Instructions.—(1) The candidate may attempt *any five* questions selecting not more than two from any section. In case the candidate answers more than five questions, only the first five questions in the chronological order of question numbers answered will be evaluated and the rest of the answers ignored.

(2) Each question carries 20 marks.

(3) Answers must be written in **English** or in **Hindi**.

(4) QUESTIONS FROM EACH SECTION SHOULD BE ANSWERED ON SEPARATE ANSWER-BOOK/SUPPLEMENTS.

(5) Answer to each question must begin on a fresh page and the question number must be written on the top.

(6) On the answer-book, Name, Roll Number etc. are to be written in the space provided for them. Name or Roll Number should not be written on the supplement.

(7) Candidates should use their own pen, pencil, eraser and pencil-sharpener and footrule.

(8) No reference books, Text books, Mathematical tables, Engineering tables or other instruments will be supplied or allowed to be used or even allowed to be kept with the candidates. Violation of this rule may lead to penalties. use of nonprogrammable electronic calculator is permitted.

(9) ALL ROUGH WORK MUST BE DONE IN THE LAST THREE OR FOUR PAGES OF THE ANSWER BOOKLET; ADDITIONAL BOOKLETS WILL BE PROVIDED ON DEMAND, WHICH SHOULD BE ATTACHED TO THE ANSWER BOOKLET BEFORE RETURNING.

A : Probability and Sampling

1. Describe stratified random sampling.

Suppose we want to estimate the population proportion P of units in the population having acharacter A; by a stratified random sampling. Let P_n denote population proportion of units having A in stratum h and P_h denote sample proportion of units with A in stratum h.

(*i*) Obtain unbaised estimator of P. Also obtain its variance under proportinal allocation.

(*ii*) Assuming that the estimated population proportion should not differ from the true population proportion by more than 10% with a probability $(1-x)$, calculate the sample size for proportional allocation.

(*b*) The following table gives the P_h and N_h values for different strata Assuming that the estimated population proporation should not differ from the true population by more than 10% with a proba bility of .95, calcukate the sample size for proportional allocation

h	P_h	N_h
1	0.3	100
2	0.2	200
3	0.4	250
4	0.2	50

2. (*a*) State and prove Holder's inequality.

(*b*) Let X_i i = 1, 2, ... 10 be independent random variables each being uniformly distributed over (0, 1) calculate $P\{\sum_{i=1}^{10} X_i \geq \rangle 7\}$

3. (*a*) Describe cumularive total method of drawing a ppswr sample.

In ppswr sampling, give an unbised estimator of population total. Also obtain its variance.

(*b*) A population consists of 5 units. The values of response variable Y*i*, size of unit X*i* are given below :

Unit No.	1	2	3	4	5
Size X*i*	5	10	15	12	8
Y*i*	32	41	25	30	35

Draw a ppswr sample of size 2 using random numbers 0.40 and 0.26 from U(0,1). estimate the population total based on yair sample. Is yaer estimate unbiased ?

B : Linear Models and Economics Statistics

4. (*a*) Define the follows and state the limitations of each :—

(*i*) Simple aggregative price index.

(*ii*) Index of price relatives.

(*iii*) Weighted aggregative price index.

(*iv*) Laspeyre's price index.

(*v*) Paasche's price index.

What is Fisher's ideal index number. Show tha it satisfies time reversal test and factor reversal test.

(*b*) Compute Laspeyre's, Paasche's and Finsher's price index numbers for the following data using 1981 as base period. The table gives prices in Rs. per Tonne, quantities in million tonnes. Figures in bracket are quantities.

Commodities	1981	1985
Wheat	554 (9.67)	673 (10.77)
Rice	427 (31.95)	622 (36.32)

5. (*a*) Consider the two-way classification model—

$y_{ij} = j\mu + x_i + \beta_j + r_{ij} + e_{ijk}$

$i = 1, 2, ...p,$ $\quad j = 1, 2, ... q \quad k = 1, 2, ... r$

Explain the terms involved in the model. State the assumptions required and derive test for the hypotheses

$H : r_{jj} = 0$ ...... for $i = 1, 2, ...p,$ $\quad j = 1, 2, ... q$

(*b*) A menu factuter wanted to study production rates for different combinations of reagents and catalists. The data given in the following table shows coaded valuesof produciton rates. Obtain ANOVA and write your conclusions. You are given reagents X catalysts interaction S. S. = 84.

	Catalyst		
Regent	1	2	3
A	4, 6	11, 7	5, 9
B	6, 4	13, 15	9, 7
C	13, 15	15, 9	13, 13
D	12, 12	12, 14	7, 9

6. (*a*) Write down multiple regression model stating all the assumptions, Describe various tests of hypotheses associated with such a model. Indiate method of testing in each case.

(*b*) Consider a linear regressiun model with six regressors having cuefficients $\beta_1, ..., \beta_6$ and intercept β_0 Data is available on 30 observations and it is given that regression S. S. = 3147.97, residual S. S. = 1149.00.

(*i*) Test the hypothesis : $\beta_1 = \beta_2 = . = \beta_6 = 0$ (F at 1 % level of significance with appropriate d.f. = 3.71).

(*ii*) Find the value of multiple correlation coefficient.

C : Statistical Inference

7. (*a*) Prove that empirical distribution function is an unbiesed estimator of F (x), df of X.

(*b*) Let $x_1, x_2, ... x_n$ be iid vvs from the following pdf.

$$f(x, \theta) = \begin{cases} Q(\theta)\, M(x) & ; a \text{ c } x \text{ c } \theta \\ O & ; \text{otherwise} \end{cases}$$

where M (*x*) is non- negative and obsolutaly continuous over (a, j) is diffeentiable everywhere. Find UMVUE of g (θ), where g (θ) is also differentiable every where.

8. (*a*) Let the VV X has the pdf f (x, θ), $\theta \leftarrow (4)$. Assume that $f(x, \theta)$ belomgs to one parameter exponential family then prove that the test Φ if U is continuous.

$$\varphi(x) = \begin{cases} 1 & \text{if } u < C_1 \;\; w \;\; u > c_2 \\ 0 & \text{otherwise} \end{cases}$$

is UMPU of size x for testing

H_U ! $\theta = \theta_U$ against $H_1 : \theta = \theta_0$, where $U = \sum_{1=1}^{n} T(4)$

(*b*) Let *X* be a random variable with pdf (x, θ), $\theta \leftarrow (4)$,

$$f(x, \theta) = \begin{cases} \dfrac{2(\theta - x)}{\theta^2} & ; 0 < x_1 < \theta \\ 0 & ; \text{otherwise} \end{cases}$$

Obtain a MP test of sine *x* to test

(*i*) $H_o : \theta = \theta_o$ against $H_1 : \theta = \theta_1 > \theta_o$

(*ii*) $H_o : \theta = \theta_o$ against $H_1 : \theta = \theta_1 < \theta_o$

9. (*a*) Lef $x_1, x_2, \ldots\ldots x_4$ be a random sample of sine *n* from a normal distribution with with mean *u* and vanance o^2. Obtain the likelined ration test for testing Ho : $\mu \leq \mu_o$ against $H_1 : \mu > \mu_o$, where σ^2 is unknown.

(*b*) In a singing competition, the judges acreed that 7 exhibits were outstanding and these were numbered from 1 to 7. Three judges have given the following raankings :—

Judge A	7	2	4	3	1	5	6
Judge B	1	3	5	2	4	7	6
Judge C	4	1	2	3	5	6	7

Compute Kandall's sample tau coefficient T from the three possible pairs of rankings.

D : Stochastic Processes

10. (*a*) Consider a time homogeneous Markov Chain x_n with finite state space and the transition mature P, obtain the probablity distribution of x_r, $x_r + 1$, ... , $x_r + n$ in terms of transition probablities and initial distribution of x_r, Also derive chapman — kolmogorov equation for computing *n* step transition probablities

(*b*) A company asesses credit worlhiners of various firmsevery quanter, the ratings are in order of decreasing merit A, B, C and D (default). Historical data support the view that the credit rating of a tyrical firm evoles as a Markov Chain with following transition probability matrix—

$$P = \begin{bmatrix} 1-a-a^2 & a & a^2 & 0 \\ a & 1-2a-a^2 & a & a^2 \\ a^2 & a & 1-2a-a^2 & a \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

for some panameter *a*.

(*i*) Determine the range of values of a for which matrix P is valia transition matrix.

(*ii*) Whether chain is inducible and apeniodic ?

(*iii*) Derive the station any probability distrubution of the chain.

(*iv*) For a = 0.2 calculate the probabilities

$P[x_3 = D/x_1 = A]$, $P[x_3 = D/x_1 = B]$, $P[x_3 = D/x_1 = C]$,

11. (*a*) Consider time homogeneous Markor Chain x_n with state place S = {1, 2, 3}, and transition matrix.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

(*i*) Given the startig value and transition matrix p explain the method of simulating path of the Markor Chain.

(*ii*) For the simulated path obtained in (*i*) above give the estimate of transition matrix *p* of the Markor Chain.

(*b*) A gambler begins with Rs. 500. Each game he may win Rs. 100 with probability 0.3 and lose with probability 0.7. He willplay until he doubles his money or losses it all use the simulation method to determine the path of the game using random numbers given below :

0.77, 0. 75, 0.14, 0.26, 0.20, 0.51, 0.72, 0.76, 0.44, 0.20, 0.67, 0.84, 0.27, 0.22, 0.07, 0.89, 0.18, 0.69, 0.10, 0.04. using the path obtained calculated the estimate of transition matrix p.

12. (*a*) A time series model is specified bym $yt = x_1\ yt-1 + x2\ yt-2$ + et where et is a wite noise process with variance σ^2,

(*i*) Determine whether the process is stationary.

(*ii*) Obtain yule walker equations.

(*iii*) Hence obtain autocorrelation of order one and two also obtain partial autocorrelations.

(*iv*) Give procedure to estimate *x*, and *x*2.

(*b*) The time series Y_t is assumed to be stationary and to follow an ARMA (2,1) process defined by

$$yt = 1 + \frac{8}{15} yt-1 - \frac{1}{15} yt - 2 + et - \frac{1}{7} et - 1$$

where et are independent N (0,1) random variables.

(*i*) Determine the roots of the characteristic polynominal and explain how their values relate to the stationanity of the process. Find autocorrelation for lags 0, 1, 2.

E : Multivariate Analysis

13. (*a*) Define population principle components show that

$V(y_p) < V(y_{p-1}) < \ldots\ldots < V(y_1)$

where yi denotes the i^{th} principle component of p-variates population.

(*b*) Obtain th first principal component y_1 of the following corrolation matrix ? also find its variance and proportion of variation explained by y_1

$$\delta = \begin{pmatrix} 1.000.63 \\ 0.631.00 \end{pmatrix}$$

14. (*a*) What is cluster analysis ? What is distance and similarity cofficient for a pair of item. Give one example of eachof them.

(*b*) The vocabulary "richness" of a text can be quantitatively described by counting the words used once, the words used twice and so forth. Based on these coutns, a linguist proposed the following distances between chapters of the Old Testament book Lamentations.

Lamentations chapter

		1	2	3	4	5
	1	0				
Lamentations	2	0.76	0			
chapter	3	2.97	0.98	0		
	4	4.88	4.17	0.21	0	
	5	3.86	1.92	1.51	0.51	0

Cluster the chapters of lamentations using the single linkage hierarchical mehod. Draw the dendrogram.

15. (*a*) Define Mahalanobis distance Δ^2, a measure of the distance between the two normal populations and its estimate D^2; based on two random samples of sizes n_1 and n_2 from the two multivariate normal population $\pi 1$ or .

(*b*) Suppose the $n_1 = 11$ and $n_2 = 12$ observations are made on two random vectors X_1 and X_2 which are assumed to have bivariate normal distribution with a common covariance matrix Σ, but possibly different mean vectots M_1 and M_2. The sample mean vectors and pooled covariance matrix are

$$x1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}', x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, S\,pooled = \begin{bmatrix} 7.3 & -1.1 \\ -1.1 & 4.8 \end{bmatrix}$$

obtained the mahalabis sample distance D^2 and obtain the linear discriminant function. Assign the observation $X_0^1 = (0 \quad 1)$ to either population π_1 or π_2. Assume equal cost and equal prior probabilities.

F : Numerical Analysis and Basic Computer Techiques.

16. (a) If $X^{(k)}$ is the approximation to the solution X of the system of linear equations AX=b then show that the next approximation $X^{(k+1)}$ of X using Jacobi's iteration method is $X^{(k+1)} = HX^{(k)} + c$, where $H = - D^{-1}(L+U)$ and $c = D^{-1}b$. k=0, 1, …,n.

Using this method obtain $X^{(1)}$ for the following system of equations

$$2x_1 + x_2 = 1$$
$$2x_1 + 3x_2 = -1$$

Take initial approximation as $X^{(0)} = [2 \ \ 2]'$

(b) Derive Trapezoidal rule from Newton Quadrature formula given by,

$$\int_a^b f(x)dx = \sum_{k=0}^{n} \lambda_k f(x_k)$$

and deduce the composite Trapezoidal rule in the form

$$\int_a^b f(x)dx = \frac{h}{2}[f(x_0) + 2(f(x_1) + f(x_2) + ... + f(x_{n-1})) + f(x_n)]$$

Evaluate the following integral using above rule. Use interval difference as 0.5.

$$\int_1^3 \frac{\cos(x)}{1+e^x} x^2 dx$$

17. (a) (i) Define forward difference operator Δ and prove that

$$\Delta^j f_0 = \sum_{k=0}^{j} (-1)^k \binom{j}{k} f_{j-k} \text{ where } j \in IN$$

(ii) Define shift operator E and central difference operator δ. Show that

$$e^x = \left(\frac{\Delta^2}{E}\right) e^x . \frac{Ee^x}{\Delta^2 e^x}, \text{ where interval of difference is one.}$$

(b) Fit the following set of data with first order (linear) splines. Evaluate the function at x =5. Also plot the spline functions.

x	3	4.5	7	9
f(x)	2.5	1	2.5	0.5

18. (a) Answer the following questions:
(i) Explain OSI reference model in networking.
(ii) Write a short note on Ring topology and Mesh topology.

(b) Answer the following questions:
(i) Explain the following terms related to Relational Database Management System with example: Primary key, Candidate key, and Foreign key.
(ii) Explain Bubble sort technique to sort the following numbers:

44, 34, 23, 22, 05