Indian Statistical Institute

B.Stat. (Hons.) III year

First Semestral Examination, 1998-99

Linear Stat. Models

Maximum marks: 60 *Date: 02.11.98* *Duration: 3 hrs*

Note : *This pape contains a maximum of 80 marks. Answer as many questions as you can. Maximum marks you can score is 60.*

1. a) State clearly under what conditions a variable can be treated as "co-variate" in the "Analysis of covariance model."

   b) Consider a one-way classification (treatment) model with a single covariate. Formulate a suitable Analysis of covariance model and explain how to carryout estimation of parameters and test of the hypothesis of absence of treatment effects.

   [5 + 15 = 20]

2. In studies of the effect of acid rain on the biomass in fresh water lakes, biologists have found that biomass decreases as acid concentreation increases. If the lakes have sources of phosphorus, however, biomass increases with an increase in the phosphorus available. In an effort to marke a thorough study, researchers took water samples from 18 randomly selected lakes and measured acidity (x_1), available phosphorus (x_2) and population density (y) of certain species of algae plant. The following statistics were computed (in the usual notation).

	corrected sum of squers and product-matrix			Mean
y	14.400	-3000	2100	1400
x_1	-3000	1600	900	2100
x_2	2100	900	3600	760

$\hat{\beta}_1 = -2.563; \hat{\beta}_2 = 1.224; \quad s^2 = 276$

where s^2 denotes the (Resedual Sum of squares/d.f)

   a) Write down the regression equation of y on x_1 and x_2.

   b) Test the hypotheses $H : \beta_1 = \beta_2 = 0$ against all alternatives basis of the above information.

   c) If acidity is increased one unit and phosphorus held fixed, wh the effect on population density? [5 + 10 + 5 =

P.T.

(i) Obtain an expression for the relative loss in effici ency measured by

contd..... 2/-

3. An investigation is made to determine how the speed of drill bits and the viscocities of a lubricant affect the speeed with which a drill can penetrate granite rock. Records were examined for wells that had been drilled using three viscocities of lubricant and two bit speeds. Information on two wells for each speed-viscocity combination was recorded and data are given below. Analyse the data

Speed of Drill	Viscocity of Lubricant		
	1	2	3
1	16.1	12.7	8.1
	15.8	15.3	9.3
2	20.2	11.2	10.7
	18.7	14.6	12.1

[20]

4. We have 3 attributes A, B, C. A takes I values, B takes J values and C takes K values. Define,

$$p_{ijk} = P(A = i \ \& \ B = j \ \& \ c = k); \quad 1 \le i \le I; 1 \le j \le J; 1 \le k \le K$$

$$p_{ijk} > 0; \quad \sum_i \sum_j \sum_k p_{ijk} = 1$$

a) Show that p_{ijk}'s satisfy the loglinear model $[AB][BC][CA]$ if and only if $\forall_i > 1, \forall_j > 1$, and $\forall_k \ge 1$

$$\frac{p_{11k}p_{ijk}}{p_{i1k}p_{1jk}} \text{ is free of } i, j, k.$$

you may use the fact that $[ABC]$ is the saturated loglinear model.

b) Compute the d.f. of the loglinear model $[AB][BC][CA]$

[12 + 8 = 20]

----x----

INDIAN STATISTICAL INSTITUTE

B.Stat. II year and B. Stat. III year

First. Semestral Examination, 1998-99

Geology

Maximum Marks: 100 Date : 03.11.98 Duration Three Hours (3 hrs)

Note : *Answer all questions. Marks allotted to each question are indicated the margin.*

1. By following an international gravity formula we know the 'g' values for any point on the earth's geoid surface taking into account the earth's shape and rotation. Indicate the other corrections necessary to determine the Bouger gravity anomaly, if any, at any particular gravity reading station situated on a topographic high. 5

2. (a) What is the magnetic inclination at the magnetic north pole of the earth ? 1

(b) What would be the magnetic declination at Calcutta if the magnetic poles of the earth coincide with geographic poles ? Briefly justify your answer. 4

3. Describe the essential feature of solid solution relationship between albite and anorthite . 5

4. Assuming a single earthquake point at the geographic north pole explain the occurrence of 'shadow zone' with the help of a sketch. 5

5. (a) Distinguish between a sill and a dyke with the help of a sketch. 2

(b) Name the commonest variety of acidic igneous rock and give its essential mineralogical composition. 1+2

6. (a) How are mechanically deposited sedimentary rocks primarily classified ? 3

(b) A calcite cemented sandstone can be described as a hybrid rock. Briefly explain why. 2

7. (a) Name one primary sole structure which can be utilized to determine the palaeoflow direction in waterlain sediments. 1

(b) Draw 3-D sketch of trough cross beds and label set boundary, foresets and flow direction in the diagram. 4

8. i) How does an antiform differ from a synform ? 1

ii) What are meant by hinge line and inflection line of a folded surface. Explain with sketches. 2

iii) Draw a sketch to show the dip isogon pattern of a similar fold. 2

9. i) What is a dip slip fault ?

ii) Draw a 3-D sketch of a dip slip fault and indicate the hanging wall block, footwall block, and net slip.

iii) Illustrate with sketches how strike fault may cause repetition of beds in outcrops.

1+2+2

P.T.O.

(i) Obtain an expression for the relative loss in effici-
ency measured by

contd..... 2/-

10. Why only the hard parts of organisms are usually preserved as fossils? Name few hard parts which are generally preserved as fossils. Explain how quick burial is helpful for fossilisation. Explain how fossils are used to date their host rocks. How fossil occurrences are used as an evidence for continental drift? 3x5=15

11. Write short notes on a) geotherm, b) regional metamorphism c)schist d) phase rules used to understand aspects of metamorphism and e)grades of metamorphism. 3x5=15

12. Why the western margin of Africa is devoid of any long mountain chain? What are cratons and mobile belts? Why seismic activities are more frequent in the Himalayan region? What do you understand by the term isostasy? Why symmetrically disposed, oppositely polarised magnetic strips are found on both sides of a mid oceanic ridge? 3x5=15

13. Select the right answer for any ten(10) of the following questions. 10

i) The mean density (gm/c.c) of the earth is in the range of
(A) 2.0 -- 3.0; (B) 3.0 -- 4.0; (C) 4.0 -- 5.0; (D) 5.0 -- 6.0

ii) The age (as per radiometric dating) of the oldest rock in the earth's crust is about
(A) 3.8 b.y ; (B) 4.8 b.y ; (C) 2.8 b.y ; (D) 5.8 b.y

iii) The most common oxide in the chemical make up of the earth's crust is
(A) Al_2O_3 (B) MgO (C) SiO_2 (D) CaO

iv) SiO_4 - tetrahedra in amphibole group of minerals are arranged in
(A) single chains; (B) double chains ; (C) sheets; (D) rings.

v) On a scarp face along an E-W trending road a thin bed of sedimentary rock occurs at fixcal elevation all along the road, the bed may be
(A) vertical ; (B) dipping towards south or north or horizontal ; (C) dipping towards east or west ; (D) dipping towards any direction.

vi) Which of the following minerals is most resistant to chemical weathering ?
(A) Olivine ; (B) Pyroxene ; (c) Mica ; (D) Quartz.

vii) The most abundant sedimentary rocks in the earth's crust is
(A) sandstone ; (B) shale ; (C) limestone; (D) conglomerate.

viii) A fold which has an interlimb angle between 120° and 70° is called a
(A) gentle fold; (B) open fold; (C) close fold; (D) tight fold.

ix) Which one of the following is most abundant in the earth's crust
A)O, B)Fe, C)Mg, D)Ca.

x) Which one of the following is most abundant in the entire earth
A)O, B)Fe, C)Mg, D)Ca.

xi) A fold which has the oldest bed at its core is called
A) anticline, B)syncline, C)antiform, D) synform

xii) Horizontal beds deposited over inclined beds indicate
A)angular unconformity B)paraconformity, C) igneous intrusion D) conformity.

xiii) The first appearance of birds was in
A)Mesozoic B) Palaeozoic C)Cenozoic D)Azoic

----------------------X----------------------

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1998-99
SEMESTRAL - I EXAMINATION
Sample Surveys

Date: 6.11.1998 Maximum Marks: 100 Time: 3 hours

Note: Attempt ALL questions.

1.(a) Suppose a sample is selected in 3 draws using a PPSWR selection method. Let ν denote the number of distinct units in the sample. Show that the following estimator is unbiased for the population total $Y = \sum_1^N y_i$.

$$t \begin{cases} = \dfrac{y_i}{p_i} & \text{if } \nu = 1 \\ = \dfrac{1}{3}\left[\dfrac{y_i}{p_i} + \dfrac{y_j}{p_j} + \dfrac{y_i + y_j}{p_i + p_j}\right] & \text{if } \nu = 2 \\ = \dfrac{1}{3}\left[\dfrac{y_i}{p_i} + \dfrac{y_j}{p_j} + \dfrac{y_k}{p_k}\right] & \text{if } \nu = 3 \end{cases}$$

where p_i's are normed size measures.

(b) A population consists of N clusters of varying sizes M_i, $i = 1, 2, \ldots, N$. Suppose that k clusters are selected at random and without replacement. Let Y_{ij} be the value taken by the study variate on the jth unit of the ith cluster, $j = 1, 2, \ldots, M_i$; $i = 1, 2, \ldots, N$. Let $\overline{Y}_i = \sum_{j=1}^{M_i} Y_{ij}/M_i$ be the ith cluster mean.

(i) Write down the conventional unbiased estimator for the mean $\overline{Y} = \sum_1^N M_i \overline{Y}_i / \sum_1^N M_i$ when M_i's are known, (ii) Suggest a method of estimating its sampling error, (iii) When M_i's are unknown derive the Hartley - Ross - type unbiased estimator for $\overline{Y}$.

[(11) + (2+3+9)] = [25]

2.(a) Denote Neyman's optimum allocation when SRSWOR design is used in all k strata by $(n_1^o, n_2^o, \ldots, n_k^o)$. Suppose that the actual allocation in practice turns out to be $(n_1^a, n_2^a, \ldots, n_k^a)$.

(i) Obtain an expression for the relative loss in efficiency measured by

contd..... 2/-

$$\{ Var_a(\hat{\bar{Y}}_{st.}) - Var_o(\hat{\bar{Y}}_{st.}) \} / Var_o(\hat{\bar{Y}}_{st.})$$

where the symbols have the usual meaning and the stratum sizes are assumed large, (ii) Further, derive a quick upper bound to the above expression in terms of $\emptyset$, the relative deviation of sample allocations given by $\emptyset = |n_i^a - n_i^o| / n_i^o$

(b) (i) What are the advantages of systematic sampling over simple random sampling ?

(ii) When the study variable in the populations show a linear trend would you prefer systematic sampling to simple random sampling w.r.t. variance criterion ? Justify your answer. [(6+6) + (3+10)] = [25]

3.(a) A survey was conducted in a ward consisting of 625 households by covering a sample of 50 households, selected using srswor scheme, to estimate the average monthly expenditure on salt. The estimate was found to be Rs.4.20 with a standard error of 0.47. Using this information, determine the sample size needed to estimate the same characteristic in a neighbouring ward on the basis of a sample to be selected by srswr scheme such that the length of the confidence interval at 99% confidence level is 10% of the true value. State clearly the assumptions involved.

(b) Verify whether $\pi_{ij} > \pi_i \pi_j$ for a probability proportional to size With Replacement design of n draws from N population units with a given size measure. (8+9) = [17]

4. A sample survey was conducted to estimate the total yield of paddy in a district. A stratified two-stage sampling design was adopted with villages as first stage units and plots within them as second stage units. From each stratum 4 villages were selected with probability proportional to area and with replacement and 4 plots were selected from each selected village with equal probability and without replacement. The data on yield for the sample plots together with information on selection probabilities are given below.

contd..... 3/-

Stratum	Sample village	Inverse of probability of selection	Total no. of plots	Yield of sample plots			
				1	2	3	4
I	1	440.21	28	104	182	148	87
	2	660.43	14	108	64	132	156
	3	31.50	240	100	115	50	172
	4	113.38	76	346	350	157	119
II	1	21.00	256	124	111	135	216
	2	16.80	288	123	177	106	138
	3	24.76	222	264	78	144	55
	4	49.99	69	300	114	68	111
III	1	67.68	189	110	281	120	114
	2	339.14	42	80	61	118	124
	3	100.00	134	121	212	174	106
	4	68.07	161	243	116	314	129

Using the above data

(i) Obtain an unbiased estimate of the total yield of paddy in the district.

(ii) Obtain an unbiased estimate of the variance of the above estimate.

(iii) What are the possible sources of non-sampling errors in the above study ? How do you plan to assess and control them ?

(10+15+8) = [33]

:bcc:

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year: 1998-99

SEMESTRAL - I EXAMINATION

Theory of Stat. Inference I

Date: 9.11.1998 Maximum Marks: 120 Time: $3\frac{1}{2}$ hours

Note: Answer ALL questions. All questions have equal marks.

1. (i) State and prove Basu's theorem.

(ii) Using (i), find $E\left[\frac{X_1}{X_1+\ldots+X_n}\right]$ when $X_1 \ldots X_n$ are iid $F_\lambda(x) = \lambda e^{-\lambda x}$, $x > 0$.

2. Suppose N is an integer valued random variable. When $N = n$, $X_N \sim \text{Bin}(n, p)$. The data is (X_N, N). Show that X/N is an unbiased estimator of p. Find its variance.

3. (i) Suppose $X \sim U(\theta, \theta+1)$. Is $X-\frac{1}{2}$ unbiased for θ ? Is it UMVUE ?

(ii) Suppose $X_1, \ldots, X_n$ are iid $U(0,\theta)$. Is $\frac{n+1}{n} X_{(n)}$ UMVUE for θ ?

4. Suppose ε_i are iid $N(0,\sigma^2)$ and we observe $Y_i = \beta x_i + \varepsilon_i$ where β and σ^2 are unknown parameters and x_i's are fixed and known. Find the MLE of β and σ^2.

5. Let $X_1, \ldots, X_n$ be i.i.d. Bin (p). Find the MLE of $p(1-p)$ and its asymptotic distribution, appropriately normalised as $n \rightarrow \infty$.

6. Suppose X and Y are independent, $X \sim F_\lambda(x) = \lambda e^{-\lambda x}$ and $Y \sim F_\mu(x) = \mu e^{-\mu x}$. Find the UMPU test of size α for testing $H_0: \lambda = \mu$ vs $H_1: \lambda \neq \mu$.

7. Suppose $X \sim F(x) = \frac{1}{\theta} e^{-\frac{(x-\gamma)}{\theta}}$, $x \geq \gamma$, $Y \sim \frac{1}{\theta} g(y/\theta)$ where θ and γ are unknown parameters and X and Y are independent, g is a fixed known distribution. Consider testing $H_0: \gamma = \gamma_0$ and the rejection region $\frac{X-\gamma_0}{Y} \leq a$ or $\frac{X-\gamma_0}{Y} \geq b$ such that

contd..... 2/-

level is α. Show that the power of this test at any $\gamma' < \gamma_0$ is given by $1 - (1-\alpha)\exp(\frac{\gamma_0 - \gamma}{\theta})$.

8. Let $p(x|\theta)$ be the probability mass function of X such that $p(x|\theta) > 0$ $\forall$ $x \in \mathcal{X}$ and $\forall$ $\theta \in \Omega$. Suppose there exists a function T(x) such that $p(x|\theta)/p(y|\theta)$ is constant as a function of θ if and only if $T(x) = T(y)$. Show that T(X) is minimal sufficient for θ.

:bcc:

INDIAN STATISTICAL INSTITUTE
B.STAT.(HONS.) III YEAR: 1998-99
SEMESTRAL-I EXAMINATION
DIFFERENTIAL EQUATIONS

Date:11.11.98 Maximum Marks:60 Time: 3 Hours

I.(a) What is an exact differential equation? Determine if

$$(2xy^3 + y\cos x)\,dx + (3x^2y^2 + \sin x)\,dy = 0$$

is exact, hence solve it.

(b) Verify that the differential equation

$$x^2y' = 3(x^2+y^2)\tan^{-1}(y/x) + xy$$

is homogeneous and solve it. [5+5=10]

II.(a) Find the general solvtion of

$$x^2y'' + xy' - y = x^2$$

by the method of variation of parameters.

(b) Find a particular solution of

$$y'' + 2y' + y = e^{-x}\log x$$ [5+5=10]

III.(a) Prove Rodrigue's formula for the Legendre's Polynomials

$$P_n(x) = \frac{1}{2^x}\sum_{k=0}^{[n/2]}(-1)^k \frac{(2n-2k)!}{k!\,(n-k)!\,(n-2k)!}x^{n-2k}$$

Using Rodrigue's formula for $P_n(x)$ and directly differentiating $P_n(x)$, show that $P_n(x)$ satisfies

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0$$

(b) Prove that Legendre's polynomials $\{P_n(x)\}$ form a sequence of orthogonal functions on $[-1,1]$. [5+5=10]

IV.(a) Prove that if $y_1(x)$ and $y_2(x)$ are two linearly independent solutions of $y''+P(x)y'+Q(x)y = 0$ where P(x) and Q(x) are continuous functions on $[a,b]$ then the zeros of these functions are distinct and occur alternately.

(b) Prove that $\frac{d}{dx}J_0(x) = -J_1(x)$ and $\frac{d}{dx}xJ_1(x) = xJ_0(x)$ where J_0 and J_1 are Bessel functions of order 0 and 1 respectively. Deduce that the positive zeros of $J_0(x)$ and $J_1(x)$ occur alternately. [5+5=10]

p.t.o.

V. (a) Let $y_1(x)$ and $y_2(x)$ be any two solutions of $y''+P(x)y'+Q(x)y$
where P, Q are continuous functions as [a,b]. Prove that $y_1($
and $y_2(x)$ are linearly dependent if and only if their wrons
$W(y_1, y_2)$ vanishes identically on [a,b].

(b) Consider the initial value problem

$$y' = x^2|y| \;;\; y(x_0) = y_0.$$

State with reason for what points (x_0, y_0) the above problem has a unique solution as some interval $|x-x_0| \leq h$. [5+5=

VI. (a) Find the normal form of Bessel's equation

$$x^2 y'' + xy' + (x^2 - p^2)y = 0$$

and use it to show that every non-trivial solution has infinitely many positive zeros.

(b) Let $y_p(x)$ be a nontrivial solution of Bessel's equation on the positive x-axis. Show, if $0 \leq p < \frac{1}{2}$, then every inte
of length π contains at least are zero of $y_p(x)$; if $p = \frac{1}{2}$,
then the distence between successive zeros of $y_p(x)$ is exa
π; if $p > \frac{1}{2}$, then every interval of length π contains at
one zero of $y_p(x)$. (Compare with $u''+u=0$). [5+5=10

VII. (a) Suppose a particle initially at rest at (0,0) slides unde
the influence of gravity along smooth curves joining (0,0
and (a,A). Use the methods of calculus of variation to fi
out the equation of the curve for which the particle take
the least time to descend to the point (a,A).

(b) A uniform flexible chain of given length hangs between tw
points. Find its shape if it hangs in such a way as to minimize its potential energy. [5+

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1998-99
SEMESTRAL - I EXAMINATION

Economics III

Date: 13.11.1998 Maximum Marks: 100 Time: 3 hours

Note: Answer any FIVE questions.

1. How can the B.O.P. crisis of 90/91 be attributed to inefficiency of the Indian economy ? Explain your answer.

[20]

2. Critically evaluate the administered interest rate policy and the directed credit programme persued by the government of India during the Mahalanobis era.

[20]

3. Critically evaluate the present monetary policy of the Govt. of India.

[20]

4. It is the export earning rather than the import tariff that determines the extent of competition faced by the domestic industry. Explain.

[20]

5. Discuss the recent currency crisis of Thailand. What lessons has India to learn from the Thai experience ?

[20]

6. Define national income. Distinguish between

(i) Gross domestic product (GDP) and Gross National Product (GNP).

(ii) GDP at factor cost and GDP at market prices.

(iii) GDP at current prices and GDP at constant prices.

[20]

:bcc:

INDIAN STATISTICAL INSTITUTE
B.STAT.(HONS.) III YEAR:1998-99
SEMESTRAL-I BACKPAPER EXAMINATION
SAMPLE SURVEYS

Date:13.1.99 | Maximum Marks:100 | Time: 3 Hours

Note: Attempt all questions.

1.(a) Define a 'sampling design' and a 'sampling scheme'. Illustrate these terms with respect to Lahiri-Midzuno-Sen method of selection of a sample.

(b) Calculate the indusion probability π_i of the unit U_i and the joint inclusion probability π_{ij} of the pair of units $U_i, U_j (i \neq j)$. Verify whether $\pi_{ij} - \pi_i \pi_j$ is non-negative.

(c) A sample survey was conducted in a village consisting of 625 households by covering a sample of 50 households selected using a srswor scheme to estimate the average weekly expenditure on soap powder. The estimate was found to be Rs.4.20 with a standard error of 0.47. Using this information, determine the sample size needed to estimate the same characteristic in a neighbouring village on the basis of a sample to be selected by srswr scheme such that the length of the confidence interval at 95% confidence level is 20% of the true value. State clearly the assumptions involved in finding the sample size.

(2+2+4)+(3+4+3)+(6+1)=25

2.(a) When the cluster sizes are varying, write down a formula for the 'intra-cluster correlation coefficient'. A population consists of 14 clusters of size 6 each. Find the bounds for the intra-cluster correlation coefficient among the elements of the cluster.

(b) Show that the relative efficiency of a simple random sample of a cluster of M units, compared to direct sampling is given by $1/(1+\overline{M-1}\rho)$, where $\rho = 1 - \left\{M\sigma_w^2/(M-1)\sigma^2\right\}$, σ_w^2 and σ^2 denoting the within and total variances respectively.

(c) Suppose that n clusters are selected at random and without replacement from a population of N clusters of sizes $M_i, i=1,2,\ldots,N$. Let $\bar{Y}_i = \sum_{j=1}^{M_i} Y_{ij} | M_i$ be the ith cluster mean, where Y_{ij} is the value taken by the study variable y on the jth unit of the ith cluster, $j=1,2,\ldots,M_i$, $i=1,2,\ldots,N$. Writedown the conventional estimator of $\bar{Y} = \sum_1^N M_i \bar{Y}_i | \sum_1^N M_i$ and its sampling error when M_i's are unknown. Further, demonstrate how you would obtain an unbiased estimator due to Hartley and Ross for $\bar{Y}$.

(2+2)+(6)+(3+4+8)=25

p.t.o.

3 (a) Suppose that n first stage units (fsu) are selected from N fsu's by PPSWR scheme. From the selected fsu's consisting of $M_1, M_2, \ldots, M_n$ second stage units (s.s.u's) suppose that m_i s are selected by srswor design.

(i) Write down an unbiased estimator for the population to of a characteristic y and demonstrate its unbiasedness

(ii) Explain how you would estimate the variance of your estimator in (i) above.

(b) Explain how you would assess and control non-sampling error at the field operations stage of a sample survey. (2+3+4)+8

4.(a) For estimating the total Y of current population in a regio two subsamples of 6 villages each are selected (circular sy tematically) from each stratum with independent random star Using the data given in the table

(i) Obtain a ratio estimate for Y taking the previous censu population (x) as auxiliary information and (ii) Compare efficiency with that of conventional unbiased estimate.

Total no. of villages (N) and sample totals of x and y

Stratum Number	No. of villages N	Sub-sample 1 x	Sub-sample 1 y	Sub-sam x
1	2044	3722	3935	3456
2	1304	3625	4033	4171
3	1265	2769	3050	3746

(Total of x for the region = 3,155,680).

(b) A population is divided into two strata of sizes 38 and 23. Two independent circular systematic samples of sizes 4 each drawn from the first stratum and the following values are observed on a study variate y:

	y-values
Sub-sample 1	247, 238, 359, 125
Sub-sample 2	256, 214, 368, 141

From the second stratum a simple random sample of two units selected without replacement and the y-values are noted as 427 and 326.

Obtain an (i) estimate of the average of y-values in the population and (ii). Obtain an unbiased estimate of variance of your estimate in (i) above. (8+12)+(5+8)=33

INDIAN STATISTICAL INSTITUTE
B.STAT.(HONS. III YEAR: 1998-99
SEMESTRAL-II EXAMINATION
ECONOMICS-IV

Date:16.4.99 Maximum Marks:100 Time: 3 Hours

Note: Answer any five questions.

1. In a normal general linear regression model with usual notations:

(a) Derive the ML estimators $\tilde{\beta}$ and $\tilde{\sigma}^2$,

(b) Prove that

$$\tilde{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1}), \quad \frac{T\tilde{\sigma}^2}{\sigma^2} \sim \chi^2_{T-K}$$

and $\tilde{\beta}$ and $\tilde{\sigma}^2$ are independent. [8+12=20]

2. In the general linear regression model:

(a) Derive the expectation and covariance matrix of the prediction error $\hat{y}_0 - y_0 = X_0\hat{\beta} - y_0$. Also show that the function $X_0\hat{\beta}$ is the best linear unbiased predictor of y_0, i.e., prediction error of any other linear unbiased predictor of y_0 of its sampling variance.

(b) Simplify the expression of the sampling v single observation with K=2 (i.e. X_0 is a 1 x 2 matrix). [

3. Write down the dummy variable regression models where (i) only slope differs, (ii) only intercept differs and (iii) both slope and intercept differ. Prove that the situation where both slope and intercept differ is s me as using separate regressions so far as estimation of regression coefficients is concerned. [20]

4. What do you mean by truncated and censored samples? Discuss one estimation procedure to estimate the regression coefficients in a censored sample. [20]

5. Describe DW test for autocorrelation giving important steps to derive the bounds of probability distribution of DW statistic under the assumption that $\rho=0$. [20]

6. (a) Explain briefly what is meant by identification problem in the context of (linear) simultaneous equations model.

(b) Consider the following three-equation model:

$$y_1 = \beta_{13} y_3 + \gamma_{12} x_2 + e_1$$

$$y_2 = \beta_{21} y_1 + \beta_{23} y_3 + \gamma_{21} x_1 + \gamma_{22} x_2 + e_2$$

$$y_3 = \beta_{31} y_1 + e_3$$

Discuss the identification status of each equation. [14+6=20]

p.t.o.

7. Suppose

$$q = \alpha p + u \quad \ldots (i)$$

and

$$q = \beta p + v \quad \ldots (ii)$$

are two relations operating simultaneously, where q and p are ables, α and β are unknown constants, and u and v are non-obse random variables with zero means, constant (unknown) variance σ_u^2 and σ_v^2 and zero covariance.

(a) Show that the LS regression coefficient of q on p is equa weighted average of α and β, the weights being σ_v^2 and σ_u^2.

(b) If in addition it is known that $\sigma_v^2 = K\,\sigma_u^2$, where K is a know non-negative constant, show how α and β might be estimated.

(c) Suppose equation (ii) is replaced by

$$p = \delta z + v \quad \ldots (iii)$$

where z is an exogenous variable. Under the same assumptio about the distribution of u and v as before, what can you about the LS regression coefficient of q on p? [8+5

INDIAN STATISTICAL INSTITUTE
B.STAT. (HONS.) III YEAR: 1998-99
STATISTICAL INFERENCE-II
SEMESTRAL-II EXAMINATION

Date: 19.4.99 Maximum Marks:120 Time: 3 Hours

Note: Answer any four. All questions carry equal marks.

1.(i) Define Wald's SPRT and under appropriate conditions, show that it stops with probability one.

(ii) State Wald's first identity and use it to derive E(N) where N is a geometric random variable.

2.(i) Define a U statistics.

(ii) Show that under appropriate conditions and with suitable normalisation, a U-statistic has an asymptotic normal distribution. (Assume that the Kernel size is 2)

(iii) Use (ii), to find the asymptotic distribution of the sample variance.

3.(i) Define the empirical distribution function.

(ii) Define the one sample Kolmogorov-Smirnov statistic and show that it is distribution free when the underlying distribution is continuous.

(iii) Call the distribution in (ii) $K_n(.)$.
Let $K_{n,F}$ be the distribution of the Kolmogorov-Smirnov statistic when the underlying distribution F is not necessarily continuous. State and prove a relation between K_n and $K_{n,F}$.

4.(i) Show that the usual lootstrap procedure may be looked upon as weighted resampling with multinomial weights.

(ii) Show that any general weights $(W_1 \ldots, W_n)$ such that $\sum_{i=1}^{n} W_i = C$(constant), leads to the same estimate for the variance of $\bar{X}$.

(iii) Let $(W_1 \ldots, W_n)$ be the delete d jackknife weights vector.

(a) Write down the distribution of $(W_1 \ldots, W_n)$

(b) Find $E(W_2 | W_1)$

(c) Find $V(W_1)$.

5.(i) State and prove the theorem done in class on asymptotic relative efficiency.

(ii) Let $X_1 \ldots, X_n$ be iid each with distribution function F_X

Consider testing

$$H_o : F_X(x) = F(x) \ \forall\, x$$

$$H_1 : F_X(x) = F(x-\theta) \ \forall\, x$$

where F is a given distribution.

Find the asymptotic efficiency of the sign test relative to the test:

T_n Reject H_o if $\Sigma X_i > t_n$.

6. Suppose $X_1, \dots, X_n$ are iid $U(0,1)$.
Let f_n be the density of $n^{1/2}(M_n - 1/2)$ where M_n is the sample median.

(i) Show that for some σ^2, $f_n(x) \longrightarrow \phi_{\sigma^2}(x)$ for all x, where ϕ_{σ^2} is the normal density with mean 0 and variance σ^2. Determine σ^2.

(ii) Can you conclude that $n^{1/2}(M_n - 1/2)$ has an asymptotic normal distribution? Give reason.

INDIAN STATISCAL INSTITUTE

Final Examination : Semester II (1998-99)

B.Stat. (Hons) : Year III

Introduction to Stochastic Processes

Date : 21.4.99 Maximum Marks : 140 Duration : 4 Hours

Note : This paper carries questions worth a **total of 165 points**. Answer as much as you can. The **maximum** you can score is **140 points**.

1. Consider a Markov Chain $\{X_n, n \geq 0\}$ with state space $S = \{1, 2, \cdots, 6\}$ and transition probability matrix

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}.$$

(a) Find $P(X_4 = 6 \mid X_0 = 2, X_1 \neq 2, X_2 = 2)$.
(b) Find $P(T_1 = 3 \mid X_0 = 3)$.
(c) Find f_{34}.
(d) If π is a stationary distribution for the chain with $\pi_6 = \frac{1}{3}$, find π.

[4 × 10 points = 40 points]

2. Consider a machine that is subject to failure with θ_i $(i \geq 1)$ as the probability that it fails during the ith day of its operation. Once a machine fails sometime during a particular day, it is replaced by a new but identical machine at the beginning of the next day. The succesive machines act independently. Consider the Markov Chain $\{X_n\}$ on $S = \{1, 2, \cdots\}$, where X_n denotes the age of the machine that is in operation at the beginning of the nth day.
(a) Find the transition probabilities of the chain.
(b) Find $f_{11}^{(n)}$ and hence find the condition for positive recurrence of the chain.
(c) If the chain has a stationary distribution π, find π_3.
(d) Find E $(T_3 \mid X_0 = 1)$.

[4 × 10 points = 40 points]

—— P.T.O. ——

3. Let $\{Y_n, n \geq 0\}$ be an irreducible Markov Chain on a state space S with all states essential and having common period 3. Put $X_n = Y_{3n}$ for $n \geq 0$. Show that $\{X_n, n \geq 0\}$ is a Markov Chain for which the state space S splits into 3 closed irreducible aperiodic classes.

[15 points]

4. Let $\{X_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ be two independent Markov Chains on the same state space S and having the same transition probabilities $p_{ij}, i, j \in S$. Put $W_n = (X_n, Y_n)$ for $n \geq 0$.
(a) Show that $\{W_n, n \geq 0\}$ is a MC and find its transition probabilities.
(b) Let $\tau = \inf\{n \geq 1 : X_n = Y_n\}$. Show that τ is a stopping time for the process $\{W_n\}$.
(c) Let $\{Z_n, n \geq 0\}$ be the process obtained by watching $\{X_n\}$ up until time τ and then switching to $\{Y_n\}$ after time τ. Show that $\{Z_n, n \geq 0\}$ is a MC and find its transition probabilities.

[3×10 points $= 30$ points]

5. Consider N objects marked $1, 2, \cdots, N$ arranged in some ordered list. At each point of time, a request is made to retrieve one of these objects — θ_i being the probability that object i is requested (independently of the past). At each stage, after the request is made, the object that is requested is put back in the list one step closer to the front, unless, of course, it was already in the front.
(a) Modelling the successive changes in ordering of the list as a Markov Chain, describe the state space and the transition probabilities. Show also that, if $\theta_i > 0$ for all $i = 1, 2, \cdots, N$, then the chain is irreducible aperiodic.
(b) Take $N = 3$ in (a) and assume that $\theta_i > 0$ for $i = 1, 2, 3$. Show that the chain is time-reversible. [Hint : Try to guess its stationary distribution.]

[2×10 points $= 20$ points]

6. Let $\{N_t, t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$.
(a) For $0 < s < u < t$ and non-negative integers $n \geq m$, find the conditional distribution of N_u given $N_s = m$ and $N_t = n$.
(b) Find the expected value of the random variable $Y = (N_t + 1) \cdot S_{N_t}$.

[2×10 points $= 20$ points]

INDIAN STATISTICAL INSTITUTE

B.STAT.(HONS) II AND III YEAR: 1998-99

Semester II (1998-99)

Sociology

Maximum Marks- 100 Date: 28.4.99 Duration 3 hours

Answer five questions with at least one from each group.
Questions carry equal value

Group : A

1. a) Which of the following three concepts Karl Marx has emphasised
   (i) Division of labour in society
   (ii) Class structure and class struggle
   (iii) Individual motivation.

   b) Discuss the relevance of your answer to (a) above in the context of studying rural development in India.

2. Choose any one of the three Indian social scientists, M.N Srinivas, Ramkrishna Mukherjee and Nirmal Kumar Bose and explain whether you consider his approach to the study of society in India is on the whole, Weberian, Durkheimian or Marxian.

3. If you are planning to undertake a project to study trend of socio-economic changes in rural India after independence, what will be your theoretical framework for the study? Discuss your answer in the light of any sociological theory.

Group : B

1. Define Social Science Research. What decisions are to be taken before designing a work-plan for social science research ?

2. What are the techniques of study? Discuss two different techniques which are used in social science research.

3. Formulate a specific research proposal of your interest to be examined by a board of experts.

Group : C

1. What are the experiences of planned economic development in India especially in the rural areas ? Would you argue for decentralised planning for rural development ? If so, why ?

2.(a) Discuss different forms of marriage generally found in our society.

(b) State how does divorce affect the stability of a family.

3. What is social network ? what are the main charecteristics of social network analysis ?

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

   (a) Ethnic group (b) Functions of a family (c) Women's status in our society
   (d) Land Reforms in India.

INDIAN STATISTICAL INSTITUTE

Final Examination : Semester II (1998-99)
B.Stat (Hon..) II & III Year
ANTHROPOLOGY

Date : 28-4-99 Maximum Marks: 100 Time : 3 Hours

Note: Use separate answerscripts for GROUP A and GROUP B. Answer *5* questions from each group

GROUP A

1. How do you define Anthropology? What are the distinguishing features of Anthropology?

2. Compare and contrast the morphological - anatomical characteristics of Australopithecine with Homo erectus.

3. Why is man unique in the animal kingdom?

4. Discuss the salient features of Darwin's theory of evolution. Bring out the weaknesses of his theory.

5. Define adaptation. What is the difference between adaptation and acclimatization?

6. Writes notes on any two of the following:
   (a) Primates
   (b) Glogger's rule
   (c) Marriage
   (d) Fertility

GROUP B

1. What is cell? Describe the functions of different cytoplasmic organelles of an animal cell. (2+8)

2. What is karyotype? State the number of chromosomes normally found in each group A- G, in the diploid set of : (a) the human male ;(b) the human female ? What are the bases by which the classification has been made? (2+6+2)

3. In man brown eyes (B) are dominant over blue (b), and dark hair (D) over red hair (d) (the genes are not linked to each other). What would be the outcome of marriages between persons of the following genotypes:
   (a) BBDd X BbDd
   (b) bbDd X BBdd (5+5)

4. Why do X-linked traits tend to skip generation? Under what circumstances would this not happen? (e.g. how could a colour blind man have a child who is also colour blind?) (8+2)

5. Define Hardy-Weinberg principle. Assume a population sample of 1000, whose estimated blood group frequencies are A=0.21, B=0.34 and O=0.45. If marriages occur at random in this population, what are the expected frequencies of the four blood groups in this population? (4+6)

6. Write short notes on any ***two*** of the following: (5+5)

(a) Crossingover
(b) Linkage
(c) Down's syndrome
(d) Genetic drift