M.Sc. Examination, 2019
Semester-I
Statistics
Course : MSC-11
(Linear Algebra and Linear Models)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin
Answer **any four** of the following questions

1. a) Define the basis of a vector space. Show that the number of vectors forming the basis of a vector space is unique.
   b) Let A be a symmetric matrix of dimension n with eigenvalues $\lambda_1, \lambda_{2,...,}\lambda_n$. Prove that $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 = \sum_{i=1}^{n} \lambda_i^2$. 6+4=10
2. a) Show that, if λ is the eigenvalue of a matrix A, then the eigenvalue of $(A+I)^{-1}$ is 1/(λ +1).
   b) Find the eigenvalues of the $n \times n$ matrix with all diagonal entries equal to a and all the remaining entries equal to b. 6+4=10
3. a) If the eigenvalues of a matrix A are positive, then show that the eigenvalues of the matrix $(A + A^{-1})$ are greater than or equal to 2.
   b) If A be a positive definite matrix, then show that there exists a non-singular matrix B such that $A = BB'$. 5+5=10
4. a) State and prove the necessary and sufficient condition for a linear parametric function $l'\beta$ to be estimable under the Gauss-Markov linear model set up.
   b) Consider the model
   $$y_1 = \beta_1 + \beta_2 + \epsilon_1$$
   $$y_2 = \beta_1 + \beta_3 + \epsilon_2$$
   $$y_3 = \beta_1 + \beta_2 + \epsilon_3$$
   Obtain a necessary and sufficient condition for $\lambda_1\beta_1 - \lambda_2\beta_2 + \lambda_3\beta_3$ estimable. 5+5=10
5. In case of multiple linear regression model, discuss the testing of the null hypothesis $H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$. Express the test statistic in terms of the multiple correlation coefficient. 10
6. a) With reference to the model $y = X\beta + e$, where $e \sim N_n(0, \Sigma)$ and $X^{n \times p}, (p < n)$ is of full column rank, describe the likelihood ratio test for $H_0 : L\beta = 0$ where the matrix L is of order $r \times p$ $(r < p)$ and rank $(L) = r$.
   b) With reference to pth degree polynomial regression model, obtain the test statistic for testing H_0: the pth coefficient is zero. 6+4=10

M.Sc. Examination, 2019
Semester-I
Statistics
Course : MSC-12
(Regression Analysis)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin
Answer **any four** of the following questions

1. Consider the simple linear regression model : $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1(1)n$, where $\varepsilon_i \sim NID(0, \sigma^2)$.
   a) Find maximum likelihood estimators of the parameters.
   b) How do you measure the quality of fit of the model?
   c) Write down the testing procedure for the slope to be equal to a constant. 4+3+3=10

2. a) Show that $I - H$ is symmetric and idempotent, where H is the hat matrix. Hence, or otherwise, find $tr(I-H)$.
   b) Write down the testing procedure for significance of regression of a multiple linear regression model.
   c) Write a short note on hidden extrapolation. 4+3+3=10

3. a) What is studentized residuals and when is it used? Show that studentized residuals (r_i) can be expressed as $r_i = \dfrac{e_i}{\sqrt{MS_{\text{Res}}\left[1-\left(\dfrac{1}{n}+\dfrac{(x_i-\bar{x})^2}{S_{xx}}\right)\right]}}, i = 1(1)n$ for simple linear regression.
   b) Describe the Box-Tidwell analytical method and its importance for transforming the regressor variables. 5+5=10

4. a) Write short notes on : Cook's Distance, DFBETAS.
   b) Show that DFFITS can be written as $DFFITS_i = \left(\dfrac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} t_i$, where t_i is R- Student residual.
   (2+3)+5=10

5. a) When do we use logistic regression model? Discuss by giving an example.
   b) Suppose we have a situation that the response variable is categorical with more than two outcomes. Then how do you model this type of phenomena?
   c) Write a short note on Poisson regression model. 3+2+5=10

6. a) What do you mean by multicollinearity and what are the reasons for multicollinearity?
   b) Discuss in detail how to deal with multicollinearity through Ridge Regression? (2+2)+6=10

M.Sc. Examination, 2019
Semester-I
Statistics
Course : MSC-13
(Stochastic Process and Distribution Theory)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin

Group - A

Answer **any two** of the following questions

1. a) Let $\{Y_n n \geq 1\}$ be a sequence of independent random variables with

$$P_r(Y_n = 1) = p = 1 - P_r(Y_n = -1)$$

Let X_n be defined by $X_0 = 0$, $X_{n+1} = X_n + Y_{n+1}$. Examine whether $\{X_n, n \geq 1\}$ is a Markov chain.

b) Find P^n and the limiting probability vector V for the chain having transition probability matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_1 & p_2 & p_3 \end{pmatrix}, \sum_{i=1}^{3} p_i = 1.$$ 5+5=10

2. a) For continuous time Markov Chain establish Feller Kolmogorov foreward and backward equations.

b) Show that for the Markov Chain with countable states S = {..., −2, −1, 0, 1, 2, ...} and transition probabilities $p_{i,i+1} = \frac{1}{2} = 1 - p_{i,i-1}, i = 0, +1, +2, \ldots$ all states are null persistent.

c) Define ergodic state. 4+4+2=10

3. a) "If the interarrival time variable sequence of a counting process with independent increment and stationarity follow exponential distribution, the process is a Poisson process" – Elucidate this comment with mathematical justification.

b) In a hardware store Bob must go to server one to get his goods, then go to server two to pay for them. Suppose the times of the two activities are exponentially distributed with means 6 minutes and 3 minutes respectively. Compute the average amount of time it takes Bob to get his goods and pay if when he enters the store, there is another one customer with server 1 and no one at server 2. 5+5=10

4. a) In a renewal process, prove the renewal equation

$$M(t) = E(t) + \int_0^t M(t-x)dF(x)$$

Where M(t) = E(N(t)); N(t) being number of renewals by time t and $F(\cdot)$ being the distribution of i.i.d. interasvival time between two menewals.

b) Let the interarrival time X between two renewal follow. A gamma distribution (2, a). Show that $M^*(s) = \frac{a^2}{s[(s+a)^2 - a^2]}$ where M*(s) is the Laplace transformation on M(t) = E(N(t)). Also find out the expected number of arrivals by time 10. 5+5=10

P.T.O.

Group – B

Answer **any two** of the following questions

5. Let $X_1, X_2, \ldots X_n$ be n independent normally distributed random variables with

$$E(Xi) = \mu_i$$
$$V(Xi) = \sigma_i^2$$

   a) Find the distribution of $\chi'^2 = \sum_{i=1}^{n} \frac{X_i^2}{\sigma_i^2}$

   b) Also find the characteristic function of χ'^2. 5+5=10

6. a) State Fisher-Cochran's theorem on quadratic forms.

   b) For $\underset{\sim}{y} \sim N_n(0, I_n)$ show that a necessary and sufficient condition that $\underline{Y}'A\underline{Y}$ has a chi-square distribution is that A is an idempotent matrix the degrees of freedom of χ'^2 being rank (A) = tr(A). 3+7=10

7. a) Show that if $A \sim Wp(n, \Sigma)$ and $A^{p\times p} = \begin{pmatrix} A_{11}^{q\times q} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ then $A_{11} \sim Wq(n, \Sigma_{11})$ where

$$\sum^{p\times p} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

   b) Show that if $A \sim Wp(n, \Sigma)$ and $L^{P\times 1}$ is a fixed vector of real elements, then

$$\underline{L}'A\underline{L} \sim (\underline{L}'\Sigma\underline{L})\chi_n^2$$ 5+5=10

8. a) State three fundamental theorems of least square theory.

   b) *Let* A_i $\underset{\sim}{\text{ind}}$ $Wp(n_i, \Sigma), i = 1, 2$. Then if $n_1 \geq p$, show

$$\frac{|A_1|}{|A_1 + A_2|} \sim \text{Beta}\left(\frac{n_1 - p + 1}{2}, \frac{n_2}{2}\right) \text{Beta}\left(\frac{n_1 - p + 2}{2}, \frac{n_2}{2}\right) \ldots \text{Beta}\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$$ 3+7=10

M.Sc. Examination, 2019
Semester-I
Statistics
Course : MSC-14
(Statistical Inference-1)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin

Answer **any four** of the following questions

1. a) State and prove Neyman-Fisher factorization theorem (discrete part only).

   b) Let $X_1, X_2, \ldots X_n$ be a random sample of size n from $N(\mu, 1)$ with unknown $\mu(-\infty < \mu < \infty)$. Find an unbiased estimator of μ^2 whose variance attains the Bhattacharya second lower bound. 4+6=10

2. a) What is completeness of a statistic? How is it related to MVU estimation? State any result in this connection.

   b) Let $(X_1, X_2, ..., X_n)$ be a random sample of size n drawn from Bernoulli population with parameter π. Find an UMVUE of $g(\pi) = 1 + n\pi + \frac{n(n-1)}{2}\pi^2$. 5+5=10

3. a) Describe Maximum likelihood method of parameter estimation. State its properties.

   b) Suppose $(X_1, X_2, \ldots, X_n)$ is a random sample of size n from a distribution with cdf

$$f(x \mid \alpha, \beta) = 0 \text{ if } x < 0$$
$$= \left(\frac{x}{\beta}\right)^{\alpha} \text{ if } 0 \le x \le \beta$$
$$= 1 \text{ if } x > \beta$$

   where the parameters α (>0) and β (>0) are unknown. Obtain the maximum likelihood estimators of α and β. 6+4=10

4. a) Consider the exponential family of distribution

$$P = \{f_\theta(x); \theta \in \Omega\},$$

   where $f_\theta(x) = k(\theta)e^{Q(\theta)t(x)}h(x)$

   Prove that if P is of the above form, then T=t(x) is complete sufficient.

   b) Let $X_1, X_2, \ldots X_n$ be a random sample of size n from $R(\theta_1, \theta_2)$. Is $T = (X_{(1)}, X_{(n)})$ sufficient for $\theta = (\theta_1, \theta_2)$? Is it complete? Give reason for your answer. 4+6=10

5. a) Define U-statistic. Find the limiting form of the variance of it.

   b) State the result regarding the asymptotic distribution of U-statistic with conditions, if any. Also find the asymptotic distribution of Wilcoxon's signed rank statistic. 6+4=10

6. a) Derive Bayes estimate of a real parametric function under $\gamma(\theta)$ under squared error loss.

   b) Suppose X_i; I = 1(1)n follows Bernoulli with parameter θ. The prior distribution of θ is Beta with parameters α and β. Find Bayes estimate of θ under squared error loss. 4+6=10

7. Write short notes on **any two** of the following : 5+5=10

   a) MVUE and Method of Covariance b) Minimum χ^2 method of estimation

   c) Kendal's τ and U-statistic d) Risk function and Bayes risk

M.Sc. Semester I Examination, 2019

Statistics (Practical)

Paper: MSC-15

Time: Four Hours | Full Marks: 40

(Computer Laboratory may be used, if necessary)

1. A random variable X takes values 0, 1, 2 with respective probabilities $\frac{\theta}{4N}+\frac{1}{2}\left(1-\frac{\theta}{N}\right)$, $\frac{\theta}{2N}+\frac{\alpha}{2}\left(1-\frac{\theta}{N}\right)$ and $\frac{\theta}{4N}+\frac{1-\alpha}{2}\left(1-\frac{\theta}{N}\right)$, where N=25 is a known number and α, θ are unknown parameters. If 75 independent observations on X yielded the values 27, 38, 10 respectively, estimate θ and α by method of moments. **7**

2. A random sample of size 20 is drawn from a population with the probability density function $f(x,\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}};x,\theta>0$ and the sample mean comes out to be 12.6. Find Maximum Likelihood Estimate of θ. How do you modify the estimate if 2 sample observations are known to exceed value 60 only and mean of the remaining observations is 6.8? What will be the estimate if it is known that observations beyond 60 are rejected? **7**

3. Suppose we have a data on response variable Y and regressor variables X_1, X_2, X_3, X_4 as follows: **15**

Index	Y	X_1	X_2	X_3	X_4	Index	Y	X_1	X_2	X_3	X_4
1	10	1985	38.9	59.7	2205	**15**	6	2140	38.8	59.2	1901
2	11	2855	38.8	55	2096	**16**	5	1730	36.6	54.4	2288
3	11	1737	40.1	65.6	1847	**17**	5	2072	35.3	49.6	2072
4	13	2905	41.6	61.4	1903	**18**	5	2929	41.1	54.3	2861
5	10	1666	39.2	66.1	1457	**19**	6	2268	38.2	58.7	2411
6	11	2927	39.7	61	1848	**20**	4	1983	39.3	51.7	2289
7	10	2341	38.1	66.1	1564	**21**	3	1792	39.7	61.9	2203
8	11	2737	37	58	1821	**22**	3	1606	39.7	52.7	2592
9	4	1414	42.1	57	2577	**23**	4	1492	35.5	57.8	2053
10	2	1838	42.3	58.9	2476	**24**	10	2835	35.3	59.7	1979
11	7	1480	37.3	67.5	1984	**25**	6	2416	38.7	54.9	2048
12	10	2191	39.5	57.2	1917	**26**	8	1638	39.9	65.3	1786
13	9	2229	37.4	58.8	1761	**27**	2	2649	37.4	43.8	2876
14	9	2204	35.1	58.6	1709	**28**	0	1503	39.3	53.5	2560

a) Fit a multiple linear Regression model between Y and X_1, X_3, X_4.
b) Using the partial F test, determine the contribution of X_3 on the model.
c) Find a 95% confidence interval on β_3.

d) Find a 95% confidence interval on mean number of Y when $X_1 = 2300$, $X_3 = 56$, and $X_4 = 2100$.

e) A 95% prediction interval on Y when $X_1 = 2300$, $X_3 = 56$, and $X_4 = 2100$.

f) What conclusion can you draw from this problem about the consequence of omitting an important regressor from the model?

g) Draw the RVH of X_1, X_3.

h) Fit the model $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon$ and test the general linear hypothesis $H_0 : \beta_2 - \beta_3 = 0, \beta_1 = 0, \beta_1 + \beta_2 + \beta_4 = 0$.

i) Find $(X_e' X_e)^{-1}$, where Xe is the unit length scaling effective matrix for the regressor X_1, X_2 and X_3.

j) Check for any possible multicollinearity of X_1, X_2 and X_3. You may use the result $VIF_j = \frac{1}{1-R_j^2}$, where R_j^2 is coefficient of determination regressing X_j on other regressor variables.

4. Show that the set of vectors are linearly independent {(1,1,1,0), (1,1,0,1), (1,0,1,1), (0,1,1,1)}. **2**

5. Let $A = \begin{bmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{bmatrix}$. Find the determinant of A and trace of A. **2**

6. Let $B = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$. Find the characteristic equation and all eigenvalues of B. **2**

7. Practical note book and Viva voce. **5**

****

M.Sc. Semester I Examination 2019
Subject: Statistics (Practical)
MSc.16

FM-40 Time 4hr

1. The number of patients arriving at a doctor's clinic can be modeled by a Poisson process with rate λ=0.8 patients per hour.
   a) Find the probability that exactly two patients arrive in each of the following time intervals viz.10 am to 12 pm and 12 pm to 2 pm.
   b) Given that 2 patients already arrived in (2pm ,4 pm),what is the probability that another 2 will arrive in (4pm,5pm)?
   c) Given that four patients already arrived in between (2pm, 5pm) what is the probability that exactly one arrived in (2pm, 3pm)?
   d) Find probability that 2 arrives in (1pm,4pm) and 3 events arrive in (3pm,5pm).
   e) How long the doctor has to wait in average to receive the fourth patient? 2+2+3+3+2
2. Given below is the transition matrix of a Markov Chain:

$$\begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

   Classify the states in terms of recurrent or transient. Justify your classification.
   Also for recurrent states find out mean recurrent time. 3+3
3. Esther and Bill enter a hair salon simultaneously, Esther to get a haircut and Bill to get a beard trimming. Suppose the time for hair cut(beard trimming) is exponentially distributed with mean 20(30 min).
   (i) What is the probability that Esther gets done first?
   (ii) What is the expected amount of time until both are done? 3+3
4. Consider the model $y_1 = \beta_1 + \beta_2 + e_1$

$$y_2 = \beta_1 + \beta_3 + e_2$$

   $y_3 = \beta_1 + \beta_2 + e_3$. $\Lambda_1\beta_1+\Lambda_2\beta_2+\Lambda_3\beta_3$ is estimable iff Λ_1= ?? 3
5. Below are given the yields in gm per plot for three varieties of seed cotton.

Variety 1	77	72	63	65	88	101
Variety 2	109	103	137	79	81	98
Variety3	46	73	72	84	71	47

   Using GM linear model, test if the varieties differ significantly among themselves. 8

6. Practical Notebook +Viva Voce 5

M.Sc. Examination, 2019
Semester-III
Statistics
Course : MSC-31
(Real Analysis and measure theory)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin
Answer **any four** of the following questions

1. a) Define σ - field. Give an example to show that all fields are not σ - field.
   b) Show that a σ - field is a monotone field. Is the converse true?
   c) Find *limsup* and *liminf* for the following sequence of sets 3+4+3=10
   $$A_n = \left\{(x, y) : 0 \le x < n, 0 \le y < \frac{1}{n}\right\}, n \ge 1$$
2. Show that any non-negative random variable can be expressed as a limit of a sequence of non-negative, non-decreasing simple random variables. 10
3. a) Using monotone convergence theorem, State and prove Fatou's lemma. Mention the strict inequality case in this regard.
   b) State and prove dominated convergence theorem. 6+4=10
4. State and prove the continuity theorem of measure. 10
5. a) Show that there does not exist a rational number r such that $r^2 = 2$.
   b) State and prove the Archimedean property of R. Hence or otherwise show that if $x \in R$, and $x > 0$, then there exist a natural number n such that $0 < \frac{1}{n} < x$. 4+6=10
6. a) Define $f : R \to R$ by
   $$f(x) = \frac{x^3}{1+x^2}$$
   Show that f is continuous on R. Is f uniformly continuous on R?
   b) Does there exist a differentiable function $f : R \to R$ such that for all $f'(0) = 0$ but $f'(x) \ge 1$ for all $x \ne 0$? 5+5=10
7. Define $f_n : [0, \infty) \to R$ by
   $$f_n(x) = \frac{\sin(nx)}{1+nx}$$
   a) Show that f_n converges pointwise on $[0, \infty)$ and find the pointwise *limit* f.
   b) Show that $f_n \to f$ uniformly on $[a, \infty)$ for every $a > 0$.
   c) Show that f_n does not converge uniformly to f on $[a, \infty)$. 4+3+3=10

M.Sc. Examination, 2019
Semester-III
Statistics
Course : MSC-32
(Categorical Data Analysis and Advanced Data Analysis Technique)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin
Answer **any four** of the following questions

1. a) Explain why odd's ratio is a suitable measure for both prospective and retrospective studies but relative risk is not so. 3

   b) Let a denote the number of calves that got a primary (t = 1) secondary (t = 2) and tertiary (t = 3) infection, b the number that received a primary and secondary infection but not a tertiary one, c the number that received a primary infection but not a secondary one, and d the number that did not receive a primary infection. Let π be the probability of a primary infection. Consider the hypothesis that the probability of infection at time t, given infection at time 1,2… t-1 is also π, for t = 2,3. Show that the ML estimate of π under the hypothesis is

   $$\hat{\pi} = \frac{3a+2b+c}{3a+3b+2c+d}$$ 4

   c) Find the deviance of a Poisson model. 3

2. a) Define Pseudo – R^2 measure for GLM and find its form for a binary response model. 3

   b) For sample proportions (p_i) and ML estimates $\{\hat{\pi}_i\}$, prove that the likelihood ratio statistic

   $G^2 = -2n\sum_i p_i \ \ln\left(\frac{\hat{\pi}_i}{p_i}\right)$ is non-negative and $G^2 = 0$ iff $p_i = \pi_i \forall i$. 3

   c) Define the power divergence statistic as

   $$T = \frac{2}{\lambda(\lambda+1)}\sum_i n_i\left[\left(\frac{n_i}{\hat{m}_i}\right)^{\lambda} - 1\right], \text{ for } -\infty < \lambda < \infty$$

   Prove that as $\lambda \to 0, T \to G^2$. 4

3. a) Define mutual, joint and conditional independence with respect to a three-way contingency table. Prove that mutual independence implies joint and conditional independence. 6

   b) Define a three–dimensional hierarchical log-linear model and interpret the model parameters. 4

4. a) Mathematically explain the E-step and M-step of EM algorithm. 6

   b) Explain how it can be applied in the estimation of the model parameters based on censored data from a normal population. 4

5. a) Explain Gibbs sampling technique. 4

   b) Suppose X denotes the number of defectives in the daily production of the product (Y) (Y is unobservable).

   We consider the model

   X|Y ~ Bin (Y, θ)

   Y ~ Poisson (λ)

   θ ~ Beta (a,b)

P.T.O.

Explain how Gibbs sampling can be implemented to simulate from the posterior distribution of $\theta \mid X$ 6

6. a) Let X_i ($i = 1(1)n$) be i.i.d with variance σ^2. Find the bootstrap and jackknife estimate of standard error of the statistic $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ for estimating σ^2 and comment. 6

   b) Write a short note on cross-validation technique. 4

7. a) Compare the importance sampling technique with accept-reject method. 4

   b) Explain how you can minimize the variance of the estimator of $E_f[h(x)]$ using importance sampling. 4

   c) What do you mean by Metropolis-Hastings algorithm? 2

M.Sc. Examination, 2019
Semester-III
Statistics
Course : MSC-33 : MSE-1
(Operations Research and Optimization Technique)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin

Answer **any five** of the following 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=8
2. 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=8
3. Briefly state the role of modeling in Operations Research. Mention different types of models and their solutions. 2+6=8
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=8
5. 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=8
6. 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=8
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 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
8. Write short notes on **any two** of the following : 4+4=8
   a) (s, S) inventory policy
   b) Duality problem in LPP
   c) Saddle point in game theory
   d) Congestion factor in Queuing model

M.Sc. Examination, 2019
Semester-III
Statistics
Course : MSC-34 : MSS-3
(Time Series Analysis)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin
Answer **any four** of the following questions

1. a) What are the two mathematical models employed for time series analysis? Can one model be considered as a particular type of the other one? Which one of the two models is considered to be more useful? 5

   b) Write down the names of different components of a time series and give at least one example of each component. 5

2. a) Define a stationary time series. 4

   b) Show that the series

$$S_t = X_1 + X_2 + \ldots + X_t,$$

   where Xi's are iid $(0, \sigma^2)$ variables, is not a stationary process. 3

   c) Let $\{X_i\}$ and $\{Y_i\}$ are independent stationary processes. Then comment on the stationarity of the process $\{aX_i+bY_i\}$ where a and b are real numbers. 3

3. a) State the properties of the auto-correlation function. 2

   b) Find the auto-correlation functions of Lag 1 and 2 for the following process.

$$Y_t = \frac{5}{6}Y_{t-1} - \frac{1}{6}Y_{t-2} + \varepsilon_t$$

   where $\varepsilon_t \sim WN(0,\sigma^2)$. 5

   c) Comment on the following series with regard to stationarity.

$$Y_t = 2.5Y_{t-1} - Y_{t-2} + \varepsilon_t$$

   where $\varepsilon_t \sim WN(0,\sigma^2)$. 3

4. a) What do you understand by invertibility of a time series? Why do we need to consider it? 3

   b) You believe that a set of data is the realization of a MA(1) process

$$Y_t = \beta\varepsilon_{t-1} + \varepsilon_t$$

   where ε_t are standard normal variates. You have calculated the sample auto-convariance function and found that $\hat{\gamma}_0 = 1$ and $\hat{\gamma}_1 = -0.25$ Estimate the parameter β.

   Which value of β do you think should be chosen and why? 7

5. Describe the different steps of fitting Box-Jenkin's ARIMA model for a time series. Explain different types of transformations used for the fitting of ARIMA model and the criterion for identifying a model as well. 10

6. a) Define power spectrum of a stationary process. Explain why do we consider it for studying time series. 4

   b) Derive the spectral density functions of the white Noise and AR(1) processes. 6

Visva Bharati University
M.Sc. Semester III Examination 2019
Subject: Statistics (Practical)
Paper: MSC-35

Students are asked to provide the printouts of the codes and outputs of the programs, clearly mentioning their roll numbers. The results should be unambiguously interpreted and the underlying theory/algorithm must be properly stated.

Full Marks: 40

Time: 4 Hrs.

1. **Instructions:**

- Please find an MS-Excel file named "$msc-35-2019-problem1.xlsx$" at the location "desktop" in your computer.
- It contains the natural tip length of 12 different fishes taken from a pond, over 12 time points, together with the amount of dissolved oxygen present in the pond over the time points.
- Denote the length of the ith fish at time point t by $Y_{it}, i = 1, 2, \ldots 12, t = 1, 2, \ldots 12$.
- Denote the amount of dissolved oxygen present at time t by Z_t.
- Define the relative growth rate of fish i at time point t as $R_{it} = \ln\left(\frac{Y_{i,t+1}}{Y_{it}}\right), i = 1, 2, \ldots 12, t = 1, 2, \ldots 11$.

(a) Plot the values of R_t against Z_t values. Fit an appropriate curve $R_t = \phi(Z_t) + \epsilon_t$, based on the data points. Draw the fitted curve over the observed data points.

(b) Find the bootstrap estimate of the standard error of the model parameters. Take the number of bootstrap replications (B) to be 50.

(c) Draw the histograms of the bootstrap replications of each of the parameters. Also check graphically whether they can be regraded to come from a Normal distribution.

4+4+4=12

2. Draw a random sample of suitable size from the Beta (4, 6) distribution using acceptance-rejection sampling (Take your proposal density to be $U(0, 1)$). Construct the histogram of the sample and fit a Beta distribution using the same sample. Draw the actual density and the fitted density over the histogram (on the same graph).
6

3. The following are the marks secured by two batches of salesmen in the final test after completion of training.

   Batch A: 26, 27, 31, 26, 19, 21, 20, 25, 30
   Batch B: 23, 28, 26, 24, 22, 19.

   Use permutation test to decide whether there is any significant difference between the mean marks of these two batches.
4

4. Suppose the random variable X follows Normal distribution with mean θ and variance 1. We assume the prior density to be Cauchy with location parameter 5 and scale parameter 1. Using Gibbs sampling technique, simulate 50 observations from the posterior distribution of θ given X. Also find the Bayes estimate of the parameter θ under squared error loss and absolute error loss.
8

5. **Instructions:**

- Please find an MS-Excel file named "$msc-35-2019-problem5.xlsx$" at the location "desktop" in your computer.
- It contains the data on 112 patients. The first column indicates whether he/she is affected by the disease DPLD or not.
- Columns 2-7 provide the corresponding values of 6 associated covariates, viz. $FEF_{25-75}(pre)$, $FEF_{25-75}(post)$, $FEV_1(pre)$, $FEV_1(post)$, $FVC(pre)$, $FVC(post)$.

Fit a binary logit model with the covariates $FEF_{25-75}(post)-FEF_{25-75}(pre)$, $FEV_1(post)-FEV_1(pre)$ and $FVC(post)-FVC(pre)$. Comment on the goodness of your fit. 5

6. Practical Note Book and Viva-Voce 5

x——— Best of Luck———x

M.Sc Semester III Examination, 2019

Statistics

MSC-36(Practical)

Time: Four Hours Full Marks: 40

Answer all the questions. Candidates can use R code for answering the questions whenever necessary. Necessary code must be put in the answer-script.

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	

8

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. 3

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. 3

5. Simulate and plot 100 observations from the AR(2) process with $\varphi_1= 0.7$ and $\varphi_2= -0.7$. Also plot the autocorrelation and partial autocorrelation functions up to five lags. 6

6. Take the data available on the computer saved as MSC36.csv on the desktop. Plot the Time series and fit an appropriate time series model. Forecast the series values for 24 months. Plot the original and the forecasted values. Also give an indication of forecast errors. 11

7. Practical notebook and Viva voce. 5

****