INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-Stream): 1988-89
Introduction to Stochastic Processes
Semestral-II Examination

Date: 16.5.89 Maximum Marks: 100 Time: 4 Hours

Note: Answer at most six questions.
The maximum you can score is 100.

1.(i) Taxis arrive at a stand according to a Poisson process with rate λ_1. Customers arrive at the stand according to another independent Poisson process with rate λ_2. If a taxi arrives and individuals are waiting, the first person in line is served; if no individuals is waiting, the taxi waits. If a person arrives and there are taxis waiting, the person requisitions the first taxi; if no taxi is available, the person waits.

Let X_t be the number of taxis waiting at time t if taxis are waiting for people, and be minus the number of people waiting at time t if people are waiting for taxis.

(a) Show that $\{X_t : t \geq 0\}$ is a compound Poisson process so that it can be represented as $X_t = \sum_{i=1}^{N_t} Y_i$.

(b) Find the parameter of $\{N_t : t \geq 0\}$.

(c) Find the distribution of Y_1.

(d) Write $p_n(t) = P(X_t=n)$; $n=0, \pm1, \pm2, \ldots$. Compute

$\sum_{n=-\infty}^{\infty} p_n(t)s^n$ for $|s| \neq 0$. [6+2+3+3=14]

(ii) Give an example of a stochastic process which is weakly stationary but not stationary. [6]

2.(i) (a) Define a stationary renewal process.

(b) Suppose $\{N_t : t \geq 0\}$ is a stationary renewal process with associated interoccurrence times denoted by $\{X_n : n \geq 1\}$. Write $S_n=X_1+\ldots+X_n, n \geq 1$.

Fix any number $s > 0$, Define $\{Y_n : n \geq 1\}$ by

$$Y_n = S_{N_s+1}-s \quad , \quad n=1$$

$$= S_{N_s+n}-S_{N_s+n-1} \quad , \quad n \geq 2.$$

Show that $\{Y_n : n \geq 1\} \overset{d}{=} \{X_n : n \geq 1\}$.

p.t.o.

2.(i) (c) Hence, or otherwise, show that a stationary renewal process has stationary increments. [3+7+4=14]

(ii) Find the renewal function corresponding to the lifetime density

$$f(x) = \lambda^2 x e^{-\lambda x}, \; x \geq 0.$$ [6]

3.(i) Suppose $\{X_n : n \geq 0\}$ is a branching process with $X_0 \equiv 1$ and associated probability generating function ψ. Suppose

$\psi(s) = \sum_{k=0}^{\infty} p_k s^k$ where every $p_k < 1$ and $p_0 + p_1 < 1$. Suppose moreover that $1 < m \stackrel{\text{def.}}{=} \sum_{k=1}^{\infty} k p_k < \infty$.

(a) State and prove a result describing the limiting behaviour of $\{X_n : n \geq 0\}$.

(b) State a result describing the limiting behaviour of

$\{W_n : n \geq 0\}$ where $W_n \stackrel{\text{def}}{=} \frac{X_n}{m^n}, \; n \geq 0.$ [10+3=13]

(ii) Solve the following equation:

$$e^x - xe = 0.$$ [7]

4.(i) Suppose $A_0(t)$ solves the renewal equation

$$A(t) = a(t) + \int_0^t A(t-y)\,dF(y)$$

where $a(t)$ is a bounded nondecreasing function with $a(0)=0$.

Show that $\lim_{t\to\infty} \frac{A_0(t)}{t} = \frac{a^*}{\mu}$, where $a^* = \lim_{t\to\infty} a(t)$ and $0 < \mu < \infty$ is the mean of $F(x)$. [9]

(ii) Suppose that in a branching process $\{X_n : n \geq 0\}$ with $X_0 \equiv 1$, the number of offspring of the initial particle has a distribution whose generating function is $f(s)$. Each member of the first generation has a number of offspring whose distribution has generating function $g(s)$. The next generation has generating function f, the next g and the functions continue to alternate in this way from generation to generation. Denote by $\varphi_n(s)$, the probability generating function of X_n.

(a) Obtain recurrence relationships between φ_n and φ_{n-2}, separately for n even and n odd.

contd. 3/-

4.(ii) (b) Hence, or otherwise, determine extinction probability of the process.

(c) Would the quantity obtained in (b) above change if we started the process with the g function, and then continued to alternate ? [3+4+4=11]

5.(a) Let i and j be two states in a Markov chain $\{X_n : n=0,1,2,\ldots\}$; assume that j is aperiodic and recurrent and it is accessible from i. State and prove a theorem concerning the limiting behaviour of $P^n_{ij} = P(X_n=j \mid X_0=i)$ as $n \to \infty$.

(b) Consider a Markov chain having only three states 0, 1, 2. Its transition probability matrix is given by

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}.$$

Show that for $n = 1,2,\ldots,$ $P^{2n} = P^2$ and $P^{2n+1} = P$ and hence comment on the recurrence or transience of the three states.

(c) Let S be the state space of a Markov chain $\{X_n : n=0,1,\ldots\}$ with $P_{ij}=P(X_n=j \mid X_{n-1}=i)$, $n \geq 1$. Write $S = S \cup T$ where R and T denote the set of recurrent and transient states in S respectively. Assume R and T to be both non-empty and also that the probability of staying forever in transient states is zero.

Prove that

$$\lambda_j = \sum_{k \in T} P_{jk} \lambda_k + 1,$$

where $\lambda_i = E(\tau_R \mid X_0 = i)$, τ_R being the hitting time of R.

[9+5+6=20]

6. Consider a Markov chain $\{X_n : n=0,1,\ldots\}$ having a finite state space $S = \{0,1,\ldots,d\}$, $d > 1$. The transition probabilities $P_{ij} = P(X_n=j \mid X_{n-1} = i)$, $n \geq 1$ are such that for $i = 0,1,2,\ldots d$,

$$\sum_{j=0}^{d} j P_{ij} = i. \qquad (*)$$

contd. 4/-

6.(a) Show that for $i = 1,2,\ldots, d$

$$i\,P_{io} + (i-1)P_{i1} + (i-2)P_{i2} + \cdots + P_{ii-1}$$

$$= P_{ii+1} + 2P_{ii+2} + \cdots + (d-i)P_{id},$$

where $P_{ij} = 0$ for $j > d$.

(b) From (*), prove that both 0 and d are absorbing states.

(c) Assume that none of the states $1,2,\ldots, d-1$ is absorbing.

(i) Using (a) above or otherwise, show that $P_{10} > 0$. Extend your arguments to establish that 0 is accessible from each of the states $1,2,\ldots, d-1$.

(ii) From (i) argue that each of the states $1,2,\ldots, d-1$ is transient.

(iii) Recall that for $n = 2,3,\ldots$ and states i,j,

$$P^{n}_{ij} = \sum_{m=1}^{n} P(\tau_j = m \mid X_o = i)\, P^{n-m}_{jj}$$

where τ_j is the hitting time of j. Use this fact to claim that for $n = 2,3,\ldots$

$$E(X_n \mid X_o = i) = \sum_{j=1}^{d-1} jP^{n}_{ij} + d\, P(\tau_d \leq n \mid X_o = i).$$

(iv) From (ii) and (iii) above, prove that

$$\lim_{n\to\infty} E(X_n \mid X_o = i) = d\, f_{id},$$

$i = 1,2,\ldots, d-1$ where f_{ij} is the probability of ever reacting j from i.

(v) Prove from (*) that $E(X_n \mid X_o = i) = i$, for $n = 1,2,\ldots$ and $i = 0, 1, 2, \ldots$; finally show that

$$f_{id} = \frac{i}{d},\ i = 0,1,\ldots,d.$$

[2+3+4+2+2+3+4=20]

contd. 5/-

7.(a) Define a Markov pure jump process and develop the Chapnan - Kolnogorov equation.

(b) Show that in a Yule process $X(t)$, $EX(t)$ grows exponentially with t.

(c) A stochastic process $X(t)$, $t \geq 0$ may take the value 0 and 1 with probabilities $p_o(t)$ and $p_1(t)$ respectively. Given that

$$P(X(t+h)=1 \mid X(t)=0) = \alpha h + o(h),$$

$$P(X(t+h)=0 \mid X(t)=1) = \beta h + o(h).$$

as $h \longrightarrow \infty$, show that

$$p_o(t) = \beta(\alpha+\beta)^{-1} + \left\{ p_o(0) - \beta(\alpha+\beta)^{-1} \right\} e^{-(\alpha+\beta)t}$$

and give corresponding expression for $p_1(t)$. [4+9+7=20]

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream): 1988-89
Semestral-II Examination
Sample Surveys and Design of Experiments

Date: 5.5.89 Maximum Marks: 100 Time: 3 Hours

Note: Answer question 4 and any one question from Group A and any two questions from Group B.

Group A

1. Show that in the class of linear homogeneous unbiased estimators the sample mean based on SRSWOR is "admissible" but not "the minimum variance estimator" for the population mean. [13+12=25]

2. From each of K strata of a stratified population, samples are independently selected by PPSWR method. Suggest an unbiased estimator for the population total. Obtain a suitable formula for its variance. Verify if your estimator of the population total is admissible in a sense to be specified by you. If inadmissible, give an alternative estimator which is admissible. [3+5+15+2=25]

3. From a population of N fsu's two fsu's are selected by PPSWOR method and from each of the sampled fsu's sub-samples of m ssu's are independently selected by SRSWOR method. Obtain an unbiased estimator for the population total, a formula for its variance and an unbiased estimator for the latter. [5+10+10=25]

4. Data relating to land holding sizes and cultivated areas are given in the following table for a certain region.

serial no of holding (i)	holding size in acres (x)	cultivated area in acres (z)
1	21.04	2.70
2	12.59	1.76
3	20.30	1.47
4	16.16	1.64
5	23.82	1.56
6	1.79	1.79
7	26.91	5.44
8	7.68	2.45
9	66.55	3.26
10	141.80	3.20
11	28.12	3.90
12	28.29	1.95
13	8.29	1.95
14	7.27	2.20
15	1.47	.48

Divide the region into three strata according to holding size classes 0-8.99, 9.00-24.99 and "25 and above". Select a stratified sample of size 9 by proportional allocation using PPSWR method taking holding size as measure of size for each stratum. Estimate average cultivated area perholding and also estimate the variance of the estimator used. Estimate also the gain in efficiency of the proposed estimator over the usual mean per unit estimator based on an SRSWOR of size 9 from these 15 holdings. [2+4+8+8+3=25]

Group B

1. When are two latin squares said to be mutually orthogonal? Give a method of construction of n-1 pairwise orthogonal latin squares of order n. (n is a prime power). Define a BIBD. Construct a BIBD with 13 treatments, block size 4, every pair appearing in exactly one block. [3+8+4+10=25]

2. (a) Consider an experiment with factors A,B,C each having 3 levels. Write down the treatment contrasts representing the following factorial effects and justify the representations:
   (i) Quadratic effect of factor A
   (ii) Interaction between A and B which is linear in A and quadratic in B.
   (iii) Interaction between A and C with is quadratic in A and linear in C.

   (b) Consider the following factorial design involving three factors with two levels each. Recognise the effect (s) which is (are) confounded with the block effects.

Block 1	(0),	(c),	(ab),	(abc)
Block 2	(a),	(b),	(ac),	(bc)

where the symbols have their usual meanings. [15+10=25]

3. An experimenter is interested in testing whether the fibres produced by 4 machines in a factory have the same strength. Supposing that the strength of the fibre may also depend on its thickness suggest a suitable plan of the experiment and describe the computational procedure to analyse data available on taking 5 observations from each machine to carry out the test.

Write down an appropriate model and obtain least squares estimates of the measure of dependence of the strength on thickness. (i) under a null hypothesis and (ii) under an alternative hypothesis to be appropriately formulated by you to carry out the test. [8+17=25]

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream): 1988-89
Semestral-II Examination
Demography

Date: 12.5.89 | Maximum Marks: 100 | Time: 3 Hours

Note: The questions are of equal value
Answer any five questions.

1.(a) Give various defenitions o. 'demography'. Indicate the importance of study of demography.

(b) Describe various sources for obtaining data relating to 'demography' with reference to India.

2. What do you mean by 'infant mortality'? Describe different methods of computing the infant mortality rate. Comment on their merits and demerits.

3.(a) What do you mean by a 'life table'? Discuss Chaing's method of constructing a complete life table.

(b) For a certain life table $l_x = 20900 - 80x - x^2$

(i) what is the ultimate age in the table?

(ii) find μ_x, q_x and ${}_{10}p_{20}$.

4. What are the different measures of fertility commonly used? Discuss in details, the merits and demerits of the crude and general fertility rates.

5. A demographer intends to project the female population of a country for the years 1991, 1996, 2001 and 2006. He intends to use a time and age interval of five years. Suggest a suitable method for determining the ${}_5K_x^{(t)}$ values.

6.(a) Define a stable population. Write down the assumptions of a stable population theory.

(b) Derive the integral equation

$$\int_{\alpha}^{\beta} e^{-rx}\, p(x)\, m(x)\, dx = 1$$

and prove that it has exactly one real root $r = r_0$, find r_0.

7. Write short notes on any three of the following

(i) Myers' Blended method

(ii) Measures of reproduction

(iii) Lexis' diagram

(iv) Stationary population.

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year(M-stream): 1988-89
Semestral-II Examination
Theory and Methods of Statistics-II

Date: 9.5.89 Maximum Marks: 100 Time: 3 Hours

Note: 1. Answer any Five Questions.
2. Precise and complete answers are given more weight.

1.(a) Give a precise definition of unbiasedness of an estimator. Show by examples that (i) an unbiased estimator need not exist (ii) an unbiased estimator may be "unreasonable" (iii) more than one unbiased estimator may exist.

(b) When is a sequence of estimators said to be consistent? Derive a sufficient condition for a sequence of estimators to be consistent. Show that a consistent estimator need not be unique. [14+6=20]

(a) If $x_1, \ldots, x_n$ is a random sample from $N(\mu, 1)$ obtain two sufficient statistics for μ. Which one do you prefer? Justify. Obtain an unbiased estimator of σ^2 of $N(0, \sigma^2)$ based on a sample of size n and based on the sufficient statistic. What should be n?

(b) If $x_1, \ldots, x_n$ is a random sample from a uniform distribution on $(0, \theta)$, $\theta > 0$, derive a sufficient statistic for θ and use it to obtain a uniformly minimum variance unbiased estimator of θ. [10+10=20]

Stating the underlying assumptions clearly, derive the Cramer-Rao lower bound for the variance of an unbiased estimator $T(x)$ of a parametric function $g(\theta)$, based on a random sample $\underset{\sim}{x} : (x_1, \ldots, x_n)$ from a population with p.d.f $f(x, \theta)$, $\theta \in \mathbb{R}$. Is the minimum variance bound always attained? Justify your answer.

(b) If $T_1(x)$ and $T_2(x)$ are two uniformly minimum variance unbiased estimators of a parametric function show that cor $(T_1, T_2) = 1$. [14+6=20]

4.(a) If $x_1, \ldots, x_n$ is a random sample from $N(\mu, 1)$ derive a Uniformly Most Powerful Test for $H_0: \mu \leq \mu_0$ against $H_1: \mu > \mu_0$, of a given size.

(b) If n=5, $\mu_0 = 3$ and $\alpha = 0.05$ draw the powerfunction of the test.

(c) Show how you can obtain a Uniformly Most Accurate upper confidence bound for μ using the above result.

(d) If you need a test of size 0.05 and of power at least 0.9 for $H_0: \mu \leq 3$ against the alternative $\mu = 4$ what is the minimum sample size you would require? [8+4+4+4=20]

p.t.o.

5.(a) If $x_1,\ldots,x_n$ is a random sample from $N(0,\sigma^2)$ show that there does not exist a UMP Test of size α for H_0: $\sigma^2=\sigma_0^2$ against H_1: $\sigma^2 \neq \sigma_0^2$.

(b) Derive a Uniformly Most Powerful Unbiased Test of size α for testing H_0: $\sigma^2 = \sigma_0^2$ against H_1: $\sigma^2 \neq \sigma_0^2$.

State precisely the main results you use in deriving the above test.

Show that the power function of this test increases as σ^2 moves away from σ_0^2. [6+14=20]

.(a) Stating clearly the regularity conditions, show that the maximum likelihood estimator which exists as a consistent solution of the likelihood equation with probability going to one as the sampl size increases, is asymptotically Normal. What can you say about the asymptotic variance of this estimator?

(b) If the family of distributions has a sufficient statistic T(x), show that the maximum likelihood estimator is an explicit function of T(x). Illustrate your answer with an example.

[14+6=20]

:SS:

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream): 1988-89
Semestral-II Examination
Linear Stat. Models and Large-sample Stat. Methods

Date: 28.4.89 Maximum Marks: 100 Time: $3\frac{1}{2}$ Hours

Note: Answer all questions. Use separate answer sheets for each part.

PART-A

Suggested time: $1\frac{1}{2}$ hours
Max. Marks: 50

1. State and Prove Gauss-Markoff theorem. [10]
2. Describe ANCOVA model in detail illustrating with an example where it can be applied. Derive test statistics for testing linear hypotheses of interest on the parameters of the model. [15]
3. Write short notes on
   (a) General regression and Linear regression
   (b) Tukey's test for nonadditivity. [5+5=10]
4. To compare the average yields of three different varieties of wheat an experiment was conducted in which each variety was tried in three plots in each of four villages selected at random. The yields are given in the table in convenient units. Analyze the data

Yields

Village	Variety V_1	V_2	V_3
1	8	7	8
	2	3	6
	5	5	10
2	4	5	11
	9	3	9
	5	2	8
3	6	8	12
	8	7	10
	5	10	15
4	7	10	8
	5	3	9
	9	8	11

[15]

5. Assignments [10]

p.t.o.

PART-B

Suggested time: 2 Hours
Maximum Marks: 50

EITHER

1. Let $x_1, x_2, \ldots, x_n$ be i.i.d. observations each following a distribution F_θ depending on an unknown real parameter θ. Let $\hat{\theta}_n$ be a strongly consistent solution of the likelihood equation.
Linearise $\hat{\theta}_n$ and find its asymptotic distribution under suitable assumptions (to be stated clearly).
Are two consistent solutions of the likelihood equation "asymptotically equivalent"? Explain. [15]

OR

Suggest suitable large sample tests for the following problems.

(a) We have a random sample of n pairs from a bivariate normal population with unknown correlation coefficient ρ. Our problem is to test $H_0 : \rho = \rho_0$ vs $H_1 : \rho \neq \rho_0$ ρ_0 is specified, $\rho_0 \neq 0$.

(b) We have two independent random samples of sizes n_1 and n_2 from two bivariate normal populations with population correlation coefficients ρ_1 and ρ_2 respectively. We are to test $H_0 : \rho_1 = \rho_2$ vs $H_1 : \rho_1 \neq \rho_2$.

What is the asymptotic distribution of the test statistic in (a) under the null hypothesis? Use this result to find the asymptotic distribution of the test statistic in (b) under H_0 when $n_1 \longrightarrow \infty$, $n_2 \longrightarrow \infty$ in such a way that $\frac{n_1}{n_2} \longrightarrow \lambda$ for some $0 < \lambda < \infty$. [15]

EITHER

2. State the representation theorem for sample quantiles and use this to prove

(i) asymptotic normality of any p-th quantile $Y_{p,n}$.

(ii) asymptotic normality of $(Y_{p_1,n}, Y_{p_2,n})$, $0 < p_i < 1$, $i = 1,2$ [20]

OR

Suppose that a population consists of k mutually exclusive classes, the proportion of members falling in the i-th class being π_i, $i = 1,2,\ldots,k$. Let a random sample of size n be drawn from the population and n_i be the number of members falling in the ith class. We want to test $H_0 : \pi_i = \pi_{i0}$, $i = 1,2,\ldots,k$. Suggest a test statistic and find its asymptotic distribution under H_0. [20]

Contd. 3/-

3. The following data relate to the distribution of 1725 school children who are classified according to their (i) economic conditions and (ii) intelligence
The standard in the latter case are (A) slow and dull (B) dull (c) slow but intelligent (D) fairly intelligent (E) distinctly capable (F) very capable.

Intelligence / Economic Condition	A	B	C	D	E	F
Very good	33	48	113	209	194	39
Good	41	100	202	255	138	15
Not good	56	71	92	71	43	5

Is there any reason to believe that there is association between economic condition and intelligence ? [15]

1988-89	524(b)

INDIAN STATISTICAL INSTITUTE

M.Stat. (M-stream) I Year : 1988-89

BACKPAPER (SEMESTRAL-I) EXAMINATION

Mathematical Analysis IM

Date: 2.1.1989 Maximum Marks: 100 Time: 3 hours

Note: Answer ALL questions. The paper carries 100 marks.

1. Let f be a continuous function on $\mathbb{R}^m$ into $\mathbb{R}^m$, say $f(\bar{x}) = (f_1(\bar{x}), \ldots, f_m(\bar{x}))$ where $\bar{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$. Define g by $g(\bar{x}) = \max_i f_i(\bar{x})$. Is g continuous ? Give reasons for your answer. [10]

2. Let f be a continuous function on $I = [0,1] \times [0,1]$. Show that there is a point (x_0, y_0) such that $\int_I f(x,y)\, d(x,y) = f(x_0, y_0)$. [10]

3.(a) Let f be a continuous real-valued function on $\mathbb{R}^2$ such that $f(x,y) = 0$ when x is rational and y irrational. Show that $f(x,y) = 0$ for all x,y. [10]

(b) Let $\{\bar{x}^{(n)}\} = \{(x_1^{(n)}, \ldots, x_m^{(n)})\}$ be a sequence in $\mathbb{R}^m$. Show that $\lim_{n \to \infty} (x_1^{(n)}, \ldots, x_m^{(n)}) = (x_1, \ldots, x_m)$ if and only if $\lim_{n \to \infty} \max_{1 \le i \le m} |x_i^{(n)} - x_i| = 0$. [15]

4. Let L be the line given by $y = ax + b$ and let f be an analytic function on $\mathbb{C}$ such that $f(z) \in L$ for all z. Show that f is constant. [10]

5.(a) Let f be such that for any closed path γ in $\mathbb{C}$, $\int_\gamma f(z)dz = 0$. Show that f is analytic. [15]

(b) Use contour integration to compute $\int_0^\infty \frac{\sin x}{x} dx$. [10]

6. If $f(z)$ is analytic everywhere and $g(z) = f(\frac{1}{z})$ has a removable singularity at $z = 0$, show that f is constant. [15]

:bcc:

1988-89 | 524

INDIAN STATISTICAL INSTITUTE

M.Stat. (M-stream) I Year : 1988-89

FIRST SEMESTRAL EXAMINATION

Mathematical Analysis II

Date: 21.11.1988 Maximum Marks. 100 Time. 3 hours

Note: The paper carries 115 marks. The maximum you can score is 100.

1. Let f be a uniformly continuous function defined on a bounded subset of $\mathbb{R}^n$ into $\mathbb{R}$. Show that f is bounded. [15]

2. Let $f(x,y) = x^2 y^2 \log(x^2 + y^2)$ if $(x,y) \neq (0,0)$

   $f(0,0) = 0$.

   Does f have a differential at $(0,0)$? Give reasons for your answer. [15]

3.(a) Let f be a real valued continuous function on $J = [0,1] \times [0,1]$. Show that $\int_J f(x,y)\, d(x,y)$ exists and

$$\int_J f(x,y)\, d(x,y) = \int_0^1 \int_0^1 f(x,y)\, dx\, dy$$

$$= \int_0^1 \int_0^1 f(x,y)\, dy\, dx.$$ [20]

(b) If in part (a), f is non-negative and $\int_J f(x,y)\, d(x,y) = 0$, show that $f(x,y) = 0$ for all $(x,y) \in J$. [15]

4.(a) Is the function f defined on $\mathbb{C}$ by $f(z) = \bar{z}$ analytic ? Give reasons for your answer. [5]

(b) If f is analytic on $\mathbb{C}$ and g is defined by $g(z) = \overline{f(\bar{z})}$, show that g is analytic and $g'(z) = \overline{f'(\bar{z})}$. [10]

5. Let $D = \{ z : |z| < 1 \}$ and A an infinite subset of $\{ z : |z| = \frac{1}{2} \}$. Is there an analytic function f on D such that $f(z) = 2$ if $z \in A$ and $f(0) = 1$? Justify your answer. [10]

6. Calculate

(a) $\int_0^\infty \frac{x^2}{(x^2 + 4)^3} dx$. [13]

(b) $\int_{|z| = 2} z^{-3} (1 - z)^{-2} dz$. [12]

1988-89	554 and 564

INDIAN STATISTICAL INSTITUTE
M.Stat. (M-stream) I Year : 1988-89

FIRST SEMESTRAL EXAMINATION

Linear Algebra and Regression and Correlation

Date: 24.11.1988 Maximum Marks: 100 Time: 3 hrs.

Note: Answers to different Parts should be in different answerscripts.

PART - I Max.Marks: 50

Linear Algebra

Note: Answer any TWO questions. Each question carries 25 marks.

1. Let J_{np} denote a matrix of order $n \times p$ with all elements equal to unity.

(a) Find the inverse of $a\, I_n + b\, J_{nn}$ when it exists.

(Hint: the inverse can be expressed in the same form).

(b) Find the inverse of

$$Q = \begin{bmatrix} K I_n & a J_{np} \\ a J_{pn} & m I_p \end{bmatrix}$$

when it exists.

(10+15) = [25]

2.(a) Prove that the characteristic roots of an idempotent matrix are either zero or unity. Find the matrix of the quodratic form $p\bar{x}^2 = x'Ax$ where x is a $p \times 1$ vector with elements x_i and $\bar{x} = \frac{1}{p} \sum_{i=1}^{p} x_i$ and show that the matrix is idempotent. Find the rank of A.

(5+4+3) = [12]

(b) If $x'Ax$ is a real quadratic form in n variables, $x' = (x_1, \ldots, x_n)$ and $\lambda_1, \ldots, \lambda_n$ are the characteristic roots of A, show that we can find an orthogonal matrix P of order $n \times n$ such that the transformation $x = Py$, where $y' = (y_1, \ldots, y_n)$ transforms $x'Ax$ to the diagonal form

Contd..... 2/-

Contd..... Q.No.2.(b)

$\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$. Illustrate the method by finding P when

$$x'Ax = 3x_1^2 + 5x_2^2 + 3x_3^2 - 2x_1x_2 + 2x_1x_3 - 2x_2x_3.$$

(6+7) = [13]

3.(a) Prove that the system $Ax = b$ is consistent iff $\rho(A) = \rho(A:b)$ where $\rho(A)$ denotes the rank of A.

(b) Prove that $Ax = b$ is consistent iff there does not exist any u such that $u^T A = 0$ and $u^T b \neq 0$.

(c) Show that any consistent system of equations. $Ax = b$ is equivalent to some system of r linear equations where r is the rank of A.

(5+10+10) = [25]

PART - II

Regression and Correlation Max.Marks: 50

1. Let $M = \begin{pmatrix} \bar{X}_0 \\ \bar{X} \\ \bar{X}_p \end{pmatrix}$ be the mean vector and $S = \begin{bmatrix} s_{00} & s'_0 & s_{0p} \\ s_0 & S & s_p \\ s_{p0} & s'_p & s_{pp} \end{bmatrix}$

be the variance-covariance matrix of the variable $\begin{pmatrix} X_0 \\ X \\ X_p \end{pmatrix}$,

where $X = (X_1\ X_2 \ldots X_{p-1})$, based on n-observations:

(i) Derive the expressions for the following interms of elements of M, S and S^{-1}.

(a) L(X), the Linear Vregression function of X_0 on X

(b) Multiple correlation coefficient $R_{0.12\ldots p-1}$,

and (c) Partial correlation coefficient $R_{0p.12\ldots p-1}$.

(ii) Further if $e = X_0 - L(X)$, the residual, show that

(a) $Cov(e,X) = \phi$ (null vector)

(b) $Cov(e,L(X)) = 0$

(c) $Cov(X_0) = V(L(X)) + V(e)$

(d) $Cov(X_0,L(X)) = V(L(X))$

Contd..... Q.No.1.(ii)

(e) $Cor(X_0, L(X)) \geq Cor(X_0, Z(X))$ where $Z(X)$ is any linear function of $X_1 \ldots X_{p-1}$.

[30]

2. The following table gives the mean vector and variance covariance matrix of 4 variables X_0, X_1, X_2, X_3 based on 100 observations.

Obtain (a) Linear regression function of X_0 on X_1 and X_2
(b) Compute multiple correlation coefficient $R_{0.12}$ and
(c) Partial correlation coefficient $R_{03.12}$.

	Cov. matrix X_0	X_1	X_2	X_3	Mean
X_0	1	1	2	2	0
X_1	1	2	3	3	0
X_2	2	3	6	7	0
X_3	2	3	7	10	0

[20]

3. Assignments. [10]

1988-89	571

INDIAN STATISTICAL INSTITUTE

M.Stat. I Year (M-stream): 1988-89

FIRST SEMESTRAL EXAMINATION

Economic Statistics

Date: 25.11.1988 Maximum Marks. 100 Time: 3 hours

GROUP - A

Note: Answer any TWO Max.Marks. 40

1.(a) Let $p_1(0)$, $p_2(0)$, ..., $p_n(0)$ and $p_1(T)$, $p_2(T)$, ..., $p_n(T)$ be the prices of n different commodities observed at time 0 and T respectively. What are the alternative approaches using which the overall price change between these two time points can be meaning fully measured ? Describe the rationale of these approaches.

(8+4) = [12]

(b) The average prices of cereal items in a certain wholesale market as observed during two different weeks are given below along with the corresponding figures of volumes transacted.

Item	Price (Rs. per maund) during the week ending on		Volume transacted(maunds) during the week ending on	
	17.11.1956	21.12.1957	17.11.1956	21.12.1957
(1)	(2)	(3)	(4)	(5)
rice	20.50	17.50	22,500	25,500
wheat	18.50	15.40	10,500	8,500
jawar	16.25	12.50	1,950	1,800
bajra	15.50	12.40	1,050	1,450

using the above data calculate two different price indices of cereals for the week ending on 21.12.1957 with the week ending on 17.11.1957 taken as the base which are based on different approaches.

[8]

2.(a) What is a cost of living index number ? Describe briefly the steps and the problems involved in the construction of such an index number.

(3+7) = [10]

Contd..... 2/-

Contd..... Q.No.2

(b) The Table below gives some incomplete information about the change in consumer prices faced by the working class households in a certain locality:

item group	group weight	consumer price index (base : 1960) for the year	
		1973	1974
food items	60	198	252
fuel and light	*	175	195
clothing	12	*	200
house rent	20	150	150
other items	*	135	212
all items	100	180	*

Where every * denotes a missing figure. It is also given that the consumer price index number for 1973 based on prices of food items, fuel and light and clothing only is 185. Calculate the all items consumer price index number for the year 1974.

[10]

3.(a) Explain the rationale of constructing chain-base index number series. Is a series of chain-base index number always superior to the corresponding series of fixed base index numbers ? Give reasons for your answer.

(6+4) = [10]

(b) The Table below gives the price and quantity of four commodities observed in four successive years.

commodity	price (Rs/Kg) in year				quantity(Kg) in year			
	0	1	2	3	0	1	2	3
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A	5	6	4	7	100	80	120	70
B	4	3	5	6	80	90	70	60
C	2	5	3	4	60	30	50	40
D	10	8	9	15	30	50	50	20

Calculate Laspeyres' fixed base and the corresponding chain base price index number for year 3 taking year 0 as the base for the above data.

(5+5) = [10]

GROUP - B Max.Marks: 40

Note: Answer any TWO

4.(a) Describe the nature of the different components that may possibly be present in a time series of monthly observations on an economic variable which covers several decades. Explain briefly why an analysis of such a time series may be useful.

(6+5) = [11]

(b) With which component of a time series would you associate the following phenomena. In each case give reason for your answer:

(i) a labour unrest in a factory hampering the volume of daily production;

(ii) a rise in demand for foodgrains due to the growth of population in a country;

(iii) a relatively higher price of vegetables in the Calcutta markets during monsoon every year.

(3 x 3) = [9]

5.(a) What are seasonal indices ? Describe an appropriate method of obtaining the constant seasonal indices of an observed time series.

(4+8) = [12]

(b) The pattern of seasonality of sales of a shop during a year is given below:

Quarter of the year	Jan.-Mar.	April-June	July-Sept.	Oct.-Dec.
Constant seasonal index	97	85	83	135

If the shopowner realizes a total sale of 15,000 units during the first quarter of a year, how much stock should be keep during the other quarter of the year in order to avoid possible shortage of supply.

[8]

6. Write short notes on the following:

(i) Additive and Multiplicative models of time series;

(ii) Logistic trend curve;

Contd..... 4/-

Contd..... Q.No.6

(iii) Cyclical variations of a time series;

(iv) Estimation of the trend by the method of moving averages.

(4 x 5) = [20]

GROUP - C

Note: Any ONE Max.Marks: 20

7. Write a note on the nature and the source of wholesale price index numbers available in India.

[20]

7. Discuss briefly the role of the central statistical organization in the official statistical system of India.

[20]

1988-89	514

INDIAN STATISTICAL INSTITUTE

M.Stat. (M-stream) I Year : 1988-89

FIRST SEMESTRAL EXAMINATION

Probability Theory IM

Date: 26.11.1988 Maximum Marks: 100 Time: $3\frac{1}{2}$ hrs.

Note. The paper carries 116 marks. Answer as much as you can. The maximum you can score is 100.

1. Consider Polya's urn scheme with r red balls and b black balls with c new balls added each time. Let X_i be 1 or 0 according as the ith ball drawn is red or black.

(a) Show that the joint distribution of $(X_1, \ldots, X_n)$ is the same as the joint distribution of $(X_{\pi(1)}, \ldots, X_{\pi(n)})$ for each permutation π of $\{1, \ldots, n\}$. [6]

(b) Compute the conditional probability that the first ball drawn is red given that the third ball is black. [4]

(c) What is the expected number of red balls drawn in n trials ? [4]

2.(a) State the Borel-Cantelli Lemma. [5]

(b) Use the Borel-Cantelli Lemma to prove that in the classical Gambler's ruin problem the game stops with probability one. [8]

3.(a) State the Borel Strong Law of Large Numbers. [5]

(b) Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common distribution function F.

Let $F_n(\omega,x) = \frac{1}{n} \# \{ i : X_i(\omega) \leq x \}$ be the empirical distribution function. Show that for each x,

$$P(\{\omega : \lim_{y \uparrow x} F_n(\omega,y) \xrightarrow[n]{} F^-(x)\}) = 1 .$$ [7]

(c) In the same set-up as above, let

$$A = \{ \omega : F_n(\omega,x) \to F(x) \text{ uniformly in } x \} .$$

Why is A an event ? [8]

Contd..... 2/-

Contd..... Q.No.3

(d) State the Glivenko-Cantelli Theorem. What does it say in relation to (b) ?

[5]

4. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. Exp.(λ) random variables and let $\{N_t\}_{t>0}$ stand for the associated Poisson process.

(a) Put $E_t = t - S_{N_t}$. Show that E_t has the same distribution as min (X_1, t).

[Hint: Express the event $\{E_t > u\}$ in terms of a Poisson increment, for $u < t$]

[12]

(b) Put $L_t = S_{N_t+1} - S_{N_t}$ be an interarrival time.

Show $\lim_{t \to \infty} E(L_t) = 2/\lambda$.

[Hint: Use (a) and results proved in the class].

[8]

5. You enter a casino to play a roulette wheel game. The wheel has 18 red, 18 black, and 2 green sectors. You decide to bet either on the red or on the black in each trial. You start with a capital of Rs.900/- and decide to play until you either reach Rs.1000/- or lose your capital. The rules specify that you play at constant stakes and you are allowed to choose any stake between Re.1 and Rs.100/-.

(a) What will your choice be ? Justify your answer in detail.

[15]

(b) For the choice you make, what is the expected duration of the game?

[12]

6. Let X and Y be two independent random variables each uniform on [0,1]

(a) Find the probability density function of X+Y explicitly.

(b) Find the (cumulative) distribution function of $Z = \min(X,Y)$.

(c) Find the joint probability density function of U and V, where $U = X^2$ and $V = X/Y$

(5+5+7) = [17]

1988-89 | 534

INDIAN STATISTICAL INSTITUTE

M.Stat. (M-stream) I Year . 1988-89

FIRST SEMESTRAL EXAMINATION

Theory and Methods of Statistics I

Date: 18.11.1988 Maximum Marks: 100 Time: $3\frac{1}{2}$ hrs.

1. Admission data for the graduate programmes in six largest departments in a certain university are given below.

Department	Men		Women	
	No. of applicants	Percent admitted	No. of applicants	Percent admitted
A	825	62	108	82
B	560	63	25	68
C	325	37	593	34
D	417	33	375	35
E	191	28	393	24
F	373	6	341	7

A politician pointed out that the university admission procedure had discriminated against Women, since 44% of the male applicants had been admitted whereas only 30% of the female applicants had been admitted. Comment on the above statement and support your comments based on the data provided.

[15]

2. In 1958 a doctor introduced a new technique for treating ulcers. He tried the method on 24 patients, and all were cured. Comment on the scientific validity of this result which tends to promote the new technique instead of the standard treatments which requires surgery. [15]

3. Represent the following data graphically. [15]

Percentage of literates in India by age and sex, 1971

Age group in years	male	Female
5 - 9	27.[illegible]	18.9
10 - 14	59.8	38.1
15 - 19	63.4	37.7
20 - 24	60.7	28.7
25 - 34	50.1	19.3
35 and above	38.0	10.7

p.t.o.

4.(a) Discuss a practical situation where the Binomial model could be used.

(b) The following table gives the frequency distribution of the number of dust nuclei in a small volume of air that fall on to a stage in a [illegible] containing moisture and filtered air :

Frequency distribution of dust nuclei

No. of dust nuclei	Frequency
1	5[illegible]
2	[illegible]
3	95
4	73
5	4[illegible]
6	17
7	5
8	3

Examine whether a suitable Poisson distribution fits the data.

(3+12) = [15]

5.(a) What are the advantages of stratified sampling ?

(b) A group of 112 students is divided into 2 strata consisting of 80 Masters and 32 Bachelors. From the first stratum a simple random sample of 8 students is selected without replacement while from the second stratum a simple random sample of size 4 is selected with replacement. The number of study hours per week (outside the class) are recorded for all the selected students as follows.

stratum 1 : 18, 16, 1[illegible], 17, 20, 19, 18, 22
stratum 2 : 10, 8, 11, 7.

Estimate the average number of study hours for this group. Also obtain an unbiased estimate of the sampling error of your estimate.

[3 + (5+7)] = [15]

6.(a) Suppose that X, the length of an item has a Normal distribution with mean 10 and variance 2. The Quality Control Manager classfies the items into three categories according as $X < 8$, $8 \leq X < 12$ and $X \geq 12$. If 21 such items are produced in a shift, what is the probability that an equal

Contd..... 3/-

Contd..... Q.No.6.(a)

number of items belong to each of the above categories.

(b) Wehn do you say that a discrete random variable has a 'hypergeometric distribution' ? For large N, explain with a numerical illustration how this can be approximated by the Binomial distribution.

(c) Let X be a random variable having an exponential distribution with parameter λ, show that $Pr.(X > s+t \mid X > s) = P(X > t)$. What is the significance of this result ?

(5+5+5) = [15]

7. Practical records. [10]

1988-89	544

INDIAN STATISTICAL INSTITUTE

M.Stat. (M-stream) I Year : 1988-89

FIRST SEMESTRAL EXAMINATION

Computational Techniques and Programming

Date: 16.11.1988 Maximum Marks: 85 Time: $3\frac{1}{2}$ hrs.

1.(a) State the general formula for errors.

(b) Find the number of trustworthy figures in the quotient of 876.3/494.2, assuming that both numbers are approximate and true only to the number of digits given.

(4+6) = [10]

2. Write a FORTRAN program to compute the following sum correct upto 3 places of decimal:

$$x + \frac{x^3}{1.3} + \frac{x^5}{1.3.5} + \frac{x^7}{1.3.5.7} + \cdots\cdots$$

$$0 < x < 1$$

[15]

3. $(x_1, x_2, x_3, x_4, x_5)$ is multinomially distributed with $n = \sum_{i=1}^{5} x_i = 197$ and the cell probabilities given by

$\left(\frac{1}{2}, \frac{1}{4}\theta, \frac{1}{4}(1-\theta), \frac{1}{4}(1-\theta), \frac{1}{4}\theta\right), \quad 0 \le \theta \le 1.$

We observe $x_1 + x_2 = 125$, $x_3 = 18$, $x_4 = 20$ and $x_5 = 34$.

Write down the EM Algorithm to find the maximum likelihood estimate of θ. Find the estimate of θ using this algorithm with the initial estimate $\theta^{(0)} = 0.5$ (with the minimum of accuracy 0.01 or 5 iterations).

[15]

4. If x is $p \times 1$ vector and A a $p \times p$ positive definite matrix show that

$$\begin{vmatrix} A & x \\ \cdots & \cdots \\ x' & 1 \end{vmatrix} = |A| \, (1 - x'A^{-1}x).$$

Hence develop a method for numerical evaluation of $x' A^{-1} x$ by sweep-out.

Contd..... Q.No.4

Use your method to compute $x' A^{-1} x$ when

$$x = (.5, .6, .7) \text{ and } A = \begin{bmatrix} 1 & .3 & .3 \\ .3 & 1 & .3 \\ .3 & .3 & 1 \end{bmatrix}$$

(5+10+10) = [25]

5.(a) What is a basic feasible solution to an LP problem ?

(b) Find an optimal solution to the following LP problem by computing all basic feasible solutions and then finding one that maximizes the objective function.

$$\begin{aligned} \max \quad z &= 2x_1 + 3x_2 + 4x_3 + 7x_4 \\ \text{s.t.} \quad & 2x_1 + 3x_2 - x_3 + 4x_4 = 8 \\ & x_1 - 2x_2 + 6x_3 - 7x_4 = -3 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

(4+16) = [20]

:bcc: