ^2, μ is a constant and θ_i's are MA parameters. Derive the auto-correlation coefficients of Y_t.

   10

4. (a) Present the Holt-Winters equations for level, trend, and seasonal components (additive version). Discuss how these components are updated recursively.

   (b) How do you modify the Holt-Winters method to handle *multiplicative* seasonality? Write down the adapted equations and discuss situations in which a multiplicative formulation is preferable.

$5 + 5 = 10$

5. (a) Describe the necessity of considering frequency domain analysis of time series with examples.

   (b) Assuming a suitable model, demonstrate how the hidden periodicity of time series can be estimated.

$4 + 6 = 10$

6. (a) Summarize the three main stages of the Box-Jenkins methodology for ARIMA modeling: *model identification*, *parameter estimation*, and *diagnostic checking*.

   (b) Write a short-note on *cross-correlation*.

$6 + 4 = 10$

7. (a) Is the time series $Y_t = Y_{t-1} - 0.5Y_{t-2} + e_t$, where e_t is white noise $(0, \sigma^2)$, covariance stationary? Answer with reasons.

   (b) Calculate Y_1, Y_2 and Y_5, Y_7.

   (c) If the coefficient of Y_{t-1} is 1.1 and that of Y_{t-2} is -0.18 then comment on the covariance stationarity of Y_t.

3+4+3

M.Sc. Examination 2024
Semester III
Subject: Statistics (Practical)
Paper: MSC-35

Full Marks: 40 Time: 4 Hrs.

Answer all questions (*Symbols have their usual meaning*)

1. Consider the dataset below and fit a multiple linear regression of Y on X_1, X_2, X_3, X_4 and X_5. Hence find the bootstrap estimate of the standard error of the regression coefficients. Draw histogram of the bootstrap replicates. 12

Index	Y	X_1	X_2	X_3	X_4	X_5
1	271.8	783.35	33.53	40.54	16.66	13.2
2	264.0	748.45	36.50	36.19	16.46	14.11
3	238.8	684.45	34.66	37.31	17.66	15.68
4	230.7	827.80	33.13	32.52	17.50	10.53
5	251.6	860.45	35.75	33.71	16.34	11.00
6	257.89	875.15	34.46	34.14	16.28	11.31
7	263.89	909.45	34.60	34.85	16.06	11.96
8	266.5	905.55	35.38	35.89	15.93	12.58
9	229.1	756.00	35.85	33.53	16.60	10.66
10	239.3	769.35	35.68	33.79	16.41	10.85
11	258.0	793.50	35.35	34.72	16.17	11.41
12	257.6	801.65	35.04	35.22	15.92	11.91
13	267.3	819.65	34.07	36.50	16.04	12.85
14	267.0	808.55	32.20	37.60	16.19	13.58
15	259.6	774.95	34.32	37.89	16.62	14.21
16	240.4	711.85	31.08	37.71	17.37	15.56
17	227.2	694.85	35.73	37.00	18.12	15.83
18	196.0	638.10	34.11	36.76	18.53	16.41
19	278.7	774.55	34.79	34.62	15.54	13.10
20	272.3	757.90	35.77	35.40	15.70	13.63
21	267.4	753.35	36.44	35.96	16.45	14.51
22	254.5	704.70	37.82	36.26	17.62	15.38
23	224.7	666.80	35.07	36.34	18.12	16.10
24	181.5	568.54	35.26	35.90	19.05	16.73
25	227.5	653.10	35.56	31.84	16.51	10.58

2. Generate Jackknife samples of size 50 from gamma distribution with your own parameters. 6

3. Simulate a random sample from the density $f(x) = \frac{2}{\sqrt{\pi}}\sqrt{x}e^{-x},\ x \geq 0$. Using a proposal density $g(x) = \lambda e^{-\lambda x},\ x \geq 0,\ \lambda = 0.4$. Also propose and justify an another $g(x)$ for simulation. 10

4. Simulate bivariate normal sample with mean $(2,3)'$ and $\Sigma = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$ by using Gibbs sampling. 7

5. Practical note book and viva-voce. 5

M.Sc Semester III Examination 2024

Statistics

MSC-36(Practical)

Time: Four Hours Full Marks: 40

1. Use the concept of dominance to convert the following game into 2x2 game and hence solve the game. 6

	Player B			
	I	II	III	IV
I	2	1	4	0
II	3	4	2	4
III	4	2	4	0
IV	0	4	0	8

2. Find out the optimal assignment and minimum cost for the assignment with the following cost matrix.

	I	II	III	IV	V
A	160	130	175	190	200
B	135	120	130	160	175
C	140	110	155	170	185
D	50	50	80	80	110
E	55	35	70	80	105

6

3. A supper market has two girls ringing up sales at the counters. If the service time for each counter is exponential with mean 4 minutes, and if people arrive in a Poisson fashion at the rate of 10 an hour, (a) what is the probability of having to wait for service? (b) What is the expected percentage of idle time for each girl? (c) Find the average queue length and the average number of units in the system. 3

4. A contractor has to supply 10,000 bearings per day to an automobile manufacturer. He finds that, when he starts a production run, he can produce 25,000 bearings per day. The cost of holding a bearing in stock for one year is 20 paise and the set up cost of a production run is Rs. 180. How frequently should production run be made? 3

5. A time series has the following sample **ACF and PACF** values:

Lag	**ACF**	**PACF**
1	0.65	0.65
2	0.38	0.12
3	0.20	0.05
4	0.12	0.02

(a) Based on the given ACF and PACF values, determine an appropriate ARIMA(p,d,q) model. Justify your answer. 3

(b) If the time series is already stationary, what would be the order of the ARIMA model? 1

6.

(a) Consider a time series with the following observed values:

Time (t)	**Observed Value (Y_t)**
1	50
2	55

You are given the following conditions:

Initial level estimate: $S_1=50$

Initial trend estimate: $b_1=3$

Smoothing parameters: $\alpha=0.4$, $\beta=0.3$

Using Holt's Double Exponential Smoothing, compute the level (S_2) and trend (b_2) for t = 2.

(b) A seasonal time series follows a multiplicative Holt-Winters model with:

Initial level estimate: $S_1=200$

Initial trend estimate: $b_1=10$

Initial seasonal index for season 1: $I_1=1.2$

Smoothing parameters: $\alpha=0.3$, $\beta=0.2$, $\gamma=0.4$.

Observed value at time t=2, $Y_2=250$

Using Triple Exponential Smoothing (Holt-Winters Method), compute the updated level (S_2).

2+3

7. A stationary time series has the following **autocovariance** values:

Lag	**$\gamma(k)$**
0	25
1	10
2	3

(a) Compute the spectral density (Periodogram) at frequency f=0.5 cycles per unit time using the Fourier Transform formula:

$$I(f)=\gamma(0) + 2\sum_{k=1}^{\infty}\gamma(k)\cos(2\pi f_k)$$

Use only given values and assume higher lags are negligible. 5

8. A time series has the following sample autocorrelation values:

Lag	**r_k**
1	0.05
2	-0.03
3	0.02
4	-0.01

Perform a white noise hypothesis test at $\alpha=0.05$ level using the approximate (Ljung-Box Test Approximation)

$$\text{Test Statistic: } Q=n\sum_{k=1}^{m}r_k^2$$

where n=20 (sample size) and m=4 (lags considered). (Given $\chi_4^2=9.49$ at 5% level of significance). Conclude if the series is white noise. 3

9. Practical Note Book and Viva-voce. 5

M.Sc. Examination, 2024
Semester-III
Statistics
Course: MSC-31
Stochastic Process

Time: 3 hrs Full Marks:40

Answer **all** questions.

1. Answer any **five** from the following. 5×2

   (a) Write down the state space and index set of a Brownian stochastic process.

   (b) For a discrete time Markov process with state space $S = \{0, 1\}$ with $p_{00} = 1$, $p_{01} = 0$, $p_{10} = 0.5$, $p_{11} = 0.5$. Does there exist any unique steady state probability? Justify your answer.

   (c) Consider a Markov chain with state space $\{0, 1, 2\}$ and transition matrix $\begin{pmatrix} \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$. Find $lim_{n\to\infty} p_{12}^{(n)}$.

   (d) Let $X(t)$ be a Poisson process with rate $\lambda = 2$. S_n is the waiting time till the n th event happens. Find the conditional expectation of waiting time of sixth event given that within first 5 minute three events have already occurred.

   (e) Customers arrive at a store with the Poisson flow having rate 10/hr. Each is either male or female with probability 1/2. Compute the probability that at least 5 men entered within 10 a.m. to 10.30 a.m.

   (f) Define an ergodic state.

2. Choose the most suitable word for the following multiple choice questions. 5×2

   (a) For a Wiener process

      i. measure of drift is double to measure of spread.
      ii. measure of drift is a function of time so is measure of spread.
      iii. measure of drift is function of time but measure of spread is function of time square.
      iv. measure of drift is equal to measure of spread.

   (b) Let π_j denote the long run proportion of time that the chain spends in state j where $\pi_j = 0$. Which of the following is false?

      i. No stationary distribution exists for j.
      ii. state j is null recurrent.
      iii. state j is positive recurrent.
      iv. If $i \longleftrightarrow j$ then i is null recurrent.

   (c) For a transition density matrix

      i. all off diagonal elements are positive
      ii. all diagonal elements are zero.
      iii. sum of elements of each row is greater than 0.
      iv. sum of the off diagonal elements is opposite in sign to the diagonal elements.

   (d) Let $W(t)$ be a standard Brownian motion. Then $P(W(1) + W(2) > 0)$ is

      i. 0
      ii. 0.5
      iii. 0.95
      iv. 1

(e) Two persons are catching fish independently with a Poisson flow at rate 2/hr. What is the expected amount of time that all of them will catch at least one fish?

i. 45 min
ii. 1 hr
iii. 30 min
iv. none of the above

3. Answer any **four** of the following. 4×5

(a) Deduce the difference equation on generalized birth and death process clearly stating the necessary assumptions.

(b) Show that if the intervals between successive occurrence of an event E are independently distributed with a common exponential distribution with mean $\frac{1}{\lambda}$, then flow of event E will follow a Poisson process.

(c) Prove that for a closed set, probability of moving out of the set is zero.

(d) Define renewal density. Show that renewal density is the derivative of renewal function.

(e) What do you mean by strict stationarity of a stochastic process? Show that strict stationarity does not imply covariance stationarity and vice versa.

M.Sc. Examination 2023
Semester III
Subject: Statistics
Paper: MSC-32 [Advanced Data Analysis Techniques]

Full Marks: 40 Time: 3 Hrs.

Answer any four questions (*Symbols have their usual meaning*)

1. a) Briefly discuss delta method for univariate case and its multivariate extension. Using this method or, otherwise approximate $Var(\frac{1}{\bar{X}})$ when $X_1, \cdots, X_n \overset{iid}{\sim} Poisson(\lambda)$.
   b) Why do we use odd's ratio? Find the asymptotic distribution of sample log-odd's ratio. (3+3)+(1+3)

2. a) What do you mean by prospective and retrospective study? Discuss advantages and disadvantages of both the studies.
   b) How does the Generalized Linear Model (GLM) differ from the Linear Model (LM), and in what situations would you choose to use a GLM over an LM when analyzing and modeling relationships in data? (1+1+4)+4

3. a) Let $f(y;\theta)$ be a exponential family with the following mathematical form: $f(y;\theta) = \exp\left[\frac{y\cdot\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right]$. Here, y is the observed value, θ is the natural parameter, ϕ is the dispersion parameter, $a(\phi)$ is the dispersion function, $b(\theta)$ is the cumulant generating function, and $c(y,\phi)$ is a normalization term. Show that, $V(Y) = a(\phi)b''(\theta)$. Hence find variance of the binomial distribution.
   b) Explain how to generate samples from $N(0, \sigma = 2)$ using the accept-reject method, with a $C(0,1)$ as the proposal distribution. (4+2)+4

4. a) Describe logistic regression model with some of it's application.
   b) Write down the detailed steps involved in fitting a logistic regression model. 3+7

5. a) Explain the basic rationale behind Gibbs sampling. Suppose $X \sim N(\theta, \sigma^2)$ with σ known, $\theta \sim C(\alpha, \beta^2)$ with α, β known. Then describe how Gibbs Sampling can be applied to simulate observations from the posterior of $\theta \mid X$.
   b) Describe the step-by-step process of using bootstrap in a linear regression model. (2+4)+4

6. a) Briefly write down the usage of EM algorithm with examples. Also discuss it's convergence.
   b) Write a short note on 'deviance' in the context of GLM. (3+3)+4

M.Sc. Examination, 2023
Semester-III
Statistics
Course: MSC-33(MSE-1)
(Operations Research and Optimization Techniques)
Time: Three Hours Full Marks: 40

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **any five** questions

1. What is a transportation problem? Is it considered to be a Linear Programming Problem? Show that a balanced transportation problem always has a feasible solution. 2+3+3

2. Briefly state the role of modeling in Operations Research. Mention different types of models and their solutions. 2+6

3. For the M/M/1 queuing system find the expected number of customers in the system in the steady state and also the expected queue length. Find the cumulative distribution function for the waiting time of a customer who has to wait in an M/M/1 queuing system. 4+4

4. What is two-person zero-sum game? Transform this game to a Linear Programming Problem. Prove that if mixed strategies be allowed, then there always exists a value of the game. 2+2+4

5. (a) Define an inventory. What are the advantages and disadvantages of having inventories?
(b) Suppose that Q^* is the optimal order quantity and K^* is the corresponding minimum annual variable cost.
Show that if a value of $Q = (1+\alpha)Q^*$ is used, $\frac{K}{K^*} = 1 + \frac{\alpha^2}{2(1+\alpha)}$, where K is the annual variable cost corresponding to an order quantity Q. 4+4

6. (a) What is replacement Problem? Give some illustrations. Discuss replacement policy of equipments that deteriorates gradually with change in time value of money.
(b) What is preventive replacement? Find out criterion for optimal replacement time in such situation. 4+4

7. (a) Distinguish between deterministic and probabilistic models of inventory.
(b) For an inventory model, if P(r) denotes the probability of requiring r units, where r is a discrete variable, C_1 is the inventory holding cost per unit of time, C_2 is the shortage cost per unit per unit of time, then show that the stock level which minimizes the total expected cost is that value of S which satisfies the conditions:

$$\sum_{r=0}^{S-1} P(r) < \frac{C_2}{C_1 + C_2} < \sum_{r=0}^{S} P(r).$$ 4+4

8. Write short notes on any two of the following: 4+4

(a) (s, S) inventory policy
(b) Duality problem in LPP
(c) Saddle point in game theory
(d) Congestion factor in Queuing model

M.Sc. Semester III Examination 2023
Subject: Statistics
Paper: MSC 34
(Time Series Analysis)

Full Marks: 40 Time: 3 hours

Answer any four of the following six questions of equal marks.
(Notations carry usual meanings)

1. (a) What is a time series? Explain the differences between time series data and cross-sectional data.
   (b) Describe with examples why time series analysis is important in real-world applications.

   5+5

2. (a) Define stationarity and invertibility of a time series.
   (b) Discuss the stationarity and invertibility of the following model:

   $$(1 - L)Y_t = (1 - 1.5L)\epsilon_t$$

   where ϵ_t follows $WN(0, \sigma^2)$ process.

   5+5

3. (a) Define auto-correlation function (ACF) of a time series. State and prove its properties.
   (b) Deduce the ACF of an MA(2) process.

   5+5

4. (a) What is a periodogram? Write down the general interpretations of a periodogram.
   (b) Define the White-noise process and justify its name with the help of frequency domain analysis.

   5+5

5. (a) Describes the steps involved in ARIMA modeling.
   (b) How do you select the appropriate orders (p, d, q) for an ARIMA model?

   6+4

6. (a) How does the exponential smoothing differ from the simple moving average method?
   (b) Compare different exponential smoothing methods for time series forecasting.

   3+7

M.Sc. Examination 2023
Semester III
Subject: Statistics (Practical)
Paper: MSC-35

Full Marks: 40 Time: 4 Hrs.

Answer all questions (*Symbols have their usual meaning*)

1. 200 boys and 100 girls were administered with intelligence test and their IQ's were determined in the following table. Write a program in R to find whether there is any evidence of sex-difference in intelligence. 5

Sex	**<80**	**80–120**	**>120**
Male	20	165	15
Female	12	77	11
Total	32	242	26

2. The table below presents the test-firing results for 21 surface-to air anti-air craft missiles at targets of varying speed. The result of each test is either a hit ($y = 1$) or a miss ($y = 0$). a) Fit a logistic regression model to the response variable y. Use a simple linear regression model as the structure for the linear predictor. b) Does the model deviance indicate that the logistic regression model from part a is adequate? c) Provide an interpretation of the parameter β_1 in this model. d) Expand the linear predictor to include a quadratic term in target speed. Is there any evidence that this quadratic term is required in the model? 10

Test	Target Speed (x)	y	Test	Target Speed (x)	y	Test	Target Speed (x)	y
1	400	0	8	470	0	15	280	1
2	220	1	9	480	0	16	210	1
3	490	0	10	310	1	17	300	1
4	210	1	11	240	1	18	470	1
5	500	0	12	490	0	19	230	0
6	270	0	13	430	0	20	430	0
7	200	1	14	330	1	21	460	0

3. Simulate n observations from the following mixture distribution $\alpha_1 N(\mu_1, \sigma_1^2) + \alpha_2 N(\mu_2, \sigma_2^2)$. Take the parameters of your choice and weights are such that i) Both are equal ii) First one is bigger than second one iii) First one is less than second one. Consider the statistics $T_n = \left[\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\right]^{1/\alpha}$, varying α as $-1.5, -1, -0.5, 0.5, 1, 1.5$. Find the bootstrap estimate of the standard error of the above estimate for all three case and draw the corresponding histogram. 10

4. Simulate a random sample from the density $f(x) = ce^{-x^2},\ x \geq 0$. Using a proposal density $g(x) = e^{-x},\ x \geq 0$. Also propose and justify an another $g(x)$ for simulation. 10

5. Practical note book and viva-voce. 5

M.Sc. Semester III Examination, 2023

Subject: Statistics
MSC-36(Practical)

Time: Four Hours | Full Marks: 40

Answer **all** questions. Candidates may use a computer for the Q. No. 5 and 6.

1. Maximize $z=5x_1+8x_2$

Such that $3x_1+2x_2 \geq 3$

$x_1+4x_2 \geq 10$

$x_1+x_2 \leq 5$

$x_1 \geq 0,\ x_2 \geq 0$

Solve the problem graphically. 4

2. Solve the transportation problem

	DI	DII	DIII	Supply
OI	4	3	2	10
OII	1	5	0	13
OIII	3	8	5	12
Demand	8	5	4	

6

3. A self service store employs one cashier at its counter. Nine customers arrive on an average of every five minute while the cashier can serve 10 customers in every five minute. Assuming Poisson distribution for arrival rate and exponential distribution for service rate, find (i) average number of customers in the system, (ii) average number of customers in the queue and (iii) average waiting time a customer spends in the system. 4

4. A fish vendor sells fish at the rate of Rs.50 per kg on the day of the catch. He pays Rs.2 per kg of fish not sold on the day of the catch for cold storage. Fish one day old is sold at the rate of Rs.30 per kg and there is unlimited demand for it. The demand of fresh fish is known to follow a uniform distribution over the range from 30 to 50.

(a) Determine the optimum quantity of fish that should be procured by the vendor.

(b) Calculate the maximum profit. Assume that the cost of procurement is Rs.35 per kg. 4

5. Consider the production data of a factory as follows:

203, 211, 219, 230, 238, 250.

a) Forecast the series values using exponential smoothing method by taking α=0.32. Plot the original and the forecasted series on the same graph, also write down the MSE.
b) You might have noticed in the results of a) that the forecasts are not that good. Improve your forecast and show the original and the forecasted values on the same graph (You free to assume any parameter value you may need for this). Also mention the MSE in this case.

5

6. Consider the following observations:

1124.18, 1099.27, 1475.66, 1927.19, 1507.88, 1333.15, 1758.01, 1714.98, 1499.96, 1584.80, 1223.13, 1005.16, 1467.27, 1058.59, 1252.02, 1778.27, 1804.83, 1985.80, 1521.48, 1549.19, 2657.69, 1820.96, 1762.64, 1138.78, 1677.46, 2062.71, 1912.33, 2611.93, 2518.59, 2292.15, 2053.17, 3016.70, 2705.78, 2039.64, 2358.18, 1504.98, 2272.55, 1940.75, 2347.31, 3082.62, 3556.77, 2919.21, 2634.30, 2858.82, 2758.46, 2580.01, 2444.16, 2286.00, 2713.02, 2442.28, 3331.12, 3422.31, 3630.87, 3539.44, 3397.08, 3716.67, 3495.05, 3140.39, 3048.92, 2595.11.

a) Read the data as a time series data. Also show its Time Series plot.
b) Comment on the stationarity of the Time Series with reasons.
c) Fit an appropriate ARIMA model to the data and comment on your quality of fitting with reasons.
d) Forecast 15 observations using the above ARIMA fitting.
e) Plot the original and the forecasted time series using along with the error bands.

2+2+5+1+2

7. Practical Note Book and Viva-voce. 5

****