INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1991-92
SEMESTRAL-II BACKPAPER EXAMINATION
PROBABILITY THEORY II

Date: 29.6.92 Maximum Marks: 100 Time: 3 Hours

1. State and prove Monotone class Theorem. [3+13=16]

2. Let $(\Omega, \mathcal{A}, P)$ be any probability space and X an absolutely continuous random variable on it with density function h. Show that if $f: \mathbb{R} \to \mathbb{R}$ is a borel measurable function, then the random variable $f(X)$ is P-integrable if and only if the function fh on $\mathbb{R}$ is λ-integrable and in that case $E(f(X)) = \int_{\mathbb{R}} f(x)h(x)dx$. [15]

3.(a) State clearly what is meant by the mutual independence of an arbitrary collection of random variables.

(b) Prove that random variables $X_1, \ldots, X_n$ are independent if and only if for every n-tuple $x_1, \ldots, x_n$ of real numbers

$$P\left(\bigcap_{i=1}^{n} (X_i \leq x_i)\right) = \prod_{i=1}^{n} P(X_i \leq x_i)$$ [4+13=17]

4. Let f be a non-negative real-valued measurable function on a measure space $(\Omega, \mathcal{A}, \mu)$. Consider the product space $(\mathbb{R} \times \Omega, \mathcal{B} \otimes \mathcal{A}, \lambda \otimes \mu)$.

(a) Show that the set $A = \{(x, \omega) \in \mathbb{R} \times \Omega : 0 \leq x < f(\omega)\}$ belongs to $\mathcal{B} \otimes \mathcal{A}$.

(b) Find $\lambda \otimes \mu (A)$. [10+10=20]

5.(a) Define what is meant by convergence in probability.

(b) Show that if $X_n \xrightarrow{P} X$, then for any continuous $f: \mathbb{R} \to \mathbb{R}$, $f(X_n) \xrightarrow{P} f(X)$ [3+13=16]

6. State and prove Kolmogorov's Maximal Inequality. [3+13=16]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1991-92
SEMESTRAL-II BACKPAPER EXAMINATION
APPLIED STOCHASTIC PROCESSES

Date: 22.6.92 | Maximum Marks: 100 | Time: $3\frac{1}{2}$ Hours

Note: Answer all the questions.

1. In the questions below, $\{X_n : n \geq 0\}$ is a Markov chain with stationary transition probabilities.

   (a) Show that the one-dimensional symmetric (i.e., $p=\frac{1}{2}$) random walk on the integers is recurrent.

   (b) Suppose x and y are two communicating states. Show that $d(x)=d(y)$.

   (c) Suppose $\underset{\sim}{\pi}$ is a stationary initial distribution for $\{X_n : n \geq 0\}$. If x is a transient or a null recurrent state, then show that $\pi(x)=0$.

   (d) Suppose the transition probability matrix $P=((p_{ij}))_{i,j \in S}$, $S=\{0,1,\dots,n\}$, is of the form $p_{00}=p_{nn}=1$, $p_{i,i+1}=p$, $p_{i,i-1}=1-p$, $i=1,\dots,n-1$, where $0<p<1$, but $p \neq \frac{1}{2}$.

   Find $u_i = P(X_n=0 \text{ for some } n \mid X_0=i)$, $i=1,\dots,n-1$.

   [10+9+11+10=40]

2. (a) Let $\{N_t : t \geq 0\}$ be a Poisson process with parameter α. Suppose each event (independently of others) is recorded with probability p, and erased with probability 1-p. Show that the random variables $\{N_t^{(1)} : t \geq 0\}$ of the number of recordedevents constitute a Poisson process with parameter αp.

   (b) State and prove a result regarding existence of various moments of N(t), $\{N(t) : t \geq 0\}$ being a renewal process.

   (c) Consider a renewal process $\{N(t) : t \geq 0\}$ with inter-arrival times $X_1, X_2, \dots$ having density

$$f(x) = \begin{cases} x e^{-x}, & x > 0 \\ 0, & x < 0. \end{cases}$$

   (i) Compute $P(N(t)=n)$, $n=0,1,2,\dots$.

   (ii) Compute $M(t)=E(N(t))$.

   (iii) Verify that $\lim_{t\to\infty} \frac{M(t)}{t} = \frac{1}{\mu}$, where $\mu = E(X_1)$. [10+12+18=40

p.t.o.

3. Answer anyone of (a) and (b):

(a) Suppose in a society, the family name is carried over from one generation to the next by the male offsprings only. Consider initially a male individual in this society. Let X_1 denote the number of his male children; thus X_1 is the number of male individuals who belong to the first generation. In general, let X_n denote the number of male individuals belonging to the nth generation of the family started by the initial male individual, $n=1,2,\ldots$, thus X_n is the total number of male children born to the male individuals belonging to the (n-1)-th generation each of whose members independently produces a random number of children (male or female), the offspring distribution being $\{p_k : k=0,1,2,\ldots\}$. Suppose

$$p_k = ap^k, \quad k=1,2,\ldots$$

where $a>0$, $0<p<1$ are constants. Assume that each child is equally likely to be a male or a female.

(i) Give a sufficient condition relating p and a for the probability of extinction of the family name to be non-nul

(ii) Hence, specify a range of values for p so that correspondingly the family name becomes extinct with certainty.

(b) Let $\{X_n : n \geq 0\}$ be a Markov chain with stationary transition probabilities. Suppose the transition probability matrix

$p = ((p_{ij}))_{i,j \geq 0}$ is of the form $p_{i0} = p_i$, $p_{i,i+1} = 1-p_i$, $i \geq 0$, where $0<p_i<1$ for $i \geq 0$.

(i) Show that $\{X_n : n \geq 0\}$ is imeducible.

(ii) Show that the states are recurrent iff $\sum_{i=0}^{\infty} p_i = \infty$.

(iii) Show that when the states are recurrent, they are positive recurrent iff $\sum_{j=0}^{\infty} \{ \prod_{j=0}^{j} (1-p_i)\} < \infty$. [20]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR:1991-92
SEMESTRAL II EXAMINATION
SAMPLE SURVEYS AND DESIGN OF EXPERIMENTS

Date:8.5.92 Maximum Marks:100 Time: 3 Hours

Sample Survey

Maximum Marks:50

Note: Attempt question 4 and any two of the rest.

1. Define an unicluster sampling design. Show that the Horvitz-Thompson estimator based on an unicluster sampling design with positive inclusion probability for each population unit is the minimum variance homogeneous linear unbiased estimator of the population total. What are the practical disadvantages of unicluster designs? [16]

2. Consider SRSWOR (N,n) sampling design. Show that,
   (i) the sample mean is an unbiased estimator of the population mean. Deduce an expression for its variance.
   (ii) the sample variance (with divisor (n-1)) is an unbiased estimator of the population variance (with divisor (N-1)).
   (iii) using (ii), find estimated standard error of the sample mean in (i). [4+4+6+2=16]

3. (a) What is stratified sampling?
   (b) What is stratified simple random sampling?
   (c) Under stratified simple random sampling,
      (i) suggest an unbiased estimate of the population proportion.
      (ii) provide an expression for estimated standard error of the estimate in (i). 2+2+(4+8)=16

4. Consider the following data set arising out of stratified simple random sampling without replacement. Compute the value of an estimate of the population mean and its estimated standard error. What is the estimated gain due to stratification over unstratified SRSWOR sample mean?

Stratum	Stratum size	Sample size	Sample total	Sample s d.
1	65	20	98.38	2.58
2	75	25	144.85	3.69
3	60	17	70.25	4.28
4	80	27	138.35	3.18
5	120	37	150.27	5.02

4+6+8=18

Design of Experiments

Maximum Marks:50

Note: Answer question 1 and any 2 of the rest.

1.(a) In a pxp Latin Square design, if the observation corresponding to the i-th row, j-th column and k-th treatment is missing, show that this missing value may be estimated by

$$y_{ijk} = \frac{p(y'_{i..}+y'_{.j.}+y'_{..k}) - 2y'_{...}}{(p-2)(p-1)}$$

where $y'_{i..}$, $y'_{.j.}$, $y'_{..k}$ indicate the totals of the available observations of the i-th row, j-th column and k-th treatment respectively and $y'_{...}$ is the grand total of the available observations.

(b) The effect of 4 different component materials (a,b,c,d) on the strength of plastics produced is being measured. The following Latin Square design is used over 4 days and with 4 different temperatures. The data on the first day with material 'a' is missing. Estimate this missing value. Analyze the data using this value and draw conclusions.

Temperatures

Day	1	2	3	4
1	a=*	b=11	c=12	d=10
2	d=11	a=11	b=13	c=15
3	b=8	c=19	d=12	a=9
4	c=11	d=15	a=8	b=9

8+10=18

2. An experiment is to be conducted to study the effect of 4 factors, each having 2 levels. The experimenter can run only 8 treatment combinations under similar conditions at one time. (a) Assuming that three-factor and higher order interactions are negligible, design an efficient experiment, including a suggested number of replications (b) Show the partitioning of the degrees of freedom in the analysis of variance. (c) Give the expression for the sum of squares for any one main effect and one 2-factor interaction in terms of the treatment combinations (d) What is the amount of information available on the different main effects and interactions?

4+4+4+4=16

3.(a) In a 3^3 experiment, show that if pairs of degrees of freedom corresponding to interactions X and Y are confounded with the blocks, then so are the pairs of d.f. corresponding to XY and XY^2.

(b) Give a suitable confounding scheme for a 3^3 experiment in blocks of 3 where you retain full information on all main effects and partial information on all the 2-factor and 3-factor interactions. For this confounding scheme give the partitioning of the degrees of freedom in the analysis of variance. 6+(6+4)=16

4.(a) In the analysis of covariance, what is the purpose of using the concomitant variables?

(b) In the case of one-way classification with a single concomitant variable, give an appropriate model for the analysis of covariance, assuming a linear relationship between the response (Y) and the concomitant variable (Z).

(c) Obtain an estimate of the linear regression coefficient of Y on Z. How will you test whether the inclusion of Z in the above model is justified? 4+2+(4+6)=16

1991-92 | 528

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1991-92
SEMESTRAL II EXAMINATION
THEORY AND METHODS OF STATISTICS II

Date: 6.5.92 Maximum Marks: 100 Time: 4 Hours

Note: Attempt all the questions. Marks for each question are given in brackets at the end of the question.

1. A random sample of size n is available from the distribution on the positive real numbers with density

$$\frac{x+1}{\theta(\theta+1)}\, e^{-\frac{x}{\theta}}$$

where θ (>0) is an unknown parameter. Examine whether each of the following functions of θ has an unbiased estimator whose variance attains Cramer- Rao lower bound. In case such an unbiased estimator exists, exhibit it explicitly.

(a) $(3 + 2\theta)\cdot(2+\theta)/(\theta+1)$

(b) $(4+\theta)\cdot(3+\theta)\cdot(2+\theta)/(\theta+1)$ [8]

2. Let $X_1, X_2, \dots, X_k$ be iid binomial (n,θ), $0 \leq \theta \leq 1$.

Obtain the UMVUE of $n\,\theta\,(1-\theta)^{n-1}$. (Notice that this is the probability of exactly one success from a binomial (n,θ).] [12]

3.(a) Show that an unbiased estimator $T(x)$ of $g(\theta)$ has minimum variance at $\theta=\theta_0$ iff $\mathrm{cov}_{\theta_0}(T(x), f(x)) = 0$ for every zero function $f(x)$ such that $V_\theta(f(x)) < \infty$.

(b) Let $Y \sim N_n(X\beta, \sigma^2 I)$ where $\beta \in \mathbb{R}^m$, $\sigma^2 > 0$ and $X_{n\times m}$ is of rank $m(<n)$. Let $\hat{\beta}$ be the least squares estimator of β. Show that $p'\hat{\beta}$ is the uniformly minimum variance unbiased estimator (UMVUE) of $p'\beta$.

(c) Consider the same set up as in (b). Show that

$(Y-X\hat{\beta})'(Y-X\hat{\beta})/(n-m)$ is the UMVUE of σ^2. [3x5=15]

p.t.o.

4.(a) Describe either the method of scoring or EM algorithm for Maximum Likelihood Method of estimating several parameters.

(b) Use either method of scoring or EM algorithm to compute Maximum Likelihood estimates or the gene-frequencies p, q and r of O, A, and B respectively using the following observed phenotypic frequencies of blood groups.

phenotype	frequency
O	180
A	190
B	62
AB	18

[6+19=25

5. Let X have the density

$$f(x \mid \theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

(a) Show that this family has MLR

(b) Based on one observation X, find the most powerful size α test for testing $H_0: \theta = 0$ against $H_1: \theta = 1$. For $\alpha = 0.2$ find the probability of type II error.

(c) Show that the test in part (b) is UMP level α test for testing $H: \theta \leq 0$ against $K: \theta > 0$. [4+6+2+6=18].

6.(a) Let f(x) be a unimodal density on positive real numbers and let $c.f(x)\,x^2$ be also a unimodel density for some $c > 0$. Let x^* be the mode of the latter distribution. Let $0 < \alpha < 1$ be a given number. Let a and b satisfy

(i) $\int_a^b f(x)\,dx = 1 - \alpha$

(ii) $f(a).a^2 = f(b).b^2$

(iii) $a < x^* \leq b$.

Show that $\frac{1}{a} - \frac{1}{b} \leq \frac{1}{u} - \frac{1}{v}$ for all u, v such that $0 \leq u \leq v$ and $\int_u^v f(x)dx = 1-\alpha$.

contd. 3/-

6.(b) Let $X_1, \dots, X_n$ be iid $N(0,\sigma^2)$. Obtain the shortest expected length confidence interval, with confidence coefficient $(1-\alpha)$, of the form

$$\left\{ \sigma^2 : \frac{\Sigma X_i^2}{a} \leq \sigma^2 \leq \frac{\Sigma X_i^2}{b} \right\} \text{ for}$$

σ^2. [9+7=16]

7. In 1000 tosses of a coin 580 heads and 420 tails appear. On it reasonable to assume that the coin is fair? Justify your answer. [6]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1991-92
SEMESTRAL II EXAMINATION
THEORY AND METHODS OF STATISTICS II

Date: 6.5.92 Maximum Marks: 100 Time: 4 Hours

Note: Attempt all the questions. Marks for each question are given in brackets at the end of the question.

1. A random sample of size n is available from the distribution on the positive real numbers with density

$$\frac{x+1}{\theta(\theta+1)} e^{-\frac{x}{\theta}}$$

where θ (>0) is an unknown parameter. Examine whether each of the following functions of θ has an unbiased estimator whose variance attains Cramer-Rao lower bound. In case such an unbiased estimator exists, exhibit it explicitly.

(a) $(3+2\theta)\cdot(2+\theta)/(\theta+1)$

(b) $(4+\theta)\cdot(3+\theta)\cdot(2+\theta)/(\theta+1)$ [8]

2. Let $X_1, X_2, \ldots, X_k$ be iid binomial (n,θ), $0 \leq \theta \leq 1$.

Obtain the UMVUE of $n\,\theta\,(1-\theta)^{n-1}$. (Notice that this is the probability of exactly one success from a binomial (n,θ).) [12]

3.(a) Show that an unbiased estimator $T(x)$ of $g(\theta)$ has minimum variance at $\theta=\theta_0$ iff $cov_{\theta_0}(T(x), f(x)) = 0$ for every zero function $f(x)$ such that $V_\theta(f(x)) < \infty$.

(b) Let $Y \sim N_n(X\beta, \sigma^2 I)$ where $\beta \in \mathbb{R}^m$, $\sigma^2 > 0$ and $X_{n\times m}$ is of rank $m(<n)$. Let $\hat{\beta}$ be the least squares estimator of β. Show that $p'\hat{\beta}$ is the uniformly minimum variance unbiased estimator (UMVUE) of $p'\beta$.

(c) Consider the same set up as in (b). Show that

$(Y-X\hat{\beta})'(Y-X\hat{\beta})/(n-m)$ is the UMVUE of σ^2. [3x5=15]

p.t.o.

4.(a) Describe either the method of scoring or EM algorithm for Maximum Likelihood Method of estimating several parameters.

(b) Use either method of scoring or EM algorithm to compute Maximum Likelihood estimates or the gene-frequencies p, q and r of O, A, and B respectively using the following observed phenotypic frequencies of blood groups.

phenotype	frequency
O	180
A	190
B	62
AB	18

[6+19=25

5. Let X have the density

$$f(x \mid \theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2} \quad , -\infty < x < \infty \ , -\infty < \theta < \infty$$

(a) Show that this family has MLR

(b) Based on one observation X, find the most powerful size α test for testing $H_0 : \theta = 0$ against $H_1 : \theta = 1$. For $\alpha = 0.2$ find the probability of type II error.

(c) Show that the test in part (b) is UMP level α test for testing $H : \theta \leq 0$ against $K : \theta > 0$. [4+6+2+6=18]

6.(a) Let $f(x)$ be a unimodal density on positive real numbers and let $c.f(x)\,x^2$ be also a unimodel density for some $c > 0$. Let x^* be the mode of the latter distribution. Let $0 < \alpha < 1$ be a given number. Let a and b satisfy

(i) $\int_a^b f(x)\,dx = 1 - \alpha$

(ii) $f(a).a^2 = f(b).b^2$

(iii) $a < x^* \leq b$.

Show that $\frac{1}{a} - \frac{1}{b} \leq \frac{1}{u} - \frac{1}{v}$ for all u, v such that $0 \leq u \leq v$ and $\int_u^v f(x)dx = 1-\alpha$.

contd. 3/-

6.(b) Let $X_1, \cdots, X_n$ be iid $N(0,\sigma^2)$. Obtain the shortest expected length confidence interval, with confidence coefficient $(1-\alpha)$, of the form

$$\left\{ \sigma^2 : \frac{\Sigma X_i^2}{a} \leq \sigma^2 \leq \frac{\Sigma x_i^2}{b} \right\} \text{ for}$$

σ^2. [9+7=16]

7. In 1000 tosses of a coin 580 heads and 420 tails appear. On it reasonable to assume that the coin is fair? Justify your answer. [6]

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1991-92

Probability Theory II
SEMESTRAL-II EXAMINATION

Date : 4.5.1992 Maximum Marks : 120 Time : $3\frac{1}{2}$ Hours.

Note : This paper carries questions worth a total of 140 points.
Answer as much as you can.
The maximum you can score is 120 points.

1. Give complete statements of the following (no proof needed) :
   (a) Monotone Class Theorem
   (b) Fubini's Theorem
   (c) Kolmogorov's maximal inequality
   (d) Lévy Continuity Theorem.

[4x4=16]

2. A subset B of $\mathbb{R}$ is called symmetric if whenever x is in B, -x also belongs to B.
   Denoting by $\mathcal{C}$ the collection of all the symmetric borel subsets of $\mathbb{R}$, show that
   (a) $\mathcal{C}$ is a σ-field on $\mathbb{R}$
   (b) a function $f : \mathbb{R} \to \mathbb{R}$ is $\mathcal{C}$-measurable if and only if f is borel measurabl. and an even function.

[10+10=20]

3. Let P be a probability on $(\mathbb{R}, \mathcal{B})$.
   (a) Show that for any $B \in \mathcal{B}$, the function $x \mapsto P(B-x)$ is a borel measurable function.
   (Hint : For $B = (-\infty, a]$, $x \mapsto P(B-x)$ is a monotone function).
   (b) Show that $Q(B) = \int_{\mathbb{R}} P(B-x)\, dP(x)$, for $B \in \mathcal{B}$, defines a probability on .... $(\mathbb{R}, \mathcal{B})$.
   (c) Show that if X and Y are independent random variables both having distribution P, then Q is the distribution of the random variable $Z = X+Y$.

[10+10+10=30]

4. Let X be a standard normal random variable. Consider the function $g(t) = E(\cos tX)$, for $t \in \mathbb{R}$.
   (a) Show that g is an everywhere differentiable function and tha it satisfies the differential equation $g'(t) = -tg(t)$ for all $t \in \mathbb{R}$.
   (b) Deduce that the characteristic function of X is $e^{-\frac{1}{2}t^2}$.

[10+6=16]

p.t.o.

5.(a) Let X_n, $n \geq 1$ and X be random variables such that $Ef(X_n) \to Ef(X)$ (as $n \to \infty$) for all bounded continuous $f : \mathbb{R} \to \mathbb{R}$. Show that X_n converge in distribution to X.

(b) If X_n, $n \geq 1$ are random variables such that X_n takes the n values $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1$ each with probability $\frac{1}{n}$, then show that X_n converge in distribution.

[10+8=18]

6. Let X_n, $n \geq 1$ be a sequence of independent random variables, and let A be the event $\{\sup_n X_n < \infty\}$.

(a) Show that $P(A)$ is either 0 or 1.

(b) Show that $P(A)$ equals 1 if and only if for some α, $\sum_{n=1}^{\infty} P(X_n > \alpha) < \infty$.

[10+10=20]

7.(a) Let X_n, $n \geq 1$ be a sequence of random variables with $P(|X_n| \leq 1) = 1$ for all n. If $X_n \xrightarrow{P} X$, then show that $P(|X| \leq 1) = 1$ and $E|X_n - X| \to 0$.

(b) Show that if X_n, $n \geq 1$ and X are random variables such that $\sum_{n=1}^{\infty} E|X_n - X| < \infty$, then $X_n \xrightarrow{a.s.} X$.

[10+10=20]

1991-92 | 548

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1991-92
SEMESTRAL II EXAMINATION
LINEAR STATISTICAL MODELS AND LARGE SAMPLE STATISTICAL METHODS

Date: 30.4.92 | Maximum Marks: 100 | Time: 3 Hours

1. Consider observations Y_1, Y_2, Y_3 and Y_4 satisfying the linear model:

$Y_1 = \beta_1+2\beta_2+e_1$ $\qquad$ $Y_3 = 3\beta_1+\beta_3+2\beta_4+e_3$

$Y_2 = \beta_1 - \beta_2 + e_2$ $\qquad$ $Y_4 = \beta_3+2\beta_4+e_4$

Here e_1, e_2, e_3 and e_4 are i.i.d random variables each with zero mean and variance σ^2. Which of the following linear parametric functions is/are ESTIMABLE? Justify your answer.

(a) $\beta_1+\beta_2$ $\qquad$ (b) $\beta_3-\beta_4$ $\qquad$ (c) $\beta_2+2\beta_3+4\beta_4$

For each of the estimable parametric function(s) above, obtain the expression for BLUE in terms of Y_1, Y_2, Y_3 and Y_4 and compute the variance of the BLUE in terms of σ^2. [25 points]

2. Let $X_1, X_2, \ldots, X_n$ be i.i.d observations with a common density $f(x) = 2\{\Gamma(1/4)\}^{-1} \exp\{-(x-\theta)^4\}$, $-\infty < x, \theta < \infty$. Let $\hat{\theta}_{n,1}$ be the sample median and $\hat{\theta}_{n,2}$ be the maximum likelihood estimate for θ based on these observations. Determine (with adequate justification) the limiting distributions of of $\sqrt{n}(\hat{\theta}_{n,1}-\theta)$ and $\sqrt{n}(\hat{\theta}_{n,2}-\theta)$. Then compute the asymptotic relative efficiency (ARE) of the sample median compared to the maximum likelihood estimate. (Assume that $\hat{\theta}_{n,2}$ is a consistent root of the maximum likelihood equation). [25 points]

3. Let Y_{ij}'s be observations satisfying the linear model described below.

$Y_{1j} = \beta_1+\beta_2+e_{1j}$ $\qquad 1 \leq j \leq m$

$Y_{2j} = \beta_2+\beta_3+e_{2j}$ $\qquad 1 \leq j \leq n$

$Y_{3j} = \beta_3+\beta_1+e_{3j}$ $\qquad 1 \leq j \leq p$

Here m,n,p are three positive integers and the e_{ij}'s are i.i.d $N(0,\sigma^2)$ random variables. Derive explicitly the form of the F-statistic (in terms of the observations Y_{ij}'s) for testing $H_0 : \beta_2=\beta_3$ against $H_A: \beta_2 \neq \beta_3$. [20 points]

p.t.o.

4. Let $X_1, X_2, \ldots, X_n$ be i.i.d observations with a common density $f(x) = \{B(\theta+1, 2-\theta)\}^{-1} x^{\theta} (1-x)^{1-\theta}$, $0 \leq x, \theta \leq 1$. Suggest a $\sqrt{n}$-consistent estimate $\hat{\theta}_n$ of θ based on the data such that $\sqrt{n}(\hat{\theta}_n - \theta)$ is asymptotically normal. Determine explicitly the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ with sufficient justification. [15 points]

5. Suppose that Y_i's are observations satisfying

$Y_i = \beta_0 + \beta_1 X_i + e_i$ $(i = 1, 2, \ldots, n, \ldots)$, where $X_i = i$ and $e_1, e_2, \ldots, e_n, \ldots$ are i.i.d random variables satisfying $E(e_1) = 0$, $Var(e_1) = \sigma^2 > 0$. Let $\hat{\beta}_{n,1}$ be the usual least squares estimate of the slope parameter β_1 based on (Y_1, X_1), $(Y_2, X_2), \ldots, (Y_n, X_n)$. Is $\hat{\beta}_{n,1}$ a consistent estimate for β ? Justify your answer. [15 points]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR:1991-92
SEMESTRAL-II EXAMINATION
DEMOGRAPHY

Date:28.4.92 Maximum Marks:100 Time: 3 Hours

Note: Answer any <u>FIVE</u> questions

1. What are the main sources of demographic data? Discuss the pitfalls of birth and death registrations in India. (5+15)=[20]

2.(a) Describe the various errors and biases in the age returns from a census. Describe Myer's procedure for detecting digit preferences.

(b) Describe a method to evaluate the coverage of vital registration and estimate the under count. (4+8+8)=[20]

3. Define rate of natural increase. Derive the exponential model for population growth. Prove that the population will quadruple in size at time 1.38/r , where r is the growth rate of the population. What is the balancing equation of the population? Interpret it. (4+8+4+2+2)=[20]

4. Define crude and standardiesed death rates. In what way are the standardised rates superior? Explain briefly the differences between the direct and indirect methods of standardising death rates. (6+6+8)=[20]

5. Define infant mortality rate. How infant mortality rate differs from infant death rate? Describe the different methods for computing the infant mortality rate, indicating their merits and limitations. (2+6+12)=[20]

6. What do you mean by a life table? Distinguish between 'complete' and 'abridged' life tables. Describe a method for constructing an abridged life table. (4+6+10)=[20]

7.(a) For a certain life table

$$l_x = x^2-50x - 5000$$

i) What is the ultimate age in the table?

ii) Find μ_x, q_x and $_{10}l_{20}$.

(b) A maternity hospital delivers 10 new born babies per week; 30% of them leave hospital within a week, 10% of the remaining one week old babies leave before they are two weeks old, 20% of the remainder leave before they are three weeks old, 40% of the remainder before they are four weeks old and the rest leave before they are six weeks old. How many cots are needed? (16+4)=[20]

8. What do you mean by popul tion projection? Explain the component method of population projection. (8+12)=[20]

9. What are the different measures of fertility commonly used? Discuss in details the merits and demerits of the crude birth rate and general fertility rates. (12+4+4)=[20]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR:1991-92
SEMESTRAL-II EXAMINATION
APPLIED STOCHASTIC PROCESSES

Date: 27.4.92 Maximum Marks:100 Time: $3\frac{1}{2}$ Hours

Note: Answer all the questions

1. In the questions below, $\{X_n : n \geq 0\}$ is a Markov chain with stationary transition probabilities.

(a) Suppose the transition probability matrix $P=((p_{ij}))_{i,j \geq 0}$ of $\{X_n : n \geq 0\}$ is of the form $p_{i,i+1}=p_i, p_{i,0}=1-p_i$, $i=0,1,2,\ldots$, $0 < p_i < 1$, $i \geq 0$. Show that the chain is irreducible, and that it is recurrent iff $\sum_{i=0}^{\infty}(1-p_i)=\infty$.

(b) Suppose the transition probability matrix $P=((p_{ij}))_{i,j \geq 0}$ of $\{X_n : n \geq 0\}$ is idempotent (i.e. $P^2=P$). Suppose, moreover, that $\{X_n : n \geq 0\}$ is irreducible. Prove that $p_{ij}=p_{jj}$ for all i and j and that $\{X_n : n \geq 0\}$ is aperiodic.

(c) (i) Consider a regular $2r+1$ polygon consisting of vertices $V_1, \ldots, V_{2r+1}$. Suppose that at each point V_k there is a nonnegative mass $W_k^{(0)}$ where $W_1^{(0)}+\ldots+W_{2r+1}^{(0)}=1$. Obtain new masses $W_1^{(1)}, \ldots, W_{2r+1}^{(1)}$ by replacing the old mass at k by the arithmetic mean of neighbouring masses, i.e.,

$W_1^{(1)}=\frac{1}{2}(W_2^{(0)}+W_{2r+1}^{(0)})$, $W_k^{(1)}=\frac{1}{2}(W_{k-1}^{(0)}+W_{k+1}^{(0)})$, $2 \leq k \leq 2r$,

$W_{2r+1}^{(1)}=\frac{1}{2}(W_{2r}^{(0)}+W_1^{(0)})$.

Do this transformation n times. Determine $\lim_{n \to \infty} W_k^{(n)}$, if it exists.

(ii) Suppose instead we consider a regular $2r$ polygon. Can we talk of $\lim_{n \to \infty} W_k^{(n)}$ for an arbitrary initial choice of $W_1^{(0)}, \ldots, W_{2r}^{(0)}$? Give reasons.

1. (d) Suppose the transition probability matrix $P=((p_{ij}))_{i,j\in S}$, $S=\{0,1,\ldots,N\}$, is of the form $p_{00}=p_{NN}=1$, $p_{i,i-1}=\mu_i, p_{i,i+1}=\lambda_i$, $p_{ii}=1-\lambda_i-\mu_i$, $i=1,\ldots,N-1$, where $\mu_i>0$, $\lambda_i>0$, $\lambda_i+\mu_i<1$, $i=1,\ldots,N-1$. Suppose that the initial state of the process is k. Determine the absorption probability at 0. [10+8+12+10=40]

2. (a) Show that a Poison process is indeed a Markov process.

   (b) Consider a Poisson process $\{X(t):t>0\}$; let S_i be the point of time at which the i-th event occurs. Obtain the conditional distribution of $S_1,\ldots,S_n$ given $X(t) = n$. Hence, obtain $E(S_k \mid X(t)=n)$, $k \leq n$.

   (c) Consider a renewal process $N(t):t > 0$ with inter arrival times $T_1,T_2,\ldots$, where $E(T_1)=\mu$. Prove that

$$\frac{E(N(t))}{t} = \frac{1}{\mu} + o(1) \quad \text{as } t\to\infty.$$

   (d) Cars arrive at a gate. Each car is of random length L having exponential distribution with mean $\frac{1}{\lambda}$ $(\lambda > 0)$. The first car arrives and parks against the gate. Each succeeding car parks behind the previous car at a distance that is random according to a uniform distribution on [0,1]. Consider the number of cars N_x that are lined up within a total distance x of the gate. Determine $\lim_{n\to\infty} \frac{E(N_x)}{x}$. [6+12+12+10=40]

3. Answer anyone of (a) and (b):

   (a) In a telephone exchange with N available channels, a connection (or conversation) is realised if the incoming call finds an idle channel; as soon as the conversation is finished, the channel being utilized becomes immediately available for a new call. An incoming call is lost if all channels are busy. Suppose the probability of an incoming call during the interval $(t,t+h)$ is $\lambda h+o(h)$, and the probability of a conversation ending during the interval $(t,t+h)$ is $\mu h+o(h)$. Let $X(t)$ be the number of channels occupied at time t. Derive appropriate differential equations for the probabilities $p_k(t)=P(X(t)= k)$. Also, obtain explicit expressions for $p_k(t)$'s whenever N is infinite.

3.(b) A person who owns r umbrellas distributes them between home and office according to the following routine. If it is raining upon departure from either place, an event that has probability p, say, then an umbrella is carried to the other location if available at the location of departure. If it is not raining, then an umbrella is not carried. Let X_n denote the number of available umbrellas at whatever place the person happens to be departing on the nth trip.

(i) Determine the transition probability matrix and the stationary distribution.

(ii) Let $0 < \alpha < 1$. How many umbrellas should the person own so that the probability of getting wet under the stationary distribution is at most α against a climate(p)? What number works against all possible climates for the probability α? [20]

1991-92	544(b)

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1991-92
SEMESTRAL-I BACKPAPER EXAMINATION

Comp. Techniques and Programming

Date: 2.1.1992 Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

Note: Answer any part of any questions.
Maximum you can score is 100.

1.(a) Given a real valued function $f(x)$ and $(n+1)$ distinct point $x_0, x_1, \ldots x_n$, show that there exists exactly one polynomial of degree $\leq n$ which interpolates $f(x)$ at $x_0, x_1, \ldots x_n$.

(b) Prove or disprove the following statement:

Addition of distinct point will always increase the degree of the interpolating polynomial.

(10+5) = [15]

2.(a) Convert base-10 number 123.625 to binary form, stating clearly the algorithm(s) used. From the binary form convert to octal and hexadecimal form.

(b) Assuming a computer with 4 decimal place mantissa, add the following numbers, first in ascending order (from smallest to highest) and then in descending order (from highest to smallest). In doing so, round off the partial sums.

Compare your results with the correct sum 0.107101023×10^5.

The numbers

$0.1580.10^0$ $0.2653.10^0$

$0.2581.10^1$ $0.4288.10^1$

$0.6266.10^2$ $0.7555.10^2$

$0.7889.10^3$ $0.7767.10^3$

$0.8999.10^4$

(14+11) = [25]

3. Reduce the following matrix to tridiagonal form using. Householder transformation

$$\begin{matrix} 1 & -2 & 4 & 4 \\ -2 & 3 & 6 & 6 \\ 4 & 6 & 6 & 3 \\ 4 & 6 & 3 & 6 \end{matrix}$$

[15]

p.t.o.

4.(a) Let A be a matrix of order n such that

$$a_{ii} = \alpha_i$$
$$a_{i,i+1} = a_{i+1,i} = \beta_{i+1} \neq 0, \quad i = 1, 2, \ldots n-1$$
$$a_{ij} = 0, \quad i = 1, 2, \ldots n; \quad j \neq i-1, i, i+1 .$$

Let $p_r(\lambda)$ be the leading principal minor of order r of $(A - \lambda I)$.

(i) Show that $p_r(\lambda) = (\alpha_r - \lambda)\, p_{r-1}(\lambda) - \beta_r^2\, p_{r-2}(\lambda)$

$$r = 2, \ldots n .$$

where $p_0(\lambda) = 1$.

(ii) Prove that

$$x_r = (-1)^{r-1} . p_{r-1}(\lambda) / \beta_2 . \beta_3 \ldots \beta_r$$

$$r = 2, 3 \ldots n$$

where $x = (x_1, x_2, \ldots x_n)'$ is an eigenvector corresponding to eigen value λ of A.

(b) Find the characteristic polynomial of the following matrix:

1	2	0	0
2	1	3	0
0	3	1	4
0	0	4	1

(5+5+5) = [15]

5.(a) Write short notes on condition number of a function and stability of a computational process.

(b) Evaluate

$$f(x) = \sqrt{x+1} - \sqrt{x}$$

at x = 12345

with six decimal arithmetic with round off using the following computational steps:

$$x_0 = 12345$$
$$x_1 = x_0 + 1$$
$$x_2 = \sqrt{x_1}$$
$$x_3 = \sqrt{x_0}$$
$$x_4 = x_2 - x_3$$

contd..... 3/-

compare your results with the actual value

$$(12345) = 0.004500032626 .$$

Is the process described above is stable ? If not, explain why not. And in this case, can you suggest any other stable process to compute $f(x)$?

(8+12) = [20]

6. Define scalar product of two functions. What do you understand by orthogonal functions ? Explain with examples.

Given $P_0(x)$, $P_1(x)$, $P_2(x)$, ... a sequence of orthogonal polynomials.

Show that $P_k(x)$ is orthogonal to any polynomial of degree $< k$.

[10]

7. Given a function f(x) and two real numbers a and b, write a FORTRAN program to compute the real root, if any, lying between a and b using method of bisection. Assume that the functional value f(x) at a point x is given by a FORTRAN Function subprogram FUN(X). State clearly, all the assumptions you are making. Give a clear algorithm for the computational steps.

[15]

:bcc:

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) I Year: 1991-92

SEMESTRAL-I BACKPAPER EXAMINATION

Comp. Techniques and Programming I

Date: 2.1.1992 Maximum Marks: 100 Time: 3 hours

Note: [illegible] up.

Testbook on Programming with FORTRAN may be used. Pocket calculators are allowed.

GROUP - A

Note: Answer question no.1 and any other <u>TWO</u> from the rest.

1. Assume that a and b are exact positive numbers. Derive error bounds on the relative errors of $u = 3(ab)$ and $v = ((a+a) + a)b$. Which of the two will give better accuracy ? Illustrate with $a = 0.4299$, $b = 0.6824$, using 4-digit floating-pt arithmetic.

(2+2+2+4) = [10]

2. Describe, in the form of a flow-chart, the method known as "Regula Falsi".

The equation $x^6 = x^4 + x^3 + 1$ has a root between 1 and 2. Find this root correct to 3 decimal places by the method of Regula Falsi.

(8+12) = [20]

3. The equation $x^3 - 2 = 0$ can be rewritten as (a) $x = x^3 + x - 2$ and (b) $x = (2 + 5x - x^3)/5$. Apply the algorithm $x_{n+1} = g(x_n)$ to both (a) and (b) with $x_0 = 1.2$. Explain the results you obtain.

(6+6+8) = [20]

4. Derive the secant method and obtain its order of convergence.

(6+14) = [20]

p.t.o.

GROUP - B

Note: Answer any THREE questions.

1.(a) Write a program to evaluate a polynomial

$$a_n x^n + a_{n-1} x^{n-1} + \dots a_1 x + a_0$$

for different values of x.

You are not allowed to use exponentiation.

(b) Write a program to check whether or not the row sums, column sums and the two diagonal sums of a square matrix are same.

(8+8) = [16]

2. Given N points (x_i, y_i), i = 1, 2, ... N, write a program to list all the triangles which can be formed by using the N points.

[16]

3. Given an integer less than 10000, write a program to express the number in words. (e.g., if the given number is 215 the output should be "TWO HUNDRED FIFTEEN")

[16]

4. Write a program to find the product of two binary integers. Clearly indicate any assumptions made by you.

[16]

5. Write a program to print values of ${}^nC_r\, p^r (1-p)^{n-r}$ in a tabulor form, for different values of n, r, p, where

$$5 \leq n \leq 10, \quad 0 \leq r \leq n, \quad p = .1(.1).9.$$

[16]

Neatness [2]

:bcc:

1991-92	574

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream): 1991-92
SEMESTRAL - I EXAMINATION

Economic Statistics

Date: 28.11.1991 Maximum Marks: 100 Time: 3 hours

Note: Answer <u>ALL</u> the questions.

1. Describe the ratio-to-moving-average method of determining moving seasonal indices from a series of monthly figures where the seasonal pattern has been changing gradually. (Assume the multiplicative model for the time series).

[25]

2. Mention different types of equations used for fitting mathematical trends to observed time series.

How does one choose the appropriate type of equation in fitting trend to any particular time series.

[35]

3. The following is a set of prices (in suitable units) for the given commodities:

	Prices					
Commodity	1960	1961	1962	1963	1964	1965
A	60	75	90			
B			80	60	90	96
C	20	30				
D		15	15	20	20	25
E	40	40	50	54	52	58

Suppose that commodities A and B fall into one group, commodities C and D into a second, and commodity E represents a third. It is found that appropriate group weights are 25, 60 and 15 respectively. Compute a weighted Index with 1960 as base for all the years from 1961 to 1965.

[30]

4. Practical Note Book. [10]

:bcc:

INDIAN STATISTICAL INSTITUTE

M.Stat. I Year (M-stream): 1991-92

SEMESTRAL - I EXAMINATION

Linear Algebra

Date: 27.11.1991 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

Note: Answer as many questions as you can. The whole paper carries 59 marks but the maximum you can score is 50. Use separate answer booklet for each group.

Group A

1.(a) State and prove the modular law about the dimension of the sum of two subspaces. [12]

(b) If S and T are 3-dimensional subspaces of $\mathbb{R}^4$, determine all the possible values for $d(S \cap T)$. Give complete proof. [5]

2.(a) If $ABC = ABD$ and $\rho(AB) = \rho(B)$, prove that $BC = BD$. [5]

(b) Prove that $\rho(BB^*) = \rho(B)$ for any complex matrix B. [5]

(c) Show that if G, H are two g-inverses of B^*B, where B is a complex matrix, then $BGB^* = BHB^*$. [3]

(d) Show that if B is an $m \times n$ complex matrix and $x \in \mathbb{C}^m$, then BGB^*x is the orthogonal projection of x into $\mathcal{C}(B)$ where G is a g-inverse of B^*B. [8]

3.(a) Let A be a real symmetric matrix. Assuming the result that the eigen values of A are real, prove that A is orthogonally similar to a diagonal matrix. [10]

(b) Show that every real n.n.d. matrix can be written as BB^T for some real matrix B. [4]

(c) Show that if the ith diagonal element of an n.n.d. matrix is 0, the ith row is null. [3]

4. Assignment. [4]

p.t.o.

Group B

Correlation and Regression

Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

Note. Answer ALL the questions.

1.(a) Define the correlation ratio e_{yx}.

(b) Show that

$$r^2 \le e^2_{yx} \le 1.$$

Interpret the cases: (i) $r^2 = e^2_{yx}$

and (ii) $r^2 < e^2_{yx} = 1$.

(2+4+4) = [10]

2.(a) Show that

$$r_{12.3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1 - r^2_{13}}\,\sqrt{1 - r^2_{23}}}.$$

(b) Consider 5 variables $x_1, x_2, \ldots, x_5$. Suppose that $r_{ij} = r$ whenever $i \neq j$. Compute $R^2_{1(2345)}$ and $r_{12.345}$.

(6+6+8) = [20]

3.(a) It is generally true that inclusion of more explanatory variables increases the value of the multiple correlation coefficient. Examine the conditions when raising the number of explanatory variables from two to three does not alter the value of the multiple correlation coefficient.

(b) For the following data set, work out the values of the linear regression coefficients and the multiple correlation coefficient of Y on X_1, X_2, X_3:

	y	X_1	X_2	X_3
Means	10	0.8	1.3	2.1
Dispersion Matrix	25	6	-4	3
		10	-2	4
			10	-3
				10

(5+15) = [20]

:bcc:

1991-92 | 544.

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1991-92
SEMESTRAL - I EXAMINATION

Computation Techniques and Programming

Date: 25.11.1991 Maximum Marks: 100 Time: 3 hours

Note: Question No.6 is compulsory. You can answer any part of any other questions.

1.(a) Let $f(x_0)$, $f(x_1)$, ..., $f(x_n)$ be the functional values corresponding to values x_0, x_1, ..., x_n of its arguments.

(i) Show that the divided differences of f(x) are symmetric function of its arguments.

(ii) $\frac{d}{dx} f[x_0, x_1 \dots x_n, x] = f[x_0, x_1, \dots x_n, x, x]$.

(b) Prove or disprove the following statement:

Addition of distinct point will always increase the degree of the interpolating polynomial.

(8+7+5) = [20]

2.(a) What is meant by the term "cancellation" ? Will cancellation always occur in addition and subtraction operation ? Why ? Why is cancellation not a problem in multiplication and division ?

(b) Let a, b, c are floating point values. Analyze the error in computing fl (a + b + c) and discuss your results. Will your results be true if cancellation occur ?

(c) Give an example which shows that

$$fl[fl(a+b) + fl(c)] \neq fl[fl(a) + fl(b+c)]$$

where a, b, c are real numbers.

(8+12+5) = [25]

3.(a) Show how the following are represented internally in VAX / VMS system :

(i) 123.624 (ii) -123.625 (iii) -1024

(iv) 127 (v) 'ABCD'

contd..... 2/-

(b) Write short notes on

(i) Compiler (ii) Interpreter (iii) Assembler
(iv) Machine Language (v) Processor.

(15+10) = [25]

4. Given a vector $w = (w_1, w_2, \ldots w_n)'$, find a vector u and a scalar β such that H.w has

(i) zeros in positions t+1, t+2, ... n

(ii) $w_1, w_2 \ldots w_{t-1}$ in positions 1, 2, ... t-1

(iii) $\pm$ S in tth position, where $S^2 = \sum_{i=t}^{n} w_i^2$

Where $H = 1 - \frac{uu'}{\beta}$.

Show that H is orthogonal under either choices of S.

[10]

5. Using simplex algorithm solve the following problem:

Minimize $2x_1 + x_2 + 3x_3 - x_4$

Subject to $2x_1 + 3x_2 - x_3 + x_4 \geq 2$

$4x_1 + 6x_2 - 2x_3 + 2x_4 = 4$

and $x_1, x_2, x_3, x_4 \geq 0$.

[15]

6. Assignments. [20]

:bcc:

1991-92 | 534

INDIAN STATISTICAL INSTITUTE

M.Stat. I Year (M-stream) : 1991-92

SEMESTRAL - I EXAMINATION

Theory and Methods of Statistics I

Date: 22.11.1991 Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

Note: Answer any FOUR questions. Marks allotted to a question are indicated in brackets [ ] at the end. Practical Records and Assignments carry 20 marks. Submit all your records and assignments by 30.11.1991.

1.(a) State and prove a result on the percentage of observations lying within $\bar{x} \pm ks$, where $\bar{x}$ is the sample mean, s is the sample standard deviation and k is a positive constant. Explain the significance of this result and discuss one use.

(7+3) = [10]

(b) Define moments and Pearson's coefficients b_1 and b_2. Prove that $b_2 \geq b_1$ and $b_2 - b_1 - 1 \geq 0$. Discuss one use of these inequalities.

(2+6+2) = [10]

2.(a) Let F(t) denote the probability that a structure fails by time t, and let $\lambda(t)\Delta(t) + O(\Delta t)$ be the conditional probability of failure in a small interval of time Δt, given that it has survived upto time t. Show that F(t) is then necessarily of the form given by:

$$1 - F(t) \propto \exp[-\int \lambda(t)\,dt].$$

What is the model called when $\lambda(t) = \lambda$, independent of t ?

Obtain the pdf and cdf of this model, and show that it "lacks memory" in the sense that

$$P[T \leq t_o + t \mid T > t_o] = P[T \leq t],$$

independent of t_o. Explain the significance of this result.

(6+8) = [14]

(b) If this distribution is truncated from above at c, then show that the mean of the truncated distribution is given by [Mean of the untruncated distribution] $- c\left[\frac{e^{-\lambda c}}{1 - e^{-\lambda c}}\right]$.

[6]

p.t.o.

3.(a) Let $X_1, X_2, \ldots, X_k$ follow a k-class multinomial distribution with the parameters n and $p_1, p_2, \ldots, p_k$ where $p_i > 0$, $\sum_{i=1}^{k} p_i = 1$. Obtain the conditional distribution of

$X_{i_1}, X_{i_2}, \ldots, X_{i_r}$, given $X_{j_1} = x_{j_1}, X_{j_2} = x_{j_2}, \ldots, X_{j_s} = x_{j_s}$

where $1 \leq i_1 \neq i_2 \neq \ldots \neq i_r \neq j_1 \neq j_2 \neq \ldots \neq j_s \leq k$, r and s being positive integers such that $r+s \leq k$.

(b) Let time-to-failure of a system follow an exponential model with parameter α. Define X = No. of such failures in a given time interval (0, T). Obtain the distribution of X and all its first four moments.

(8+12) = [20]

4.(a) Let X have the beta distribution with parameters m and n. Prove that

$$P[X \leq p] = P[Y \geq m],$$

where Y has the binomial distribution with parameters m+n-1 and p ; $0 < p < 1$.

(b) Suppose the pmf f(x) satisfies the relation.

$$\frac{f(x+1)}{f(x)} = \frac{\alpha + \beta x}{x+1} ; \quad x = 0, 1, 2, \ldots ,$$

where α and β are constants. Find the mode of X, E(X), V(X) and $E|X - E(X)|$ in terms of α and β.

(8+12) = [20]

5.(a) Describe stratified random sampling. Show that for stratified simple random sampling an optimum allocation for n_h is given by

$$n_h \propto N_h S_h / \sqrt{c_h} ,$$

where the symbols have their usual significance. Obtain also the expression for the constant of proportionality in terms of N_h, S_h, c_h and given V or C.

(b) Prove that $V_{opt} \leq V_{prop} \leq V_{ran}$, and interpret the result.

(10+10) = [20]

p.t.o.

6.(a) Define the maximum likelihood estimator (MLE) and state some of its important properties known to you.

Suppose X has uniform distribution in the interval $[0, \alpha]$. Obtain MLE of α, and show that it is a biased but consistent estimator of α.

(b) Obtain a $100(1 - \alpha)\%$ confidence interval for α.

(12+8) = [20]

:bcc:

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1991-92
SEMESTRAL - I EXAMINATION

Mathematical Analysis I M

Date: 20.11.1991 Maximum Marks: 100 Time: 3 hours

Note: Answer FIVE questions.

1. For each of the following statements, either give a proof or provide a counter example:

(a) In a metric space if $\{A_n\}$ is a sequence of closed sets such that $A_{n+1} \subseteq A_n$, $n = 1, 2, \ldots$, then $\bigcap_{n=1}^{\infty} A_n$ is nonempty.

(b) In a metric space (X,d) if $f : X \to X$ is a function such that $d(f(x), f(y)) < \frac{1}{2} d(x,y)$ for all $x, y \in X$, then for any $x \in X$, the sequence $x, f(x), f(f(x)), \ldots$ is Cauchy.

(c) If f is a continuous function: $(X,d) \to (Y, \rho)$ then $f^{-1}(C)$ is closed in X whenever C is closed in Y.

(d) If f is a continuous function: $(X,d) \to (Y, \rho)$ then $f(U)$ is open in Y whenever U is open in X.

(5 x 4) = [20]

2.(a) Show that an absolutely convergent series Σa_n of complex nos. is convergent.

(b) Prove that if $\Sigma a_n z^n$ is convergent at $z = z_o$, then $\Sigma a_n z^n$ converges uniformly on $|z| < |z_o|$.

(c) Prove that $f(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n^3}$ is a differentiable function on $\mathbb{R}$.

(6+6+8) = [20]

3.(a) Define f on $\mathbb{R}^3$ by

$$f(x, y_1, y_2) = x^2 y_1 + e^x + y_2 .$$

Show that there exists a differentiable function g in some neighbourhood V of $(1, -1)$ in $\mathbb{R}^2$ such that $g(1, -1) = 0$ and

contd..... 2/-

$$f(g(y_1,y_2), y_1, y_2) = 0$$

for all $(y_1,y_2) \in V$.

(b) Find $D_1 g(1,-1)$.

(c) What are the values of k for which g is a C^k - function ?

(7+7+6) = [20]

4. For a piecewise differentiable path γ lying in an open set $\Omega \subseteq \mathbb{R}^2$, define

$$\int_\gamma P dx + Q dy$$

where P and Q are continuous functions on Ω, in the usual way.

(a) Show that if $\int_\gamma P dx + Q dy$ depend only on the endpoints of γ, then there exists f on Ω s.t. $\frac{\partial f}{\partial x} = P$ and $\frac{\partial f}{\partial y} = Q$ on Ω [assume Ω to be path-connected]

(b) Show that if g is an analytic function on the open set $\Omega \subseteq \mathbb{C}$ and if

$$\int_\gamma g(z)\, dz$$

depends on on the end points of γ, then there exists an analytic function G on Ω such that $F'(z) = g(z)$ for all $z \in \Omega$.

(10+10) = [20]

5.(a) Let Ω be an open subset of $\mathbb{R}^2$ and $(u,v) : \Omega \to \mathbb{R}^2$ be differentiable (in the real sense). Show that if (u,v) satisfies the Cauchy Riemann equations, then $f = u + iv$ is an analytic function on Ω (regarded as a subset of $\mathbb{C}$).

(b) Let Ω be an open subset of $\mathbb{C}$ and f an analytic function on Ω such that $f'(z) \neq 0$ for each $z \in \Omega$. Prove that $f(\Omega)$ is open in $\mathbb{C}$.

(10+10) = [20]

6.(a) Prove that if f is an analytic function on $\Omega \subseteq \mathbb{C}$ and, for some $k \geq 0$, $f^{(k)}(z_0) \neq 0$, then there exists a neighbourhood V of z_0 such that $f(z) \neq 0$ if $z \in V$ and $z \neq z_0$.

(b) Compute $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(x^2+4)^2}$.

(10+10) = [20]

:bcc:

1991-92 | 514

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1991-92
SEMESTRAL - I EXAMINATION

Probability Theory IM

Date: 18.11.1991 Maximum Marks: 100 Time: 3 hours

Note: Answer any THREE questions from Group A and any TWO from Group B.

GROUP A (3x16) = [48]

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ variables. Find the joint density of

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

and show that they are independent.

Also show that $X_1 - \bar{X}$ and $\bar{X}$ are independent.

2. (i) State and prove Chebyshev's inequality.

(ii) If X follows Poisson distribution with parameter λ, show that

$$P(X \leq \frac{\lambda}{2}) \leq \frac{4}{\lambda} \quad \text{and} \quad P(X \geq 2\lambda) \leq \frac{1}{\lambda}.$$

(iii) If X is a positive random variable such that both $E(X)$ and $E(\frac{1}{X})$ exist, show that $E(\frac{1}{X}) \geq \frac{1}{E(X)}$.

3. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with common density f and d.f. F and $X_{(1)} \leq \cdots \leq X_{(n)}$ be the corresponding order statistics. Assume that f is continuous and equal to F'.

(i) Find the densities of $X_{(1)}$ and $X_{(n)}$.

(ii) Take $f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0. \end{cases}$

Find the densities of $X_{(1)}$ and $X_{(n)}$. Find the joint density of $X_{(1)}, X_{(2)} - X_{(1)}, \ldots, X_{(n)} - X_{(n-1)}$ and show that they are independent. Find $E(X_{(k)})$, $K = 1, 2, \ldots, n$.

p.t.o.

4. (i) Define a random variable, density function of a continuous random variable and conditional expectation of a ramdom variable given another random variable.

(ii) Let X and Y be two random variables such that $E(Y|X = x) = m(x)$ exists for all x. Show that $E(Y - g(X))^2 \geq E(Y - m(X))^2$ for any function g.

(iii) Suppose that Y is a random variable with finite mean θ. Let X be another random variable and $\phi(X) = E(Y|X)$ be well defined.

Show that $E\,\phi(X) = \theta$

and $V[\phi(X)] \leq V(Y)$.

5. (i) Let Y_i's be independent and $Y_i \sim$ Gamma (α_i, λ), $i = 1, 2, \ldots, K+1$. Show that the random variables

$$X_i = \frac{Y_i}{\sum_{i=1}^{K+1} Y_i}, \quad i = 1, 2, \ldots, K$$

jointly follow the Dirichlet distribution.

(ii) Suppose that $(X_1, X_2, \ldots, X_{10}) \sim D_{10}(\alpha_1, \ldots, \alpha_{10}; \alpha_{11})$. Set $Z_1 = X_1 + X_2$, $Z_2 = X_3 + X_4$, $Z_3 = X_5$.
Show that $(Z_1, Z_2, Z_3) \sim D_3(\alpha_1+\alpha_2, \alpha_3+\alpha_4, \alpha_5; \alpha_6+\ldots+\alpha_{11})$.

GROUP B (2 x 12) = [24]

1. Let T_1 and T_2 be two random variables and θ a real number such that

$$P[|T_1 - \theta| > t] \leq P[|T_2 - \theta| > t] \quad \text{for all } t.$$

Show that $E\,\phi(|T_1 - \theta|) \leq E\,\phi(|T_2 - \theta|)$ for any strictly increasing function $\phi(.)$ [e.g., $\phi(x) = x^2$].

2. Let X be a continuous random variable having distribution function F. Show that $Y = F(X)$ has $U(0, 1)$ distribution.

[Hint: Define $F^{-1}(y) = \inf\{x : F(x) \geq y\}$].

3. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ variables. Show that $(X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_{n-1} - \overline{X})$ and $\overline{X}$ are independent.

:bcc: