INDIAN STATISTICAL INSTITUTE

M. Stat - I year (2012-2013)

First Semestral

Mid-Semestral Examination

Measure Theoretic Probability

Date : 07.09.2012 | Total Marks - 100 | Time : 3 Hours

NOTE : Answer all questions Maximum you can score is 100.

. (a) The Borel σ-field in $\mathbb{R}$ is, by definition, the σ-field generated by the family of all intervals (of $\mathbb{R}$) of the form (a, b) with $-\infty < a < b < \infty$. Show that all countable sets are Borel. [5]

(b) Let $\mathscr{A}$ be a semi-ring in Ω. Let $A, A_1, \ldots, A_n$ be in $\mathscr{A}$. Show that $A \cap A_1^c \cap \ldots \cap A_n^c$ is a finite disjoint union of sets in $\mathscr{A}$. [5]

Show that a sub-σ-field of a countably generated σ-field need not be countably generated, by first stating **all** the main steps and then proving **all** these with perfect and **complete** arguments. [30]

. Let $\mathscr{A}$ be a family of subsets of Ω such that $\phi \in \mathscr{A}$. Let $\alpha : \mathscr{A} \to [0,\infty]$ be a set function such that $\alpha(\phi) = 0$. Let μ^* be "the canonical outer measure generated by α".

a) Show that μ^* is, indeed, an outer measure satisfying

$$\mu^*(A) \leq \alpha(A) \quad \forall A \in \mathscr{A}$$ [10]

b) Show that $\mu^*(A) = \alpha(A)\, \forall A \in \mathscr{A}$ iff α enjoys the following property : $A \subset \bigcup_{n=1}^{\infty} A_n$, A, A_n are in $\mathscr{A}$ $\forall n \geq 1 \Rightarrow \alpha(A) \leq \sum_{n=1}^{\infty} \alpha(A_n)$. [10]

c) Assume furthermore that $\mathscr{A}$ is a semi-ring in Ω and α is finitely super-additive in the sense that $A = \bigcup_{i=1}^{n} A_i$ for some $n \geq 2$, A, A_i are in $\mathscr{A}$ for i=1, ... n and the A_i are pairwise disjoint together imply that $\alpha(A) \geq \sum_{i=1}^{n} \alpha(A_i)$.

Show that all sets in $\mathscr{A}$ are μ^*-measurable. [10]

(a) Define a λ – system. [2]

(b) Show that $\mathscr{A}$ is a λ-system in Ω iff $\Omega \in \mathscr{A}$, $\mathscr{A}$ is closed under complementation and $\mathscr{A}$ is closed under countable disjoint unions. [10]

(c) Let P, Q be two probability measures on $(\Omega, \sigma(\mathscr{A}))$ where $\mathscr{A}$ is a π-system in Ω. Show that

$$P(A) = Q(A) \quad \forall A \in \mathscr{A}, \Rightarrow P(A) = Q(A) \quad \forall A \in \sigma(\mathscr{A}).$$

[You may assume the Sierpinski π - λ theorem.] [10]

(d) Let $(\Omega, \sigma(\mathscr{A}), \mu)$ be a finite measure space where $\mathscr{A}$ is a field in Ω. Show that for each $A \in \sigma(\mathscr{A})$ and for each $\varepsilon > 0$, there exists a set $B \in \mathscr{A}$ such that $\mu(A \Delta B) < \varepsilon$. [10]

(e) Use (d), to show that if two finite measures agree on a field $\mathscr{A}$, then they agree on $\sigma(\mathscr{A})$.

[Assume that the two measures are defined on $\sigma(\mathscr{A})$.] [10]

*****

INDIAN STATISTICAL INSTITUTE

M.Stat. (1 st year)—2012-2013

Mid -semestral Examination

Subject: Applied Stochastic Process

Exam date: 12.09.2012

Full marks: 35

Duration: 1 hour 15 minutes

Each question carries 8 marks with a total of 40 marks. Attempt all questions. Maximum marks you can score is 35.

1. Consider deterministic model in case of simple epidemic and initial number of susceptible= n and initial number of infected=1.Time is discrete and at any time point rate of infection is proportional to the product of the number of infected and the number of susceptible. Show that if T=duration of epidemic then $T \to 0$ as $n \to \infty$.

2. Let $P= ((p_{ij}))$ and $p_{ij}= C_i \,|i+1-j|$, where C_i is constant with i,j= 1,2, ...,n.

Let $\{X_n\}$ be a stochastic process on the state space {1,2......,n} with transition matrix P. Let T=min $\{n \geq 1: X_n=X_{n-1}\}$. Show that T is finite with probability 1 starting from any state.

3. Let $\{X_n\}$ be a Branching Process with progeny generating function G(s) and $G(s)= p_0+p_1s+p_2s^2+ .....$ with $p_0 >0$ and $p_0+p_1<1$. Show that extinction probability of the process is minimum non-negative real number that satisfy G(s) =s.

4. Let in a city there are n families with certain family name. Let each family takes the policy to have exactly three children through generations whereas all children gets married with probability one excepting that the eldest child marries with probability ½. Find the probability of eventual survival of the family name.

5. There are two types of particles A and B. After each unit time each particle A vanishes and gives one particle A with probability ¼, one particle B with probability 2/3 and no particle with probability 1/12 whereas each particle B vanishes giving one particle A with probability 2/3 , one particle B with probability ¼ and no particle with probability 1/12.

(X_n ,Y_n)=number of particles after time n. Given $(X_0,Y_0)=(2,1)$,find the expression for $E(X_n)$.Hence show that $X_n \to 0$ in probability.

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2012-13

Course Name : **M.Stat. 1st Year**

Subject Name : **Sample Survey and Design of Experiments**

Date : Sept 14, 2012 **Maximum Marks : 30 + 40 = 70** **Duration :** $2\frac{1}{2}$ *hrs.*

Note: Use separate answer sheets for two groups.

Group – Sample Survey. (Total Marks = 40)

Answer any four.

1. (a) Given any design p, prove that

   (i) $\sum_{i=1}^{N} \pi_i = E(\nu(s))$, and

   (ii) $\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \pi_{ij} = Var(\nu(s)) + E(\nu(s))(E(\nu(s)) - 1)$,

   where π_i and π_{ij}s are the first and second order inclusion probabilities and $\nu(s)$ is the effective sample size of a sample s.

   (b) For fixed effective sample size (n) designs, prove that $\sum_{j=1, j \neq i}^{N} \pi_{ij} = (n-1)\pi_i$.

   $(6 + 4 = 10)$

2. (a) Prove that for any given sampling design, there exists a sampling scheme that realises this design.

   (b) Write the two forms of variance of the Horvitz-Thompson estimator of population total Y and show that under some conditions (to be explicitely stated) both the forms are equivalent.

   $(6 + 4 = 10)$

3. (a) Given any design p and an unbiased estimator t for Y depending on order and/or multiplicity of units in sample s, derive an improved estimator for Y through Rao-Blackwellization.

   (b) Let $P_i (0 < P_i < 1, \sum_{i=1}^{N} P_i = 1)$ be known numbers associated with the units i of a population U. Suppose on the first draw a unit i is chosen from U with probability P_i and on the second draw a unit $j (\neq i)$ is chosen with probability $\frac{P_j}{1 - P_i}$.

For a sample of size 2 drawn under this sampling scheme, write down Des Raj's (1956) unbiased estimator for Y. Improve that estimator through Rao-Blackwellization.

$(6 + 4 = 10)$

4. (a) Define admissibility of an estimator in a class of estimators.

   (b) Prove that for any sampling design p with $\pi_i > 0 \forall i$, the Horvitz-Thompson estimator of population total Y of a variable y is admissible in the class of all homogeneous linear unbiased estimators.

$(3 + 7 = 10)$

5. State and prove Godambe's (1955) theorem regarding the existence of uniformly minimum variance estimator for Y within the class of homogeneous linear unbiased estimators.

(10)

Mid-semester Exam
Sample Survey and Design of Experiments
M.Stat. I yr. 2012-13
Part B: Design of Experiments

Date: 14/09/12 Time : 1 hour for Part B Marks: 30

Part B : Use separate answer booklet

1. Show that for a connected block design, every treatment contrast is estimable and the variance of the BLUE of an elementary treatment contrast always lies between $2\sigma^2/\lambda_{max}$ and $2\sigma^2/\lambda_{min}$ where λ_{max} and λ_{min} are the maximum and minimum nonzero eigen values of the C- matrix. [5+5=10]

2. Consider a Balanced Incomplete Block Design(BIBD) d with parameters b, v, r, k, λ. Let d_0 be a new block design obtained by deleting the last block of the design d. Is this block design d_0 connected? Orthogonal? Justify your answer. For the design d_0, obtain the C- matrix and the BLUE of the elementary treatment contrasts $\tau_i - \tau_j, \; i \neq j$. (Leave your expression of the BLUES in terms of $Q's$).

[5 +2 + 5 + 8= 20]

Indian Statistical Institute
Semestral Examination First Semester (2012-2013)
M.Stat. First Year
Large Sample Statistical Methods

Maximum Marks: 60 Date : 20.11.2012 Duration :- 3 hours

This question paper carries 66 marks. Answer as many questions as you can. The maximum you can score in this exam is 60.

1. Suppose $X_n = O_p(1)$ and $X_n - Y_n = o_p(1)$. Show that for any continuous function $g : \mathbf{R} \to \mathbf{R}$, $g(X_n) - g(Y_n) = o_p(1)$. [7]

2. Let $\{X_n\}$ be iid $N(\mu, \sigma^2)$, where both μ and σ are unknown. Let $S_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Find a variance stabilizing transformation for S_n and hence get a confidence interval for σ^2 with asymptotic coverage probability 0.95. Prove your answers. (You may assume the asymptotic distribution of (suitably centred and scaled) S_n.) [4+2=6]

3. Let $X_1, \ldots, X_n$ be iid from a distribution F, given by $F(x) = (1 - \frac{1}{\ln x})I_{x \geq e}$. Let $X_{(n)}$ denote the sample maximum. It is well known that $X_{(n)}$ can't have a non-degenerate limit distribution under any normalization. Does there exist a measurable transformation $g(.)$ such that the maximum of $Y_i = g(X_i)), i = 1, 2, \ldots, n$ has a non-degenerate limit distribution under appropriate normalization ? Prove your assertion. [6]

4. Let $X_1, \ldots, X_n$ be iid with common density $f(x, \theta)$ where $\theta \in \Theta$, Θ being an open subset of the real line. Stating appropriate regularity assumptions, show that, if there is a unique solution of the likelihood equation for all n and all sample values $(x_1, \ldots, x_n)$, with probability tending to one (under the true value of the parameter) as $n \to \infty$, this unique solution is the global maximum of the likelihood function. [6]

5. Give an example of an inconsistent maximum likelihood estimator and prove that it is indeed inconsistent. [5]

6. Let $X_1, \ldots, X_n$ be iid observations from a distribution given by $\theta G + (1 - \theta)H$ where $0 < \theta < 1$ is unknown while G and H are *known* distribution functions and are not identical. Can you suggest a $\sqrt{n}$-consistent estimator for the unknown θ based on the observations ? Prove your answer. [6]

7. Suppose you have an iid sample of size n from a multinomial population with k classes. Let π_i, $i = 1, 2, \ldots, k$ denote the probabilities of the classes. Find the asymptotic distribution of $T_n = \sum_{i=1}^{k} \frac{(n_i - n\pi_i)^2}{n\pi_i}$, where $n_i, i = 1, 2, \ldots, k$ denote the number of members in the sample falling in the i-th class. [10]

8. (a) What is the Jackknife estimator of variance of an estimator ? Is there any heuristic justification in using this esimator of variance ? [2+2=4]

(b) Prove that the Jackknife estimator of variance is consistent for estimating the variance of certain types of functions of the sample mean under appropriate assumptions. [8]

9. Let a_n be a given sequence of constants and $X \sim N(0,1)$. Construct a sequence of random variables X_n such that X_n converges almost surely to X and $E(X_n) = a_n$ for all n. [8]

INDIAN STATISTICAL INSTITUTE

First Semester Examination : 2012-13

Course Name : **M.Stat. 1st Year**

Subject Name : **Sample Surveys and Design of Experiments**

Date : Nov 22, 2012 **Maximum Marks : 50 + 50 = 100** **Duration : 4 *hrs.***

Note: Use separate answer sheets for two groups.

Group A – Sample Surveys. (Total Marks = 50)

Answer any four.

1. (a) Prove that for any given sampling design, there exists a sampling scheme that realises this design.

   (b) Consider Midzuno (1952)'s scheme of sampling design with θ_i as the selection probability of unit i. Determine the probabilities θ_i's such that the scheme happens to be an IPPS design with size measures x_i's where $x_i > 0$ for all i's. Also prove that for this IPPS design, if p_i's are the normed size measures corresponding to x_i's, then $\frac{n-1}{n}\frac{1}{N-1} < p_i < \frac{1}{n}$.

   $(6 + 6\frac{1}{2} = 12\frac{1}{2})$

2. (a) Prove that for a given sample s, if s^* denotes the reduced set equivalent to s obtained by ignoring the order and multiplicities of the units appearing in s, and if d^* denotes the data corresponding to s^*, then d^* is a minimal sufficient statistic.

   (b) Let $P_i(0 < P_i < 1, \sum_{i=1}^{N} P_i = 1)$ be known numbers associated with the units i of a population U. Suppose on the first draw a unit i is chosen from U with probability P_i and given that the first unit is i on the second draw a unit $j(\neq i)$ is chosen with probability $\frac{P_j}{1-P_i}$.

   For a sample of size 2 drawn under this sampling scheme, write down Des Raj's (1956) unbiased estimator for Y. Improve that estimator applying the concept of minimal sufficient statistic.

   $(8 + 4\frac{1}{2} = 12\frac{1}{2})$

3. (a) State and prove Rao (1979)'s theorem on the General Mean Square Error formula for a homogeneous linear estimator (HLE) for population total Y.

   (b) Use the theorem to obtain the variance of Horvitz and Thompson (1952)'s estimator by deriving the proper condition to be satisfied.

$(8 + 4\frac{1}{2} = 12\frac{1}{2})$

4. (a) Suppose in a survey population of N units, the values of a variable y is known to lie within a range guessed by some prior experience. The population may be divided into H disjoint strata. The h^{th} stratum consists of the set of units bearing y values in the range (a_{h-1}, a_h), but not knowing which units and how many of them have y values within this stipulated range. In such a situation, deduce an unbiased estimator and the variance of that for population mean $\bar{Y}$ by double sampling.

   (b) In the above situation, if the number of units falling in the stipulated ranges are known, find an unbiased estimator for $\bar{Y}$.

$(8 + 4\frac{1}{2} = 12\frac{1}{2})$

5. (a) Describe Politz and Simmon (1949, 1950)'s technique of using 'at-home-probabilities' to estimate the population mean $\bar{Y}$ by SRSWR scheme with an estimator for variance.

   (b) Describe Greenberg et al. (1969)'s unrelated question model to estimate a sensitive population proportion by a sampling design with $p(s)$ being the selection probability of a sample s of respondents. Also give a variance estimator.

$(6 + 6\frac{1}{2} = 12\frac{1}{2})$

First Semestral Examination
Sample Survey and Design of Experiments
M.Stat. I yr. 2012-13
Group B: Design of Experiments

Date: Nov-22-12 Time : 2 hours

Part B : Use separate answer booklet.

This question paper carries 55 marks and the maximum score is 50. Marks will be deducted if the answer is not to the point.

1. Prove or disprove with full justification: A proper connected balanced block design is always equireplicate. [7]

2. (a) For a BIBD with parameters b, v, r, k, λ, prove Fisher's inequality and hence or otherwise show that $b \geq v + r - k$.

   (b) Construct a BIBD with $b = v = 13$, $r = k = 9, \lambda = 6$. (Briefly outline your construction method) [7+10=17]

3. (a) Write down a Latin Square design d of order 5. Construct a new row column design d_0 from d, by deleting the last row. Identify this design d_0. Under an appropriate fixed effects additive linear model, specified by you, choose a full set of independent treatment contrasts and provide an unbiased estimator, if it exists, for each of the treatment contrasts of your set. Hence conclude about the connectedness of d_0. [2+1+8+1=12]

   (b) What is the relationship (justify your answer) between the C- matrix of a row column design d and that of the corresponding block design obtained from d, by treating the columns as blocks? Under what structural property of d, will both the C-matrices be equal?

   [3+2=5]

4. A subset of 27 treatments in a non-key block of a $(3^5, 3^3)$ confounding scheme is presented below. Identify the full set of confounded interaction components of the factorial design. Also obtain two independent treatment combinations for each of the remaining blocks of the confounded design.

0	0	0	0	2
1	0	0	1	0
0	1	1	0	0
0	1	0	2	2
1	2	2	1	2
0	2	1	2	0
1	0	1	2	2

[7+7=14]

INDIAN STATISTICAL INSTITUTE
First Semester Examination: 2012-13
M. Stat. I Year
Measure Theoretic Probability

Date: 27.11.2012 **Maximum Marks: 100** **Duration: 4 Hours**

Note : a) Answer both Group A and Group B.
b) All proofs should be rigorous, complete and clear. State clearly the result(s) you are using.
c) The notations are as usual.
d) This paper carries 120 marks. Maximum you can score is 100.

Group A

Answer either Question 1 or Question 2.
Question 3 is compulsory.
Maximum you can score from this Group is 30.

1 a) State the Monotone Convergence Theorem, Fatou's Lemma, Dominated Convergence Theorem and Fubini's Theorem.

[1+1+2+3=7]

b) State Holder's and Minkowski's inequalities.

[1+1=2]

c) Let $M(t) = E(e^{tX})$ for $t \in R$, where X is a random variables on a probability space ($\Omega, \mathscr{A}$, P). Show that if $0 < \theta < 1$, and $t, u, \in R$ then

$$M(\theta t + (1-\theta)u) \le (\mathrm{M(t)})^{\theta} \, (\mathrm{M(u)})^{1-\theta}.$$

Give details.

[4]

d) Let ($\Omega, \mathscr{A}$, P) be a probability space. Let f_n and f be real-valued measurable and integrable functions on Ω. Assume that $f_n - f$ is well-defined. Let

$$\alpha_n = \sup_{A \in \mathrm{A},} \left| \int_{\mathrm{A}} \mathrm{f_n} d\mu - \int_A f d\mu \right|.$$

Show that $f_n \to^{L^1} f \Leftrightarrow \alpha_n \to 0$ as $n \to \infty$.
[Hint: Use $\int |f_n - f| \le 2\, \alpha_n$] .

[7]

2. Let ($\Omega, \underline{A}$, P) be a probability space.

a) Let $\{\underline{A}_n\}_{n \ge 1}$ be a given sequence of σ – fields in Ω such that $\underline{A}_n \subset \underline{\mathrm{A}}$ $\forall \mathrm{n} \ge 1$. Define the tail σ – field of $\{\underline{A}_n\}_{n \ge 1}$. Show that if $\{\underline{A}_n\}_{n \ge 1}$ are independent, then the tail σ – field is P- trivial. State clearly the result(s) you are using.

[1+(5+1)=7]

P.T.O.

b) Let $\mathscr{B} = \{ A \in \underline{A} : P(A) = 0 \text{ or } 1\}$. Show that if $f : \Omega \to \overline{R}$ is $\mathscr{B}$ – measurable, then there exists $c \in \overline{\mathrm{R}}$ such that $f = c$ a.s. Here $\overline{R} = \mathrm{R} \cup \{+\infty, -\infty\}$.

[4]

c) Let $\{X_n\}_{n \geq 1}$ be independent. Show that lim sup X_n is almost surely a constant. Give details.

[5]

d) Let $\{X_t\}_{t \in T}$ be independent where T is infinite. Let $\underline{F}$ be the set of all non-empty finite subsets of T. Let $g = \bigcap_{S \in \underline{F}} \sigma\left(\bigcup_{t \notin S} \sigma(\mathrm{X}_t)\right)$. Show that g is P- trivial.

[4]

3 a) Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Let $f_n \to^{\mu} f$. If S is a closed subset of R such that $\mu\left(\mathrm{f_n}^{-1}(S^c)\right) = 0 \quad \forall n \geq 1$, show that $\mu\left(\mathrm{f}^{-1}(S^c)\right) = 0$. State the result(s) you are using.

[Hint Consider a mode of convergence for which the result is obviously true].

[4]

b) Let $(\Omega, \mathscr{A}, \mathrm{P})$ be a probability space. Show, using (a), that the pseudo- metric d defined by

$$d(A, B) = P(A \Delta B), \quad A, \ B \in \mathscr{A}$$

is complete (the notation is as usual).

[Hint: Recall that $I_{A \Delta \mathrm{B}} = |\mathrm{I_A} - I_B|$, and show that $\{ \mathrm{I_A} : A \in \mathscr{A} \}$ is a closed subset of $L^1(\Omega, \mathscr{A}, \mathrm{P})$]

[6]

Group B

Answer any three Questions. Maximum you can score from this Group is 70.

4 a) State and prove the two parts of the Kolmogorov Inequality.

[10+10=20]

b) Let $X_k > 0 \ \forall \mathrm{k} = 1, \ldots, \mathrm{n}$, and the X_k be independent. If the $\mu_k := E(X_k)$ satisfy $0 < \mu_k < \infty \quad \forall k = 1, \ldots, n$ and $T_k = X_1 \ldots X_k$, show that

$$P\left(\mathrm{T_k} \geq \lambda \prod_{\mathrm{i=1}}^{\mathrm{k}} \mu_i \text{ for some } \mathrm{k} \in \{1, \ldots, \mathrm{n}\}\right) \geq 1/\lambda$$

Where $\lambda > 0$. [Hint: Proceed as in the proof of Kolmogorov's inequalitiy]

[10]

5. In this problem, all the X_k are independent.

a) If $\lambda > 0$, show that

$$P\left(\max_{1 \leq k \leq n} |\mathrm{S_k}| \geq 3\lambda\right) \leq 3 \max_{1 \leq k \leq n} P\left(|S_k| \geq \lambda\right).$$

Here, $S_k = X_1 + \ldots + X_k$ for $k \geq 1$.

[8]

P.T.O.

b) If $S_n \to^P S$, show that $S_n \to S$ a.s. Here the S_n are as in (a).

[8]

c) State the three-series theorem of Kolmogorov.

[2]

d) If $E(X_n) = 0 \quad \forall n \geq 1$, and for some $c > 0$

$$\sum_{n=1}^{\infty} E(X_n^2 : |X_n| < c) < \infty, \quad \sum_{n=1}^{\infty} E(|X_n| : |X_n| \geq c) < \infty$$

show that $\sum X_n$ converges a.s.

[3]

e) If $X_n \geq 0 \quad \forall\, n \geq 1$, show that

$$\sum X_n \text{ converges a.s.} \Leftrightarrow \sum_{n=1}^{\infty} \left(P(X_n > 1) + E(X_n : X_n \leq 1)\right) < \infty$$

$$\Leftrightarrow \sum_{n=1}^{\infty} E\left(\frac{X_n}{1 + X_n}\right) < \infty.$$

[9]

6 a) Define the uniform integrability of a family of measurable functions. Show that, in a general measure space, it does not imply the ' integrability'

[1+2=3]

b) Show that, in a probability space, $\{X_t\}_{t \in T}$ is uniformly integrable iff both the following conditions (i) and (ii) hold :

(i) $\sup_{t \in T} E(|X_t|) < \infty.$

(ii) Given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$P(A) < \delta \Rightarrow \sup_{t \in T} E(|X_t| : A) < \varepsilon.$$

c) Show that if $\{X_n\}$ is uniformly integrabale, then so is $\{\overline{X}_n\}$. Here $\overline{X}_n = n^{-1}(X_1 + \dots + X_n)$, $n \geq 1$.

[5]

d) If $\{|X_n|^p\}$ is uniformly integrable for some p > 1, Show that $\{|\overline{X}_n|^p\}$ is also so. Here the $\overline{X}_n$ are as in (c)

$$\left[Hint: show\ that\ |\overline{X}_n|^p \leq n^{-1}\left(|X_1|^p + \dots + |X_n|^p\right)\right]$$

[5]

e) Let $X_n \to^d X$. Show that $\liminf \operatorname{var}(X_n) \geq \operatorname{var}(X)$.
Here, for any random variable Y, we define

P.T.O.

$$Var(Y) = E(Y-E(Y))^2 \quad if \quad E(|Y|) < \infty;$$
$$= \infty$$

[9]

7 a) State and prove the Helly-Bray First Lemma.

[1+7=8]

b) Let $X_1, X_2, \ldots$ be i.i.d. following the standard Exponential distribution. Show that

$$\limsup\left(\frac{X_n}{\log n}\right) = 1 \text{ a.s.}, \quad \liminf\left(\frac{X_n}{\log n}\right) = 0 \text{ a.s.}$$

[10]

c) Let $X_n \to^d X$ and $Y_n \to^d X$. If $P(X_n - Y_n < -\varepsilon) \to 0 \ \forall \varepsilon > 0,$

then show that $X_n - Y_n \to^P 0$. Give details.

[12]

8 a) Show that if each of the functions $\phi_1, \ldots, \phi_k$ is a characteristic function (of a probability measure) and $0 < \lambda_i < 1 \quad \forall i = 1, \ldots, k$ and $\sum_{i=1}^{k} \lambda_i = 1$, then $\sum_{i=1}^{k} \lambda_i \phi_i$ is also a characteristic function (of a probability measure). Hence deduce that the real part of a characteristic function is again a characteristic function. Give details.

[3+2=5]

b) Show that if ϕ_n is the characteristic function of X_n, $n \geq 1$ and $\varphi_n(t) \to \varphi(t) \ \forall t \in R$ and ϕ is continuous at $t = 0$, then ϕ is a characteristic function of a random variable X and $X_n \to^d X$.

[You may assume the unicity property of characteristic functions ; but you have to state and prove all the other result(s) you using]

[12]

c) Show that if ϕ is a characteristic function, then so is

$$e^{\lambda(\phi-1)}$$

for any $\lambda > 0$.

[5]

d) If $E(|X|^{n+\delta}) < \infty$ for some $n \geq 1$ and $0 < \delta' \leq 1$, then

$$\left| \phi_X(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} E(X^k) \right| \leq \frac{|t|^{n+\delta} E(|X|^{n+\delta}) 2^{1-\delta}}{(1+\delta)\ldots(n+\delta)}.$$

[8]

*****

INDIAN STATISTICAL INSTITUTE

M.Stat (1st year, B-stream) 2012-13

Semestral Examination

Applied Stochastic Process (Date of Exam. : 30/11/2012)

Time-3 hours Full Marks- 50

The questions carries a total of 55 marks. Attempt all questions. You can score a maximum of 50.

1.a)Let for an X-linked gene with alleles A and a the ratio AA : Aa : aa be r : 2s : t in females and the ratio A: a be p : q in males. Under the assumption of random mating through generations, find the imiting proportion of the genotypes.

b) Consider the model of self-fertilization through generations with an initial proportion

AA : Aa : aa = 1/3 : 1/3 : 1/3

Find how many generations will be needed to make the proportion of each homozygote greater than 49%

[9+6]

2.a) Assume that there are no deserts and there is full mobility for spread of epidemic in Neyman - Scott model in R^2. Show that the epidemic originating at u will get extinct eventually, either for all $u \varepsilon R^2$ or, for no $u \varepsilon R^2$.

b) In Neyman -Scott model let the mobility function be $f_u(x) = f(x-u)$ where f is p.d.f. of bivariate normal variable with parameters(0,0,1,1,0) and p.g.f. of number of infected persons from one infectious person at u is given by

$g(t|u)=(1/2)\, t+(1/2)\, t^2$ if $u_1 u_2 \neq 0$

$=(1/2) +(1/4)\, t+(1/4)\, t^2$ if $u_1 u_2 = 0$, where $u, x \varepsilon R^2$, $u=(u_1,u_2)$, $x=(x_1,x_2)$

What are the extinction probabilities of the epidemic for different $u \varepsilon R^2$?

[10+5]

3.a)Write down the model for a Simple Stochastic Epidemic, where X_t and Y_t denote the number of susceptible and the number of infected at time t. $X_0 = n$,$Y_0=1$ and $t \varepsilon [o,\infty)$.Let T denote the duration of the epidemic. Find the expression for E(T).

P.T.O.

b) Write down the deterministic model in general epidemic using Kermack-Mckendrick Equations with $x(t)$, $y(t)$, $z(t)$ denoting numbers of susceptible, infected and removals respectively.

Show that $y(t) \to 0$ as $t \to \infty$.

[9+9]

4. Consider the single server queuing model. Time = 0,1,2,.......... . At each time point the number of customers arriving follows Bernoulli (p). Service time has the probability generating function given by $(1/4)\, t + (1/4)\, t^2 + (1/4)\, t^3 + (1/4)\, t^4$

Determine the set of values of p for which the busy period will be finite with probably one.

[7]

##########################

Indian Statistical Institute
Backpaper
M.Stat. First Year
First Semester, 2012-13 Academic Year
Large Sample Statistical Methods

Total Marks: 100 Duration :- 4 hours

17 01 13

Answer all questions

1. State and prove the Glivenko-Cantelli Theorem. [20]

2. Suppose X_i, $i = 1, 2, \ldots$ are independent N(0,1) random variables defined on the same probability space. Show that the sequence can't converge in probability. [10]

3. State and prove the Weak Bahadur Representation of sample quantiles under iid sampling from a common distribution. [20]

4. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Poisson(λ), $\lambda > 0$. Can you find two sequences $a_n > 0$ and b_n of real constants (possibly dependent on λ) such that $a_n \left((1 - \frac{1}{n})^{n\bar{X}_n} - b_n\right)$ converges in distribution to a non-degenerate random variable? Prove your assertion. [10]

5. Give an example where the maximum likelihood estimator is inconsistent. Prove your answer. [10]

6. Suppose $X_1, \ldots, X_n$ are iid having density $f(x, \theta)$, where $\theta \in \mathbf{R}$. Invoking appropriate regularity assumptions, derive the asymptotic null distribution of the likelihood ratio test statistic for testing $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$. [10]

7. Suppose you have obtained an iid sample from a certain distribution F. Stating appropriate assumptions, prove the consistency of Bootstrap in estimating the distribution of a suitably scaled and centred version of sample mean. [20]

INDIAN STATISTICAL INSTITUTE
First Semester Back Paper Examination: 2012-13
M. Stat. I Year
Measure Theoretic Probability

Date: 18.01.2013 **Maximum Marks: 100** **Duration: 3 Hours**

Note : a) All proofs should be rigorous, complete and clear
b) Notations are as usual.
c) Maximum you can score is 100.

ANSWER ANY FOUR QUESTIONS

1 a) Let $f : \Omega \to \Omega'$ be a function, and let ζ be a non-empty family of subsets of Ω'. Show that

$$\sigma\left(f^{-1}(\zeta)\right) = f^{-1}\left(\sigma(\zeta)\right).$$

Hence deduce that if $\underline{D}$ is a non-empty family of subsets of Ω and $A \subset \Omega$, then $\sigma_A\left(\underline{D} \cap A\right) = \sigma_\Omega\left(\underline{D}\right) \cap A$.

[10+2=12]

b) Let $(\Omega, \mathscr{A}, P)$ be a probability space, Let $A_n \in \mathscr{A}$ $\forall_{n \geq 1}$. Show that

$$P(\liminf A_n) \leq \liminf P(A_n)$$
$$\leq \limsup P(A_n) \leq P(\limsup A_n)$$

State and prove the result(s) you are using. [8]

c) Let $(\Omega, \mathscr{A}, \mu)$ be a σ - finite measure space , and $\mathscr{U}$ a family of sets $A \in \mathscr{A}$ such that $\mu(A) > 0$ and whenever $A \neq B$, one has $\mu(A \cap B) = 0$. Show that $\mathscr{U}$ is constable.

[5]

2. a) *Let* $T : \Omega \to \Omega'$ be an $(\mathscr{U}, \mathscr{U}')$ measurable function. Let μ be a measure on $(\Omega, \mathscr{U})$. Let $f : \Omega' \to \overline{R}$ be a non-negative $(\mathscr{U}' \mathscr{B}(\overline{R}))$ measurable function. Show that

$$\int_{T^{-1}(A')} f(T(\omega))\, \mu(d\omega) \equiv \int_{A'} f(\omega')\, \mu T^{-1}(\omega')$$

for each $A' \in \mathscr{U}'$

[12]

b) **Let each** μ_n **be a measure on** $(\Omega, \mathscr{A})$. Let f : $\Omega \to \overline{R}$ be a non-negative $(\mathscr{A}\ \mathscr{B}(\overline{R}))$ - measurable function. Show that

$$\int f(\omega)\, \mu(d\omega) = \sum_{n=1}^{\infty} \lambda_n \int f(\omega)\, \mu_n(d\omega)$$

Where $\mu = \sum_{n=1}^{\infty} \lambda_n \mu_n$ with $\lambda_n \geq 0$ $\forall_n \geq 1$.

[8]

P.T.O

c) Let $f:\Omega \to \overline{R}$ be a non-negative function which is $(\mathscr{A}, \mathscr{B}(\overline{R})))$- measurable. Show that there exists a non-decreasing sequence $\{g_n\}$ of non-negative simple fuctions on R and such that $g_n(f(w)) \to f(w) \ \forall \omega \in \Omega$. [5]

3(a) Let X, Y be two independent random variables with the respective distribution functions F and G . Show that

$$E(F(Y)) + E(G(X)) = 1 + P(X = Y).$$ [8]

(b) Show that if the distribution function of $\times$ *is* F, then

$$E\,(F(\times)) = \frac{1}{2} + \sum_x (P(X = x))^2 ,$$

The summation being taken over all discontinuity points of F.

[5]

(c) Let $(\Omega, \mathscr{A}\ \mu)$ be a measure space . Let $f_n \to^{\mu} f$. Show that

(i) $\{f_n^2\}$ need not converge to f^2 in μ-measure ;

(ii) if g is uniformly continuous , then $g\,(f_n) \to^{\mu} g(f)$;

(iii) if g is continuous and $\mu(|f| \geq a) \to 0$ as $a \to \infty$, then $g\,(f_n) \to^{\mu} g\,(f)$

$[2+3+7=12]$

4(a). Show that if μ is σ - finite and $a > 0$,

$$\int |f|\, I_{[|f| \geq a]}\, d\mu = = \int_{[a,\infty)} \mu(|f| \geq x)dx + a\mu(|f| \geq a).$$

[6]

(b) Show that $\{X_t\}_{t\in T}$ is uniformly integrable on a probability space $(\Omega, \mathscr{A}, P)$ $\Leftrightarrow \sup_{t\in T} \int_{[a,\infty]} P(|X_t| \geq x)dx \to 0$ .as $a \to \infty$

[6]

(c) State and prove holder's inequality

[1+(5+1)=7]

(d) State and prove minkowski's inequality.

[1+5=6]

5 (a) State and prove helly's selection theorem .

[1+6=7]

(b) Let $\{\mu_n\}_{n\geq 1}$ be a sequence of sub- probability measures .

Show that $\{\mu_n\}_{n\geq 1}$ is vaguely convergent iff there exists a dense subset D of $\Re$ such that fir every $a_1, b \in D$ With a<b,

$\{\mu n\,((a,b])\}_{n\geq 1}$ converges.

[10]

(c) let $\mathscr{U} := \{\mu_\alpha\}\alpha \in A$ be a family of probability measures . Show that $\mathscr{U}$ is uniformly tight iff every sequence in $\mathscr{U}$ has a subsequence that is vaguely convergent to a probability measure.

[8]

P.T.O

6 (a) Let $\{X_n\}_{n\geq 1}$ be a sequence of independent L^2 - random variables . Show that if $\sum E(X_n)$ and $\sum$ var (X_n) are both convergent then $\sum X_n$ converges a.s.

The converse is true , provided the X_n are uniformly bounded ; show this .

[10+10=20]

b) Let $\{X_n\}_{n\geq 1}$ be a sequence of meen-zero random variables such that $\sup_{n\geq 1} E(X_n^4) < \infty$. If S_n and X_{n+1} are independent $\forall\, n \geq 1$, show that $n^{-1}S_n \to 0$ a.s. Here $S_n = X_1 + ... + X_n$, $n \geq 1$.

[5]

7(a) Show that if ϕ is a characteristic function, then so is Re (ϕ).

[4]

(b) State the inversion formula for a characteristic function. Deduce from it the inversion formula for the density of a probability measure.

[2+6=8]

(c) State and prove the Levy – Cramer continuity theorem.

[1+ (1+6) = 8]

(d) Show that a sequence of univariate normal distributions is uniformly tight iff the means and variances are both bounded.

[5]

*****

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2012 – 13

M. Stat (1st Year)

Time Series Analysis

Date: 18 February 2013 Maximum Marks: 30 Duration: 1½Hours

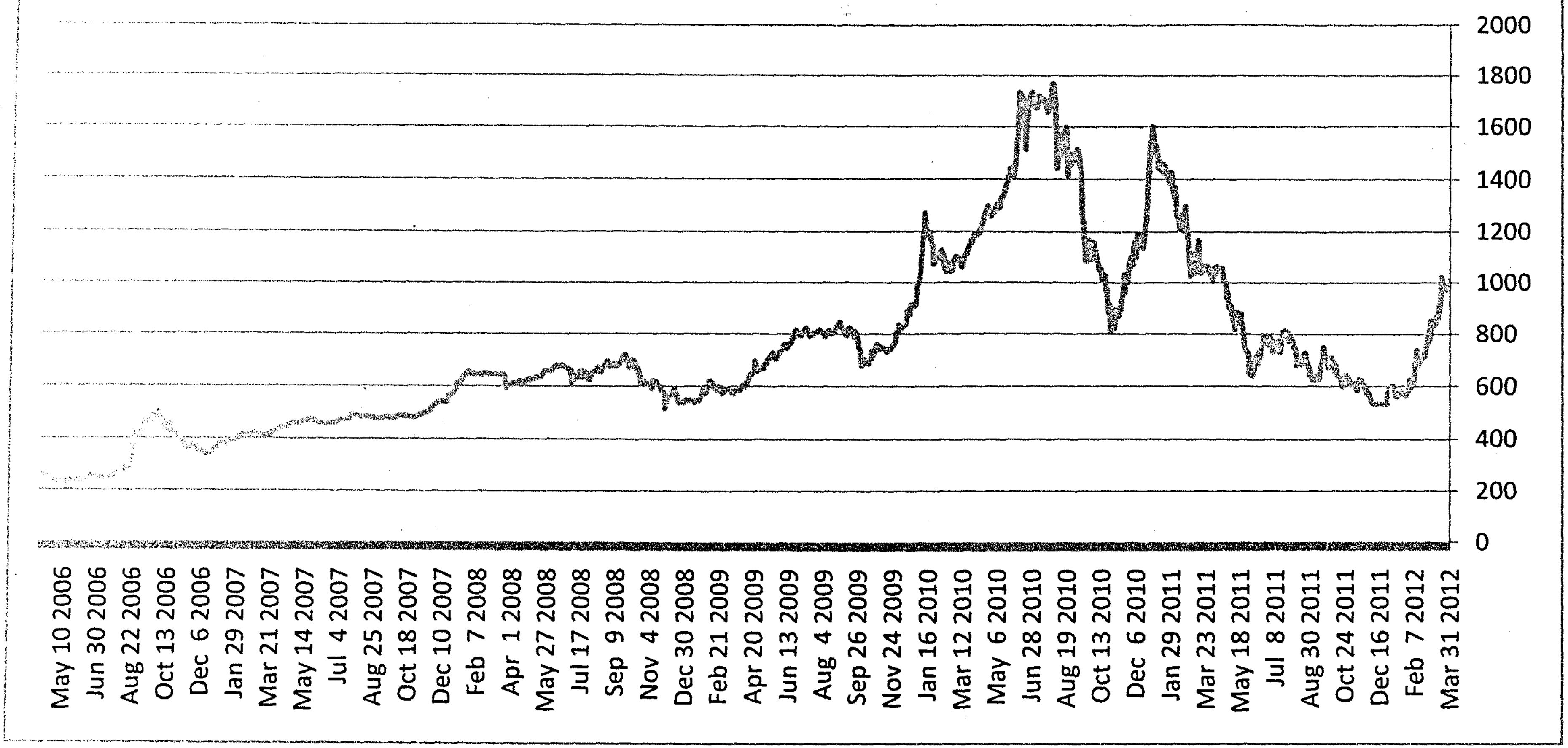
1. From the Scatter plot given on page 2, comment on the characteristics of the Cardamom Spot price data. Your answer should be relevant in the context of Time series analysis.

2. From the ACF and PACF graph and table on page 3, can you figure out the possible nature of the ARMA model that would fit this data? Justify your answer.

3. If a data generated by a Trend Stationary Process (TSP) is by mistake analysed using a Difference Stationary Process (DSP); analytically show the resultant problem in the modeling of the error term.

4. Compute the ACF and the PACF of the ARMA (1, 1) model.

Cardamom Spot Price(Rs.)



Page

Correlogram

```
==============================================================
 Autocorrelation Partial Correlation      AC       PAC
==============================================================
     .   |*****    |       .   |*****    |  1   0.652  0.652
     .   |***      |       .  *|   .     |  2   0.382 -0.075
     .   |*  .     |       .**|    .     |  3   0.095 -0.217
     .  *|   .     |       .  *|   .     |  4  -0.140 -0.170
      ***|   .     |        ***|   .     |  5  -0.423 -0.363
     ****|   .     |       .**|    .     |  6  -0.579 -0.229
     ****|   .     |       .   |*  .     |  7  -0.459  0.182
      ***|   .     |       .   |   .     |  8  -0.334 -0.053
     .  *|   .     |       .   |   .     |  9  -0.135  0.048
     .   |   .     |       .  *|   .     | 10  -0.008 -0.132
     .   |*  .     |       .  *|   .     | 11   0.143 -0.122
     .   |**.      |       .   |   .     | 12   0.248  0.047
==============================================================
The vertical lines are 95% confidence bounds
```

Indian Statistical Institute

Mid-Semester Examination 2012-13

Course: M. Stat. I Year Subject: Optimisation Techniques

Date: 21·02·13 Maximum Marks : 30 Duration: 2hrs.

Attempt all the questions. The maximum you can score is 30.

1. Consider the linear programming problem: Find y_1 and y_2 to minimise $y_1 + y_2$ subject to
   $y_1 + 2y_2 \geq 3$
   $2y_1 + y_2 \geq 5$
   $y_2 \geq 0$
   Solve the problem graphically. (4)
2. Obtain the dual to the problem in question 1 above. Solve the dual using duality theorem. (3+3)

3. Consider the problem minimise $8x_1 - 100x_2$ such that
   $2x_1 - 10x_2 \geq 2$
   $x_1 - 11x_2 \geq -3,\ x_1, x_2 \geq 0.$
   Write down the dual of the problem and solve. Using duality theory what conclusion can be made about the primal problem. Verify your conclusion. (1+3+2+3)
4. The simplex table for a maximisation problem is given by

z	x_1	x_2	x_3	s_1	s_2	C
1	-5	-2	-2	0	0	0
0	1	2	-2	1	0	30
0	1	3	1	0	1	36

Which variable should come in the next table and why? Which one should go out and why? (4)

5. Consider the following balanced transportation problem.

	D_1	D_2	D_3	D_4	
S_1	9	8	5	7	12
S_2	4	6	8	7	14
S_3	5	8	9	5	16
	8	18	13	3	42

Find an initial basic feasible solution using Vogel's approximation method and check whether it is optimal or not. (7+5)

6. Find the minimum assignment cost from the following cost matrix. (5)

	I	II	III	IV
A	9	6	6	5
B	8	7	5	6
C	8	6	5	7
D	9	9	8	8

7. Consider the following game with payoffs

			Player B
		H	T
Player A	H	-1,1	1,-1
	T	1,-1	-1,1

Find the Nash equilibrium (in pure or mixed strategy), if any, for the above game. (5)

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : 2012-13(First Semester)

M. Stat. I year

Discrete Maths.

Date: 22.02.13 Maximum Marks : 80 Duration : 2 1/2 hrs

Note: Answer as many as you can. The maximum you can score is 80. Notation is as used in the class.

1. (a) Let a_n denote the number of ways of multiplying to form the product $x_1 * x_2 * \ldots * x_n$. For instance

$$x_1 * x_2 * x_3 = (x_1 * (x_2 * x_3)) = ((x_1 * x_2) * x_3)$$

and hence $a_3 = 2$. Find a recurrence relation for the sequence a_n and solve it.

(b) Define the Stirling number of the first kind $s(n, k)$. Show that

$$[x]^n = \Sigma_{k=1}^{n} |s(n,k)| x^n.$$

Hence, or otherwise, find the expected number of cycles in a randomly chosen permutation of S_n

[10+10]

2. Let D_n denote the number of derangements of $1, 2, \ldots, n$. Derive a recurrence relation for D_n. [5]

3. (a) Consider the recurrence

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_p a_{n-p}, \quad n \geq p,$$

where $a_i \in I\!R$ and $a_p \neq 0$. What is its characteristic polynomial? Show that if α is a root of this polynomial, then $\{\alpha^n\}$ is a solution of the recurrence relation

Also show that the set of solutions of the recurrence relation forms a vector space over $I\!R$

(b) Solve the Fibonacci sequence by the method of characteristic equations.

(c) If, for $n \geq 0$,

$$C_{n+1} = 2nC_n + 2C_n + 2,$$

and $C_0 = 1$, find C_n

[6+5+7]

4. State the Inclusion-Exclusion Principle.
Given a graph G, define its Chromatic Polynomial $P(G, x)$.
Show that the Chromatic Polynomial is a monic polynomial of degree n, where n is the number of vertices of G.
Find the chromatic polynomial of a tree with n vertices. [3+7+4]

5. Let $p(n)$ denote the number of partitions of n. Find the generating function for the sequence $p(n)$.

Let $p(n, k)$ denote the number of partitions of n into (exactly) k parts. Derive the following recurrence relation

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

Show that the number of (non-congruent) triangles with integer sides and perimeter $2n$ is equal to $p(n, 3)$ [7+6+5]

6. Given integers $p, q \geq 2$ define the Ramsey number $R(p, q)$.
Show directly that a graph with six vertices either contains a triangle or an independent set of three vertices.
Show that $R(3, 4) = 9$. [2+5+6]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: (2012–2013)

M. Stat First Year

Regression Techniques

Date: 26/02/2013 Marks: 30 Duration: 3 hours

Attempt all questions

1. (a) Is the sum of the residuals always zero? Justify.

(b) Consider the following data set: $(X_1, Y_1) = (2.5, 6)$, $(X_2, Y_2) = (4, 9)$. If we wish to fit one of the two models given by (i) $Y = \beta_0 + \beta_1 X + \epsilon$ and (ii) $Y = \beta_1 X + \beta_2 X^2 + \epsilon$, which of the two models is expected to perform better? Justify your answer.

(c) Show that in the presence of pure error the square of the multiple correlation coefficient $R^2 < 1$.

(d) Consider the regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ (assume that it contains an intercept), where $\boldsymbol{\beta}$ is a $p \times 1$ vector and $\mathbf{X}$ is $n \times p$ and has full rank. Consider testing $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, where $\mathbf{A}$ is $q \times p$ $(q < p)$, has full rank and does not involve the intercept β_0. Show geometrically that the F-statistic for testing this hypothesis can be written in terms of the difference between the R^2-statistics, where R^2 stands for the coefficient of determination.

Marks: 2+2+2+4=10

2. (a) Consider fitting the linear model having the form $y_i = \sum_{j=1}^{p} \beta_j x_{ij} + \epsilon_i$; $i = 1, \ldots, n$. Assume that, for $j, k = 1, \ldots, p$, $\sum_{i=1}^{n} x_{ij} x_{ik} = n\delta_{jk}$, where $\delta_{jk} = 1$ if $j = k$ and 0 otherwise. Suppose that p_0 denotes the true model dimension, and $\tilde{\lambda}$ is the minimizer of $\tilde{P}(\lambda) = n^{-1} RSS(\lambda) + n^{-1} a \lambda \sigma^2$. Choose a appropriately such that $\tilde{\lambda} \xrightarrow{P} p_0$. Provide the motivation behind your choice and prove the convergence result.

(b) Assume that there are infinitely many non-zero β_j's in the model $y_i = \sum_{j=1}^{\infty} \beta_j x_{ij} + \epsilon_i$, $i = 1, \ldots, n$, where $\epsilon_i; i = 1, \ldots, n$, are normal random errors, $\sum_{j=1}^{\infty} \beta_j^2 < \infty$ and, for each n, for $(j, k) = 1, \ldots, n$, $\sum_{i=1}^{n} x_{ij} x_{ik} = n\delta_{jk}$, where $\delta_{jk} = 1$ if $j = k$ and 0 otherwise. Let $\hat{\lambda}$ be the minimizer of $\hat{P}(\lambda) = n^{-1} RSS(\lambda) + 2n^{-1} \sigma^2 tr(S_\lambda)$ over $\Lambda_n = \{1, \ldots, a_n\}$ with $a_n \to \infty$ as $n \to \infty$. Show that $P(\hat{\lambda} \le x) \to 0$ for any finite x.

Marks: 5+5=10

3. (a) Consider the case where $\mathbf{y}$, conditional on $\boldsymbol{\beta}$, has an n-variate normal distribution with mean $\mathbf{X}\boldsymbol{\beta}$ and variance-covariance matrix $\sigma^2 \mathbf{I}$. Let

$\boldsymbol{\beta}$ be p-variate normal with mean $\mathbf{0}$ and variance-covariance matrix $\sigma_s^2\mathbf{I}$. Show that the ridge estimator $\mathbf{b}_\lambda$ with $\lambda = \sigma^2/\sigma_s^2$ is the mean of $\boldsymbol{\beta}$ given $\mathbf{y}$.

(b) In the context of ridge regression set up, establish the existence of biased estimators which have smaller risks than unbiased ones.

Marks: 5+5=10

INDIAN STATISTICAL INSTITUTE

M-Stat I, Second Semester, 2012-13

MULTIVARIATE ANALYSIS

Date: 01.03.2013 Midsemester Examination Time: 3 hours

(Total Point 84. Answer as many as you can. Maximum you can score is 80)

1. Let $S \sim W_p(k, \Sigma)$ (p dimensional central Wishart with degrees of freedom k, and parameter Σ). Let $S^{-1} = (s^{ij})_{p\times p}$ represent the inverse of $S = (s_{ij})_{p\times p}$; similarly $\Sigma^{-1} = (\sigma^{ij})_{p\times p}$ be the inverse of $\Sigma = (\sigma_{ij})_{p\times p}$. Show that

   (a) $\frac{\sigma^{pp}}{s^{pp}} \sim \chi^2(k-p+1)$, and is independent of (s_{ij}), $i, j = 1, \ldots, p-1$.

   (b) $\frac{l^T\Sigma^{-1}l}{l^TS^{-1}l} \sim \chi^2(k-p+1)$ for any fixed vector l. [7+7=14]

2. Suppose that bivariate observations are available on individuals from three different treatment groups. The data are given below:

$$\text{Treatment 1}: \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$\text{Treatment 2}: \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

$$\text{Treatment 3}: \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

   Write down the appropriate hypothesis to be tested for the one way MANOVA model and perform the test; use $\alpha = 0.05$. Clearly state the necessary assumptions for this procedure. [14]

3. Let $\boldsymbol{R}$ be the correlation matrix of the $p \times 1$ random vector $\boldsymbol{X}$, and consider the evaluation of the principal components of $\boldsymbol{X}$ based on this correlation matrix. An interesting feature of these principal components is that they depend on the ratios of the correlations, and not on their absolute values. To see this, divide all the off-diagonal elements of the correlation matrix $\boldsymbol{R}$ by a constant k such that $k > 1$. Call the resulting matrix $\boldsymbol{R}^*$.

   (a) Show that the $\boldsymbol{R}$ and $\boldsymbol{R}^*$ matrices have different eigen values, but the same eigen vectors, so that the principal components do not change.

   (b) How are the eigen values of $\boldsymbol{R}$ related to those of $\boldsymbol{R}^*$?

   (c) What could be a possible problem in choosing $k < 1$ in the above formulation?

[7+7+2=16]

4. Consider the bivariate normal distribution with parameters $\boldsymbol{\mu} = 0$ and Σ, and consider the derivation of the principal components for a random variable $\boldsymbol{X}$ having the above distribution. A constant density ellipse for this model has the form

$$\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x} = c^2,$$

where c represents the appropriate constant. Show that finding the principal components of $\boldsymbol{X}$ corresponds to rotating the rectangular axes to make them lie along the major axis, minor axis directions of the above ellipse. [12]

5. Suppose that a p_i dimensional multivariate normal random vector X_i has a positive definite covariance matrix Σ_{ii}, $i = 1, 2$, such that

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

is the covariance matrix of the augmented vector $X = (X_1^T, X_2^T)^T$.

Recall that X_1 and X_2 are independent if and only if $\Sigma_{12} = 0$. Argue how a test for the null hypothesis that X_1 and X_2 are independent can be constructed in terms of the sample version of the first canonical correlation. [14]

6. (a) Let

$$\rho_{12} = \begin{bmatrix} \rho & \rho \\ \rho & \rho \end{bmatrix}, \quad \rho_{11} = \rho_{22} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

represent the components of the correlation matrix of $X = (X_1^T, X_2^T)^T$ where $X_1$and X_2 each have two components. Determine the canonical variables corresponding to all non zero canonical correlations.

(b) Generalize the result to the case where X_1 has p components and X_2 has $q \geq p$ components. [7+7=14]

MStat-I: Metric Topology and Complex Analysis (Midterm Exam)

04 March, 2013.

Duration: **Two Hours (i.e. 02.30pm to 04.30pm).**

Each question carries five marks. Solve any eight of them.

1. Define equivalent metric. Prove that there is a homeomorphism between $\mathbb{R}^n$ with Euclidean metric and $\mathbb{R}^{n-1} \times \mathbb{R}$ with prduct metric of Euclidean metric on $\mathbb{R}^{n-1}$ and absolute metric on $\mathbb{R}$.

2. Write down equivalant definition of compact in terms of cluster point, complete and sequence. Explain all terms used. Prove that if X is not compact, there exists a real valued function on X which is unbounded.

3. Define continuous map. $A \subseteq X$. Then $d(\_, A) \colon X \longrightarrow \mathbb{R}$ is a continuous map.

4. Let X be a compact metric. Prove that $B(X)$= set of bounded real valued functions is a metric space w.r.t. sup norm $|| \cdot ||_\infty$. Check whether $(C(X), || \cdot ||_\infty)$ is subspace of $B(X)$ under sup norm or not. If yes, check whether $C(X)$ is open in $B(X)$ under sup norm.

5. Let (X, d) be a metric space. We know that $\delta : X \times X \longrightarrow \mathbb{R}$ defined by $\delta(x, y) = \min\{1, d(x, y)\}$ defines a metric on X. Define complete metric space and prove that (X, d) is complete if and only if (X, δ) is complete.

6. Define a metric on a set. Prove that metric function $d \colon X \times X \longrightarrow \mathbb{R}$ is a continuous function where $X \times X$ is equipped with product metric i.e. $d((x_1, y_1), (x_2, y_2)) = \max\{d(x_1, x_2), d(y_1, y_2)\}$.

7. Prove that a compact subset A of a metric space (X, d) is closed in (X, d). Prove that if A, B are both compact subsets of a metric space

(X, d) such that $A \cap B = \phi$ then we can find U_A open set containing A and U_B open set containing B such that $U_A \cap U_B = \phi$.

8. Define connected metric space. Let $\{A_i | i \in I\}$ be a collection of connected subsets of a metric space X with the property that for all $i, j \in I$ we have $A_i \cap A_j \neq \phi$. Then $A = \cup_{i \in I} A_i$ is connected.

9. Define Cauchy sequence. Let d_1, d_2 be two metric on X such that there exist $a, b > 0$ so that for all $x, y \in X$, $ad_1(x, y) \leq d_2(x, y) \leq bd_1(x, y)$. Then prove that (X, d_1) and (X, d_2) have same Cauchy sequences.

10. Consider $(C[0, 1], d)$ where $d(f, g) = \displaystyle\int_0^1 |f(t) - g(t)| dt$. Prove that $(C[0, 1], d)$ is a metric space. Check whether $(C[0, 1], d)$ is complete or not.

INDIAN STATISTICAL INSTITUTE
Second Semestral Examination : 2012-13
M. STAT. I YEAR
Optimization Techniques

Date: 19 April 2013 | Maximum Marks: 100 | Duration: 3 hours

There are two groups: A and B. *Use separate answer booklet for each group.* In Group A, you can attempt questions of 30 marks only. You may attempt all of Group B that has 80 marks, but maximum you can score is 70.

Group A

(Answer any **three** of the following questions.)

1 Prove that if both the primal and the dual LPP have feasible solutions, then they both have optimal solutions. [10]

2 Establish the relation between extreme points and basic feasible solutions in the context of a LPP. [10]

3 Set up a transportation problem of your choice. Form its dual and interpret the dual problem. [10]

4 Show that any two person zero sum game can be looked upon as a standard LPP and its dual. Hence prove the minimax theorem. [10]

Group B

1 (a) An auctioneer contemplates selling m items. Let $M = \{1, 2, \ldots, m\}$ be the set of items. From among the prospective buyers, n bids are received. A bid is represented by $B_j = (S_j, p_j)$ for $j = 1, 2, \ldots, n$, where S_j is the specified set of items the bidder wishes to acquire by paying the price p_j, $S_j \neq \phi$ and $S_j \subseteq M$. For example, $B_1 = (\{1, 3, 5\}, 13)$, $B_2 = (\{1, 2, 5\}, 20)$, and so on. The auctioneer intends to select the bids in order to maximize his/her revenue. Formulate this as an optimization problem.

P.T.O.

(b) Consider the following optimization problem:

$$\max\ z = x_1 + 2x_2$$

subject to

$$\begin{aligned} x_1 + x_2 &\leq 8, \\ -x_1 + x_2 &\leq 2, \\ x_1 - x_2 &\leq 4, \\ x_1 &= 0, 1, 4 \text{ or } 6, \\ x_2 &\geq 0 \text{ and integer.} \end{aligned}$$

(*i*) Reformulate it as an equivalent integer linear programming problem.

(*ii*) How would you change the answer to part (*i*) if the objective function were changed to: $z = x_1^2 + 2x_2$?

(c) Consider the least-cost assignment problem with the following cost-matrix. Solve it by branch-and-bound method.

		Persons			
		A	B	C	D
	1	94	74	62	11
Jobs	2	1	10	88	74
	3	54	88	8	81
	4	68	82	76	21

[5 + (5+5) + 15 = 30]

2 (a) For $k > 0$, consider the non-linear programming problem:

$$\max\ z = x_1 x_2^3 x_3$$

subject to

$$\begin{aligned} x_1 + x_2 + x_3^3 &\leq k, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

(*i*) Formulate recursive relations of dynamic programming for this problem.

(*ii*) Solve the problem using the relations of (*i*).

(b) Given functions $g_1(x), g_2(x), \ldots, g_N(x)$, consider the optimization problem:

$$\min\ z = \max\{g_1(x_1), g_2(x_2), \ldots, g_N(x_N)\}$$

subject to

$$\sum_{j=1}^{N} x_j \leq b,$$
$$x_j \geq 0 \; \forall j.$$

- Formulate the above as a dynamic programming problem.

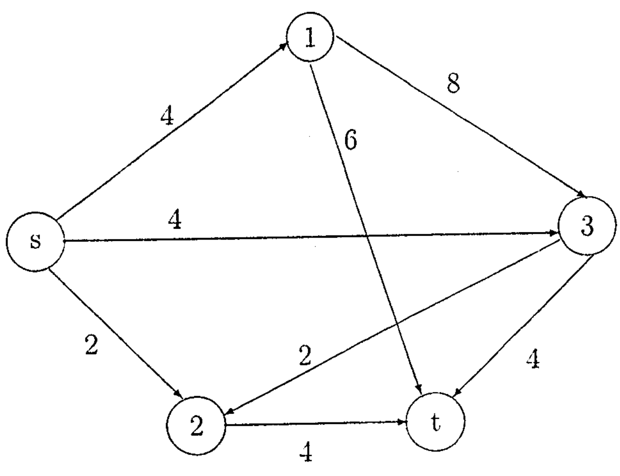
[(5+10) + 10 = 25]

3 (a) Consider the setup of Max-flow problem, and prove the following.

(*i*) A cut $(X, \bar{X})$ is minimal if and only if for every maximal flow f, $f(X, \bar{X}) = c(X, \bar{X})$ and $f(\bar{X}, X) = 0$. [$c(X, \bar{X})$ is the capacity of $(X, \bar{X})$.]

(*ii*) Let $(Y, \bar{Y})$ be any minimal cut, f be a maximal flow, and $(X, \bar{X})$ be the cut (relative to f) as defined in the proof of Max-flow Min-cut theorem. Then $X \subseteq Y$.

(b) Solve the Max-flow problem for the following network (the number adjacent to an arc being its capacity) by labeling algorithm. What is the corresponding minimal cut?



[(5+10) + 10 = 25]

_____*** xXx ***_____

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2012-13(Second Semester)

M. Stat. I year

Discrete Maths.

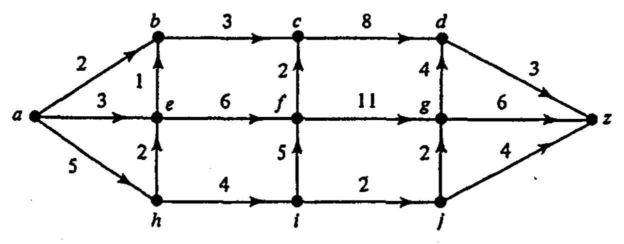
Date: 22.04.13 Maximum Marks : 90 Duration : 3 hrs

Note: Answer as many as you can. The maximum you can score is 90. Notation is as used in the class.

1. State Polya's theorem.
   Suppose we label the vertices of a graph G with n vertices with the labels $1, 2, \ldots, n$. Any labeling of G can be thought of as a permutation of $\{1, 2, \ldots, n\}$ if we start with a fixed labeling. Every automorphism π can be thought of as taking a labeling σ to another labeling $\pi \circ \sigma$. Show that the number of distinct labelings is $n!/|\Gamma(G)|$, where $\Gamma(G)$ is the automorphism group of G. Verify it with the circuit C_4. [3+7=10.

2. (i) Write down Kruskal's algorithm for finding a minimum weight spanning tree of a weighted graph. What is its worst case time complexity? Show that if the graph with n vertices is disconnected, then the algorithm halts without finding a subgraph T with $n-1$ edges.

   (ii) In the directed network below, demostrate Dijkstra's algorithm to find a shortest path from a to z.



[8+7=15]

3. (i) Let $n > 0$ be an odd composite integer. With usual notation, define

$$G(n) = \{a \in Z_n^* : \left(\frac{a}{n}\right) = a^{\frac{n-1}{2}} \bmod n\}.$$

Show that $G(n)$ is a subgroup of Z_n^*.
Suppose that $n = p_1 p_2 ... p_k$, where the p_i's are distinct primes. Show that in this case, $G(n) \neq Z_n^*$ by producing an integer a such that $\left(\frac{a}{n}\right) = -1$ but $a^{\frac{n-1}{2}} \neq -1 \bmod n$. Conclude that $|G(n)| \leq \frac{n-1}{2}$.

(ii) Write down the Solovay-Strassen primality testing algorithm.

(iii) Assume that $|G(n)| \leq \frac{n-1}{2}$ in all cases. Consider the two events:
$\mathbf{E_1}$: a random odd integer n of a specified size is composite
$\mathbf{E_2}$: The algorithm answer "n is prime" m times in succession
Assume that $\Pr[\mathbf{E_1}]$ is approximately $1 - \frac{2}{\log n}$. Find a non-trivial bound for $\Pr[\mathbf{E_1}|\mathbf{E_2}]$. [10+5+7=22]

4. (i) Label the vertices of a tree with n vertices with the labels $1, 2, \ldots, n$. Establish a 1-1 correspondence with such trees and sequences $(b_1, b_2, \ldots, b_{n-2})$, $1 \leq b_i \leq n$, of labels. Show that a vertex v appears m times in the sequence $(b_1, b_2, \ldots, b_{n-2})$ iff $deg(v) = m+1$

(ii) Let u_n denote the number of unlabeled rooted trees with n vertices. Find a recurrence relation for u_n. Hence compute u_4. [8+6=14]

5. Let $Q(x_1, \ldots, x_n) \in \mathbf{F}[x_1, \ldots, x_n]$ be a non-zero multivariate polynomial of total degree d. Fix a finite set $S \subseteq \mathbf{F}$. Let $r_1, \ldots, r_n$ be chosen independently and uniformly at random from the set S. Show that

$$Pr[Q(r_1, \ldots, r_n) = 0] \leq \frac{d}{|S|}.$$

Hence, construct a randomized algorithm for testing whether a bipartite graph $G(U, V, E)$, where U, V are independent sets of n vertices, has a perfect matching. [10]

6. (i) Let A be the adjacency matrix of a k-regular graph. Show that k is an eigenvalue whose multiplicity equals the number of connected components.

(ii) Define a strongly regular graph. Define a Paley graph $P(q)$, where q is a power of a prime. Show that it is strongly regular and find its parameters.

(iii) Show that a strongly regular graph on $p+1$ vertices is imprimitive, where p is prime.

(iv) When is an orthogonal array $OA(k,n)$ said to be extendible? Let G be the graph corresponding to an $OA(k,n)$. Show that the chromatic number $\chi(G) = n$ iff $OA(k,n)$ is extendible. [6+7+4+7=24]

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2012–13

M. Stat (1st Year)

Time Series Analysis

Date: 24 April 2013 | Maximum Marks: 100 | Duration: 3 Hours

Answer ALL questions. Marks allotted to each question are given within parentheses

1. From the Scatter Plot given on page 3, comment on the characteristics (Trend, Seasonality and Volatility) of the US Dollar : Rupee exchange rate data. Your answer should be relevant in the context of time series analysis.
[5 + 3 + 4 = 12]

2. From the ACF and PACF graphs and tables on pages 4 & 5, can you figure out the possible nature of the ARMA model that would fit this inflation data? Justify your answer. [12]

3. Find the ACF and PACF for the following SARMA model:
$$(1 - \phi_S L^{12})(1 - \phi_1 L - \phi_2 L^2)x_t = (1 + \theta_S L^{12})\epsilon_t$$
where ϵ_t is white noise. [6 + 6 = 12]

4. (a) Obtain the spectral density for the ARMA (1, 1) process with usual notations.
(b) Demonstrate the linkage between the spectral density and the autocorrelation function. [8 + 6 = 14]

5. Consider the time series data $\{x_1, x_2, \dots, x_{21}\}$. For this data consider a spectral model with three sine and cosine components each. Write down the Spectral ANOVA table for this model. [12]

6. Obtain the unit root test statistic for an $ARMA(2,2)$-type model given by:

$$(1 - \phi_1 L - \phi_2 L^2)x_t = (1 + \theta_1 L + \theta_2 L^2)\epsilon_t$$

where $\epsilon_t \sim N(0, \sigma_e^2)$ is white noise.

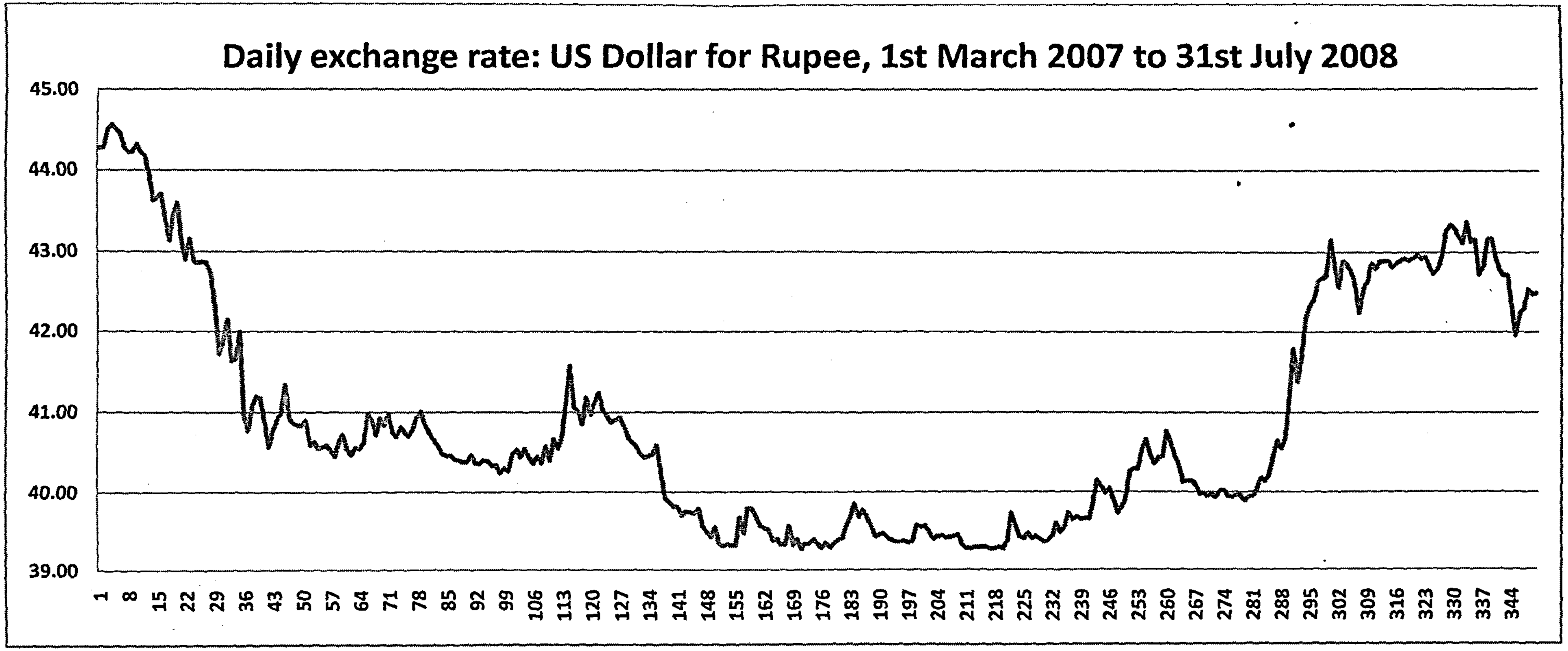
Indicate the kind of modifications needed if we introduce conditional heteroscedasticity in the model, i.e. $\epsilon_t \sim N(0, \sigma_t^2)$? [8 + 4 = 12]

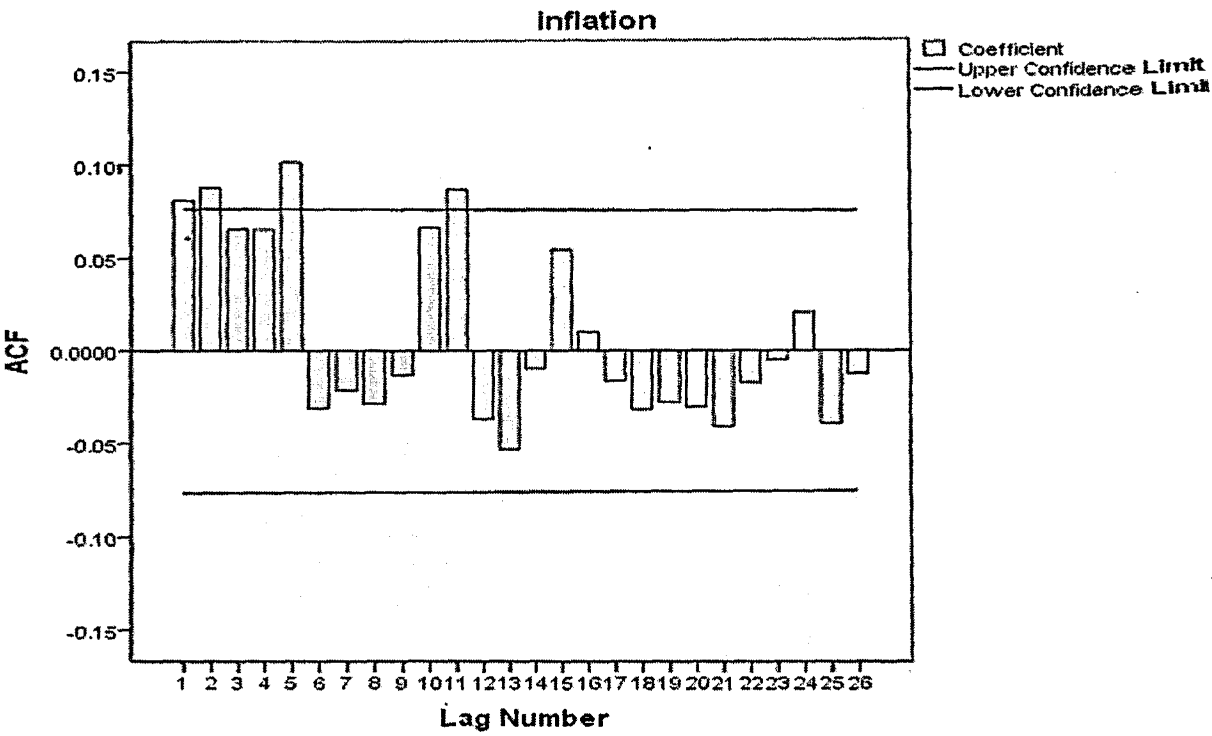
7. Define the GARCH model along with all its assumptions.
What are the causality conditions for GARCH (2, 1)? Give explanations for your answer. [4 + 8 = 12]

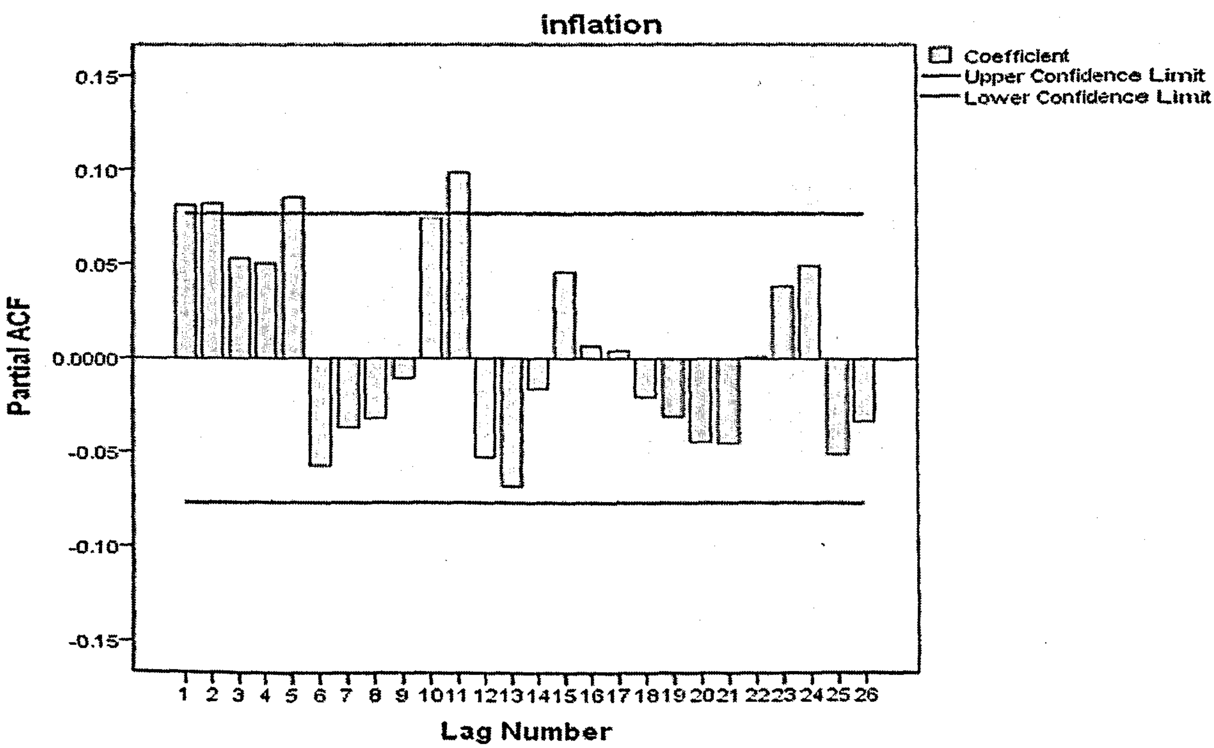
8. For the state space model:

$$x_t = \phi x_{t-1} + \omega_t$$
$$y_t = 3x_t + v_t$$

Write down the expressions for the Kalman **filter** and **gain**, the **innovations** equation and the **covariance** expression using the standard notations.
[4 + 3 + 3 + 4 = 14]


Daily exchange rate: US Dollar for Rupee, 1st March 2007 to 31st July 2008
45.00
44.00
43.00
42.00
41.00
40.00
39.00
1
8
15
22
29
36
43
50
57
64
71
78
85
92
99
106
113
120
127
134
141
148
155
162
169
176
183
190
197
204
211
218
225
232
239
246
253
260
267
274
281
288
295
302
309
316
323
330
337
344


inflation
Coefficient
Upper Confidence Limit
Lower Confidence Limit
ACF
0.15
0.10
0.05
0.0000
-0.05
-0.10
-0.15
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Lag Number


inflation
Coefficient
Upper Confidence Limit
Lower Confidence Limit
Partial ACF
0.15
0.10
0.05
0.0000
-0.05
-0.10
-0.15
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Lag Number

Autocorrelation
Series:inflation

Lag	Autocorrelation	Std. Error	Box-Ljung Statistic		
			Value	df	Sig.
1	.081	.038	4.481	1	.034
2	.088	.038	9.753	2	.008
3	.065	.038	12.684	3	.005
4	.065	.038	15.607	4	.004
5	.101	.038	22.639	5	.000
6	-.031	.038	23.300	6	.001
7	-.021	.038	23.616	7	.001
8	-.029	.038	24.185	8	.002
9	-.013	.038	24.302	9	.004
10	.066	.038	27.358	10	.002
11	.087	.038	32.563	11	.001
12	-.037	.038	33.518	12	.001
13	-.053	.038	35.473	13	.001
14	-.009	.038	35.534	14	.001
15	.055	.038	37.609	15	.001
16	.010	.038	37.680	16	.002
17	-.016	.038	37.868	17	.003
18	-.032	.038	38.565	18	.003
19	-.028	.038	39.113	19	.004
20	-.030	.038	39.765	20	.005
21	-.041	.038	40.944	21	.006
22	-.017	.038	41.148	22	.008
23	-.005	.038	41.165	23	.011
24	.021	.038	41.473	24	.015
25	-.040	.038	42.579	25	.016
26	-.012	.038	42.687	26	.021

Partial Autocorrelation
Series:inflation

Lag	Partial Autocorrelation	Std. Error
1	.081	.038
2	.082	.038
3	.053	.038
4	.050	.038
5	.085	.038
6	-.057	.038
7	-.037	.038
8	-.032	.038
9	--.011	.038
10	.074	.038
11	.098	.038
12	-.052	.038
13	-.068	.038
14	-.016	.038
15	.046	.038
16	.007	.038
17	.004	.038
18	-.021	.038
19	-.031	.038
20	-.044	.038
21	-.045	.038
22	.001	.038
23	.038	.038
24	.049	.038
25	-.051	.038
26	-.033	.038

INDIAN STATISTICAL INSTITUTE
Semestral Examination: 2012-13 (Second Semester)

M.STAT. I YEAR
Metric Topology and Complex Analysis

Date: 27.04.2013 | Maximum Marks: 60 | Duration: Three hours

ch question carries ten marks.

1. (i) Let $f, g\colon (X, d_X) \longrightarrow (Y, d_Y)$ be continuous.
Prove that $A = \{x \in X | f(x) = g(x)\}$ is closed in X. [3 marks]
(ii) Let $M(n, \mathbb{R})$ denote the metric space of $n \times n$ matrices over $\mathbb{R}$. The metric d denotes pull back metric on $M(n, \mathbb{R})$ obtained from $\mathbb{R}^{n^2}$ using bijection between $M(n, \mathbb{R})$ and $\mathbb{R}^{n^2}$. Consider $p\colon (M(n, \mathbb{R}), d) \longrightarrow (M(n, \mathbb{R}), d)$ be given by $p(A) = P(A, A^t)$ where P is a polynomial in two variables with real coefficients and A^t denote transpose of a matrix. Prove that p is continuous. [7 marks]

2. (i) Define entire function. [1 mark]
(ii) Write down Cauchy integral formula for higher derivatives and explain terms involved. [3 marks]
(iii) Prove that if f is an entire function which satisfies $|f(z)| \leq A|z|^k$ for all $z \in \mathbb{C}$, where A is a positive real number and $k \in \mathbb{N}$ then f is a polynomial of degree not exceeding k. [6 marks]

3. (i) Write down statement of Rouche's formula and explain terms involved. [4 marks]
(ii) Hence or otherwise prove open mapping theorem. [6 marks]

4. (i) Define winding number $W_\alpha(z)$ of a loop α around a point z_0. Prove that if α is a closed curve and z, w are in same connected component of $\mathbb{C} \setminus \alpha$ then $W_\alpha(z) = W_\alpha(w)$. [5 marks]
(ii) Let $\alpha\colon [0, 1] \longrightarrow \mathbb{C}$ be a curve given by $e^{2\pi i t}$ for $0 \leq t \leq \frac{1}{2}$ and $e^{-2\pi i t}$ for $\frac{1}{2} \leq t \leq 1$. Then calculate $W_\alpha(-0.74138734)$. [5 marks]

5. (i) Write down the definition of topology on a set. [2 marks]
(ii) What do we mean by discrete topology on a set? [2 marks]
(iii) Let σ be a topology on $\mathbb{R}$ such that every interval of the form $[x, y)$ is in σ where $x < y$. If $a\colon (\mathbb{R}, \sigma) \longrightarrow (\mathbb{R}, \sigma)$ defined by $a(x) = -x$ is continuous then prove that σ is the discrete topology on $\mathbb{R}$. [6 marks]

6. (i) When do we say two spaces are homotopy equivalent? [2 marks]
(ii) Check (and explain) whether S^1 with induced metric from $\mathbb{R}^2$ and $A(a, b) = \{z \in \mathbb{C} | a < |z| < b\}$ where $a < b$ with induced metric from $\mathbb{C}$ are homotopy equivalent or not. [4 marks]
(iii) Let U be a connected open subset of $\mathbb{C}$. If $x, y \in U$ then prove that we can find a piecewise smooth path joining x and y. [4 marks]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2012–2013)

M. Stat First Year

Regression Techniques

Date: 30.4.13 Full Marks: 100 Duration: 3 hours

Attempt all questions

[You may use the following result if necessary, without proof: If $\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$, then $\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{pmatrix}$ where $\mathbf{A}^{11} = \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} = \mathbf{A}_{11}^{-1} - \mathbf{A}^{12}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$, $\mathbf{A}^{12} = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}^{22}$, $\mathbf{A}^{22} = \left(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1}$, $\mathbf{A}^{21} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}^{11}$.]

1. Let the full model $\mathbf{y}^{(n\times 1)} = \mathbf{X}_1^{(n\times \overline{k+1})}\boldsymbol{\beta}_1^{(\overline{k+1}\times 1)} + \mathbf{X}_2^{(n\times \overline{m-k})}\boldsymbol{\beta}_2^{(\overline{m-k}\times 1)} + \boldsymbol{\epsilon}^{(n\times 1)}$ be the true regression model. The components of ϵ are distributed independently as $Normal(0, \sigma^2)$.

   Also, consider the reduced model: $\mathbf{y}^{(n\times 1)} = \mathbf{X}_1^{(n\times \overline{k+1})}\boldsymbol{\theta}_1^{(\overline{k+1}\times 1)} + \boldsymbol{\epsilon}^{(n\times 1)}$.

   Let $\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$ and, $\hat{\boldsymbol{\theta}}_1 = \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{y}$, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$.

   Further, let $\mathbf{x}^*$ and $\mathbf{x}_1^*$ denote two vectors of input having $m+1$ and $k+1$ components respectively. Let $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$.

   Prove the following:

   (a) $\hat{\boldsymbol{\theta}}_1$ is biased for $\boldsymbol{\beta}_1$ and $\mathbf{x^*}_1'\hat{\boldsymbol{\theta}}_1$ is biased for $\mathbf{x^*}_1'\boldsymbol{\beta}_1$ unless $\boldsymbol{\beta}_2 = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$. Also prove that $\mathbf{x^*}_1'\hat{\boldsymbol{\theta}}_1$ is biased for $\mathbf{x}^{*\prime}\boldsymbol{\beta}$ unless $\boldsymbol{\beta}_2 = \mathbf{0}$ or $\mathbf{x}_2^* = \mathbf{A}'\mathbf{x}_1^*$.

   (b) Prove that $Var(\hat{\boldsymbol{\beta}}_1) - Var(\hat{\boldsymbol{\theta}}_1)$ is positive semi-definite, and $Var(\mathbf{x}^{*\prime}\hat{\boldsymbol{\beta}}) \geq Var(\mathbf{x^*}_1'\hat{\boldsymbol{\theta}}_1)$.

   (c) Let RSS_1 and RSS be the residual sums of squares for the reduced and the full model respectively. Also, let $s_1^2 = RSS_1/(n-k-1)$. Then show that $RSS_1 \geq RSS$ and that s_1^2 is a biased estimator of σ^2 in general.

   (d) Let $H_1 = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$. Then, under the condition that

   $$\frac{1}{\sigma^2}\boldsymbol{\beta}_2'\mathbf{X}_2'(\mathbf{I} - \mathbf{H}_1)\mathbf{X}_2\boldsymbol{\beta}_2 \leq 1,$$

   show that $Var\left(\hat{\boldsymbol{\beta}}_1\right) - MSE\left(\hat{\boldsymbol{\theta}}_1\right)$ is positive semi-definite and $Var\left(\mathbf{x}^{*\prime}\hat{\boldsymbol{\beta}}\right) \geq MSE\left(\mathbf{x^*}_1'\hat{\boldsymbol{\theta}}_1\right)$. Here MSE stands for mean square error.

Marks: 6+6+6+7=25

2. (a) Suppose that data $\{y_1, \ldots, y_n\}$ is available where $E(y_i) = \mu(t_i)$; $t_i = (2i-1)/2n$. Assuming that $\mu(\cdot)$ is unknown, consider its regressogram estimator, given by $\mu_\lambda(t) = \sum_{i=1}^{n} K(t, t_i; \lambda) y_i$, where $K(t, t_i; \lambda) = \frac{\sum_{r=1}^{\lambda} I_{P_r}(t) I_{P_r}(t_i)}{\sum_{j=1}^{n} \sum_{r=1}^{\lambda} I_{P_r}(t) I_{P_r}(t_j)}$. Here $I_{P_r}(\cdot)$ denotes the indicator function for the interval $P_r = [\frac{r-1}{\lambda}, \frac{r}{\lambda})$, $r = 1, \ldots, \lambda - 1$, and $P_\lambda = [\frac{\lambda-1}{\lambda}, 1]$. Show that, under suitable assumptions, the risk of the regressogram estimator converges to zero at the rate $n^{-2/3}$, where n is the sample size.

(b) Consider the following extension of the regressogram estimator. Define the partitions $P_j = [\frac{j-1}{\lambda}, \frac{j}{\lambda})$, $j = 1, \ldots, \lambda - 1$, and $P_\lambda = [\frac{\lambda-1}{\lambda}, 1]$. For $t \in P_j$ let $\mu_\lambda(t) = b_{0j} + b_{1j}(t - \bar{t}_j)$ for $b_{1j} = \sum_{t_i \in P_j} y_i(t_i - \bar{t}_j) / \sum_{t_i \in P_j} (t_i - \bar{t}_j)^2$, $b_{0j} = \bar{y}_j = n_j^{-1} \sum_{t_i \in P_j} y_i$, $\bar{t}_j = n_j^{-1} \sum_{t_i \in P_j} t_i$ and $n_j = \sum_{i=1}^{n} I_{P_j}(t_i)$. Show that if μ has two continuous derivatives and the number of partitions is allowed to grow at the rate $n^{1/5}$, then the squared error risk for μ_λ decays to zero at the rate $n^{-4/5}$.

Marks: 10+15=25

3. (a) Consider the proportional hazards model $h(t; x) = \lambda(t) \exp\{\beta' \mathrm{x}\}$. Let $\Lambda(t) = \int_{-\infty}^{t} \lambda(u) du = \exp\{\alpha t\}$. Suppose there are n uncensored and m censored observations. Discuss a methodology for computing the maximum likelihood estimates of the parameters α and β *(complete mathematical details not required)*.

(b) Consider the case where, for $i = 1, \ldots, n$, $y_i \in \{1, \ldots, C\}$, where $C > 1$ is a known integer. Suppose that, for $j = 1, \ldots, C$,

$$P(y_i \leq j \mid \mathrm{x}_i) = \Phi(\alpha_j - \mathrm{x}_i' \beta),$$

where $\Phi(\cdot)$ is the cdf of a $Normal(0,1)$ random variable. Assume that $-\infty = \alpha_0 < \alpha_1 < \ldots < \alpha_{C-1} < \alpha_C = \infty$, and that $\alpha_k; k = 1, \ldots, C-1$, are known. Develop an EM algorithm to compute the maximum likelihood estimate of β.

Marks: 10+15=25

4. (a) In the least squares set-up, show that the ordinary residuals e_i and the leverage values p_{ii} satisfy the inequality

$$p_{ii} + \frac{e_i^2}{SSE} \leq 1,$$

where SSE is the sum of squares due to error.

(b) Show that the variance of the regression coefficient $\hat{\beta}_j$ $(j = 1, \ldots, p)$ in the usual linear regression set up can be written in terms of the variance inflation factor VIF_j.

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination 2012-13

Course Name: M. Stat. (Ist Yr.)

Subject Name: Multivariate Analysis

Date: 03·05·13 | Maximum Marks: 100 | Time: $3\frac{1}{2}$ hours

(Total Point 105. Answer as many as you can. Maximum you can score is 100)

1. Suppose that a random sample of size 10 is available from a trivariate normal distribution, with unknown mean vector $\mu = (\mu_1, \mu_2, \mu_3)^T$, and unknown covariance matrix $\Sigma_{3\times3}$. The sample statistics are

$$\bar{X} = (1.23, 1.96, 0.84)^T$$

and

$$S = \begin{pmatrix} 0.2994 & 0.1774 & 0.1278 \\ 0.1774 & 0.1075 & 0.0747 \\ 0.1278 & 0.0747 & 0.0693 \end{pmatrix}.$$

Test for the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$, against the alternative that all the μ_i's are not equal at level $\alpha = 0.05$. [15]

2. Show that for the one way MANOVA model the likelihood ratio criterion leads to a test which is equivalent to the test given by the Wilk's Lambda criterion. [16]

3. Let $\boldsymbol{X} = (X_1, X_2)$ represent the head length and head breadth measurements of the first son in the family. Similarly let $\boldsymbol{Y} = (Y_1, Y_2)$ represent the head length and head breadth of the second son in the family. In a sample of 25 families, the matrix of correlations of $(\boldsymbol{X}, \boldsymbol{Y})$ is given by

$$\boldsymbol{R} = \left(\begin{array}{cc|cc} 1.0000 & 0.7346 & 0.7108 & 0.7040 \\ 0.7346 & 1.0000 & 0.7470 & 0.7086 \\ \hline 0.7108 & 0.7470 & 1.0000 & 0.8392 \\ 0.7040 & 0.7086 & 0.8392 & 1.0000 \end{array}\right).$$

Find the two canonical correlations between $\boldsymbol{X}$ and $\boldsymbol{Y}$, and the canonical variable pairs. Can you argue from the form of the correlation matrix that the second canonical correlation should be close to zero? [13+1=14]

4. (a) A diagnostic test for gout is based on serum uric acid level. Using appropriate units the level is approximately distributed as $N_1(5, 10)$ among healthy adults and $N_1(8.5, 1)$ among adults with gout. If 1% of the population has gout, how should a patient with level x be classified so that the total probability of misclassification is minimized? Also calculate the posterior probabilities when $x = 7$.

(b) In the two group discrimination problem suppose that the probability mass functions of X in the two populations are given by

$$f_i(x) = \binom{n}{x} \theta_i^x (1-\theta_i)^{n-x}, \quad 0 < \theta_i < 1, \; i = 1,2$$

where θ_1 and θ_2 are known. If p_1 and p_2 are the usual prior probabilities, show that the optimal classification rule leads to a linear discriminant function. Give expressions for both types of probabilities of misclassification (Assume $\theta_1 > \theta_2$).

[7+8=15]

5. The following matrix represents the matrix of Euclidean distances between 9 cities (labeled simply as 1 to 9) in Wisconsin, USA on the basis of some appropriate p-dimensional variable. Construct the complete linkage dendrogram for this distance matrix and interpret it.

$$\begin{pmatrix} 0 & & & & & & & & \\ 120 & 0 & & & & & & & \\ 98 & 33 & 0 & & & & & & \\ 102 & 50 & 36 & 0 & & & & & \\ 103 & 85 & 64 & 38 & 0 & & & & \\ 100 & 73 & 54 & 77 & 184 & 0 & & & \\ 149 & 33 & 58 & 47 & 170 & 107 & 0 & & \\ 315 & 377 & 359 & 330 & 219 & 394 & 362 & 0 & \\ 91 & 186 & 166 & 139 & 45 & 181 & 186 & 223 & 0 \end{pmatrix}.$$

[12+2=14]

6. Consider the Cressie-Read multivariate goodness-of-fit statistics

$$2nI^\lambda = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{k} n_i \left[\left(\frac{n_i}{n\pi_i^0} \right)^\lambda - 1 \right], \quad \lambda \in \mathbb{R},$$

where $\pi_0 = \{\pi_1^0, \ldots, \pi_k^0\}$ are the probabilities under the null and $N = \{n_1, \ldots, n_k\}$ are the frequencies based on a random sample of size $n = \sum_{i=1}^k n_i$ on a discrete distribution with support $\{1, 2, \ldots, k\}$. We want to test $H_0 : \pi_1 = \pi_1^0, \ldots, \pi_k = \pi_k^0$, where $(\pi_1, \ldots, \pi_k)$ is the vector of theoretical probabilities of the k-classes.

First consider the case $\lambda = 1$ (notice that $2nI^1$ is the Pearson's chi-square). Let D_{π_0} be the diagonal matrix with i-th diagonal element given by π_i^0. The following result is well known: as $n \to \infty$, $W_n = \sqrt{n}(N/n - \pi_0)$ converges in distribution to W which has a k-variate multinormal distribution under the null hypothesis with mean vector zero, and covariance matrix $D_{\pi_0} - \pi_0 \pi_0^T$.

(a) Show that $2nI^1$ can be written as a quadratic form in W_n. What is the matrix of the quadratic form?

(b) Show that the corresponding quadratic form in W (and hence the statistic $2nI^1$) has a chi-square distribution under the null hypothesis.

(c) Finally show that for any real λ, $2nI^\lambda = 2nI^1 + o_p(1)$, and argue that the Cressie-Read goodness-of-fit statistics have asymptotic $\chi^2(k-1)$ distributions for all real λ.

[4+4+8=16]

7. A principal factor analysis of a correlation matrix (corresponding to 276 observations) obtained with three factors gave the factor loadings as in the following matrix. The variables are Skull length, Skull breadth, Humerus length, Ulna length, Femur length and Tibia length respectively, of white leghorns.

$$\hat{\Gamma} = \begin{bmatrix} 0.685 & 0.361 & 0.007 \\ 0.639 & 0.409 & -0.030 \\ 0.951 & -0.083 & 0.162 \\ 0.946 & -0.152 & 0.168 \\ 0.930 & -0.180 & -0.166 \\ 0.942 & -0.124 & -0.154 \end{bmatrix}.$$

It is proposed that an orthogonal transformation matrix

$$T = \begin{bmatrix} mq & mp & l \\ -lq & -lp & m \\ p & -q & 0 \end{bmatrix} \qquad \begin{matrix} l^2 & + & m^2 & = & 1 \\ p^2 & + & q^2 & = & 1 \end{matrix}$$

be used to rotate $\hat{\Gamma}$ into a matrix approximately of the form

$$G = \begin{bmatrix} h_1 & 0 & 0 \\ h_2 & 0 & 0 \\ 0 & w_1 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & l_1 \\ 0 & 0 & l_2 \end{bmatrix},$$

in which the h's, w's and l's are regarded as 'head', 'wing' and 'leg' loadings respectively.

Determine the unknowns in T from the relation $\hat{\Gamma}T = G$, compute the 18 elements of the matrix G, and decide whether the 3-factor interpretation is defensible. [15]

[12+2=14]

4. Let $\boldsymbol{R}$ be the correlation matrix of the $p \times 1$ random vector $\boldsymbol{X}$, and consider the evaluation of the principal components of $\boldsymbol{X}$ based on this correlation matrix. Divide all the off-diagonal elements of the correlation matrix $\boldsymbol{R}$ by a constant k such that $k > 1$. Call the resulting matrix $\boldsymbol{R}^*$.

   (a) Show that the $\boldsymbol{R}$ and $\boldsymbol{R}^*$ matrices have different eigen values, but the same eigen vectors.

   (b) How are the eigen values of $\boldsymbol{R}$ related to those of $\boldsymbol{R}^*$?.

   (c) What could be a possible problem in choosing $k < 1$ in the above formulation?

[16]

5. Given an $n \times n$ dissimilarity matrix D, give a necessary and sufficient condition for the matrix to be Euclidean. Prove the necessity and sufficiency of the condition that you propose. [15]

6. Consider the general discriminant analysis problem with g populations, $\pi_1, \ldots \pi_g$. Let p_i be the prior probability for π_i, and let $c(j|i)$ be the cost of allocating any item from π_i to π_j. Prove that the classification regions that minimize the Expected Cost of Misclassification (ECM) are defined by: allocate an individual with observation x to the population π_k, $k = 1, 2, \ldots, g$ if the sum $\sum p_i f_i(x) c(j|i)$ is minimized at $j = k$, where the sum is over $i = 1, \ldots, g$, $i \neq j$, and $f_i(x)$ is the density of X under π_i. [14]

7. The following matrix represents dissimilarities between pairs of companies; the dissimilarity measure is the number of variables in which two companies have different scores out of a set of 32 binary variables.

Company	A	B	C	D	E	F	G	H
A	-	3	16	16	18	26	21	25
B		-	15	15	19	26	24	22
C			-	6	22	24	23	19
D				-	22	20	21	21
E					-	21	17	13
F						-	13	15
G							-	8

Applying the Single Linkage and Complete Linkage methods, obtain dendrograms for clustering these companies, and comment on the form of the dendrograms. What similarities and differences do you see between the two dengrograms? [12+2=14]

INDIAN STATISTICAL INSTITUTE

Back Paper Examination 2012-13 (Semester II)

Course Name: M. Stat (Ist Yr.)

Subject Name: Multivariate Analysis

Date: 03.07.13 Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

1. Let $S \sim W_p(k, \Sigma)$ (p dimensional central Wishart with degrees of freedom k, and parameter Σ, assumed positive definite). Let $S^{-1} = (s^{ij})_{p\times p}$ represent the inverse of $S = (s_{ij})_{p\times p}$; similarly $\Sigma^{-1} = (\sigma^{ij})_{p\times p}$ be the inverse of $\Sigma = (\sigma_{ij})_{p\times p}$. Show that

   (a) $\frac{\sigma^{pp}}{s^{pp}} \sim \chi^2(k-p+1)$, and is independent of (s_{ij}), $i, j = 1, \ldots, p-1$.

   (b) $\frac{\mathbf{l}^T\Sigma^{-1}\mathbf{l}}{\mathbf{l}^T S^{-1}\mathbf{l}} \sim \chi^2(k-p+1)$ for any fixed vector $\mathbf{l}$. [6+7=13]

2. Suppose $z_0, z_1, \ldots, z_p$ are independently and identically distributed with mean 0 and variance σ^2. Let $X_i = z_0 + z_i$, $i = 1, 2, \ldots, p$. Verify that there is a principal component of $X = (X_1, \ldots, X_p)$ which is proportional to $\bar{X}$. Argue that this is the principal component with the maximum variance, i.e. it is the first principal component. [14]

3. A study was performed on the market value measures of profitability. With 8 variables, the following rotated principal components estimates of factor loadings for a 3 factor model is obtained:

	Estimated factor loadings		
Variable	F_1	F_1	F_3
Historical return on assets	0.433	0.612	0.499
Historical return on equity	0.125	0.892	0.234
Historical return on sales	0.296	0.238	0.887
Replacement return on assets	0.406	0.708	0.483
Replacement return on equity	0.198	0.895	0.283
Replacement return on sales	0.331	0.414	0.789
Market Q ratio	0.928	0.160	0.294
Market relative excess value	0.910	0.079	0.355

   (a) Determine the specific variances, communalities, and the proportion of total variance explained by each factor. Does a 3 factor model appear to be appropriate to you?

   (b) Assuming that estimated loadings below 0.4 are small, can you interpret the three factors?