Indian Statistical Institute
Midterm Examination
First Semester, 2016-2017 Academic Year
M.Stat. 2nd Year
Statistical Inference II

Date: 05 September, 2016 Total Marks : 35 Duration: 2 hours

Answer all questions

1. Let X_1, X_2 be i.i.d. $\sim U(\theta - 1/2, \theta + 1/2)$, $\theta \in R$. A frequentist 95% confidence interval is $(\bar{X} - 0.056, \bar{X} + 0.056)$ where $\bar{X} = (X_1 + X_2)/2$. Show that if X_1 and X_2 are sufficiently apart, say $X_2 - X_1 > c$ (find c), then θ must lie in this confidence interval. Justify your answer. [3]

2. Suppose X has density $e^{-(x-\theta)} I(x > \theta)$ and the prior density of θ is $\pi(\theta) = [\pi(1 + \theta^2)]^{-1}$. Consider a loss function $L(\theta, a) = I(|\theta - a| > \delta)$ and find the Bayes estimate of θ. [5]

3. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be two independent random samples from two normal populations with distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ respectively. Assume that the prior distribution of $(\mu_1, \mu_2, \log \sigma^2)$ is improper uniform where μ_1, μ_2, and σ^2 are independent. Find the posterior distribution of $\mu_1 - \mu_2$. [8]

4. Let the sample space be $\{1, 2, \ldots, k\}$. Let $P = (p_1, \ldots, p_k)$ be a random probability distribution on ($p_i \geq 0 \ \forall i$, $\sum p_i = 1$). Let $X_1, \ldots, X_n$ be i.i.d. $\sim P$, and $P \sim$ Dirichlet$(\alpha_1, \ldots, \alpha_k)$, $\alpha_i > 0 \ \forall i$. Show that for any subset A of the sample space, the posterior mean of $P(A)$ is a weighted average of its prior mean and $P_n(A)$ where P_n denotes the empirical distribution of $X_1, \ldots, X_n$. [8]

5. Define the Jeffreys prior. Calculate this prior when the sampling model is multinomial. Briefly argue why the Jeffreys prior may be considered a noninformative/low information prior. [1+3+7=11]

MSTAT II - Branching Processes

Midsem. Exam. / Semester I 2016-17

Time - 2 hours/ Maximum Score - 30

06.09.16

NOTE : SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.

1. (7 marks)

   Let $\{X_n : n \geq 0\}$. be a Galton-Watson Branching Process with offspring distribution $p_j = 1/(j+1)^3$, for $j \geq 1$. and $p_0 = 1 - \sum_{j\geq 1} p_j$. Does there exist a non-negative random variable W s.t. $X_n/m^n \to W$. Justify your answer, clearly giving the reasons.

2. (7 marks)

   Fix a positive integer, say, k. Calculate the limit, $\lim_{n\to\infty} P(X_n = j \mid X_n > 0, X_{n+k} = 0, X_0 = 1)$.

3. (6+6=12 marks)

   Let $\gamma = f'(q)$ and $m > 1$.

   (a) Show that, $\sum_j a_j < \infty$, where $a_j = |\gamma - f'(f_j(s))|$.

   (b) Hence or otherwise, show that $\lim_{n\to\infty} f_n'(s)/\gamma^n$ exists and finite for every $0 \leq s < 1$.

4. (7 marks)

   A family name existed for 100 years. Determine the probability (justifying your steps) that the family will stay another 200 years, given that the expected no. of offspring who bears the name is 1 and has finite variance, say σ^2.

All the best.

INDIAN STATISTICAL INSTITUTE
Mid-Semestral Examination: 2016-2017
MStat II Year
Time Series Analysis

Date: 7th September 2017 Maximum Marks 30 Duration 2 hours

All notations are self-explanatory. You can answer any part of any question.

1. Construct an example of a strong stationary but not weak stationary process. Construct an example of a weak stationary but not strong stationary process. [3+2=5]

2. Let $\{X_t, t = 1,2, \ldots T\}$ be a sequence of random variables with mean zero and with unit variance. Define $Z_t = X_t X_{t-1}$. Examine whether Z_t is weakly stationary or not for the following cases:

   (a) $\{X_t, t = 1,2, \ldots T\}$ is white noise,
   (b) $\{X_t, t = 1,2, \ldots T\}$ is i.i.d process. [3+2 =5]

3. Let $\{X_t, t = 1,2, \ldots T\}$ be a quarterly observed time series with three classical components, viz., trend, Seasonality; and irregular component. Describe a regression method to estimate all these three components. Assume that three components are in additive form. [9]

4. Define Ergodicity for a stationary time series. Let $\{X_t, t = 1,2, \ldots T\}$ be a covariance stationary and ergodic time series with mean μ. Provide a consistent estimator of μ. Prove that your proposed estimator is consistent. [2+1+5=8]

5. Let $\{X_t, t = 1,2, \ldots T\}$ be a time series which follows a linear time trend model: $X_t = \alpha + \beta\, t + \varepsilon_t$. ε_t 's are i.i.d. Normal random variables with mean 0 and with variance σ^2. Prove that the least square estimators of α and β are consistent. Derive the asymptotic distributions of the least square estimators of α and β. Provide the rate of convergence. [3+3+2=8]

INDIAN STATISTICAL INSTITUTE
M. Stat. (II year) 2016 – 17, First Semester
Mid-Semester Examination
Pattern Recognition

Date: September 8, 2016 **Maximum marks: 60** **Duration: 120 minutes**

Note: Answer all the questions

1. State the Bayes decision rule for three-class classification problem and show that it minimizes the probability of misclassification. [2+10=12]

2. Let

$$p_i(x) = ie^{-ix} \; ; x > 0 \qquad \text{for } i=1,2.$$
$$\quad = 0 \;\; otherwise.$$

Let there be two classes with prior probability for class1 being P. Let the density function for the ith class be p_i for i=1,2. Find the Bayes decision rule for the classification problem, and find its probability of misclassification. [15]

3. (a) State the k-nearest neighbor density estimation procedure.
(b) Derive the k-nearest neighbor decision rule from the above procedure.
(c) State an algorithm for reducing the size of the training set for k-nearest neighbor decision rule. [4+6+6=16]

4. (i) State the minimum within cluster distance criterion.
(ii) Describe the k-means algorithm for clustering.
(iii) Give an example of a data set and two sets of initial seed points for which the resultant clusterings would be different when k-means algorithm is applied.
[4+4+4=12]

5. Describe the Complete linkage clustering procedure. [5]

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, First Semester, 2016-17

Mid-Semestral Examination 10·09·16

Full Marks: 60 | Statistical Computing-1 | Time: $3\frac{1}{2}$ Hours

(Answer as many as you can. The maximum you can score is 60.)

1. Two persons were asked to estimate the value of π by generating at most n observations from an appropriate distribution.

   (a) One person generated n observations from the uniform distribution on $S = \{(x,y): -1 \le x, y \le 1\}$ and counted how many of them had Euclidean norm smaller than unity. Suggest an unbiased estimator of π based on this experiment. [2]

   (b) Another person generated observations $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$ from the uniform distribution on $C = \{(x,y): x^2 + y^2 \le 1\}$ and checked whether they lie inside the region $R = \{(x,y): -1/\sqrt{2} \le x, y \le 1/\sqrt{2}\}$. Let T denote the number of observations lying inside R and $R_T = \max\{i: \mathbf{z}_i \in R\}$. Construct an unbiased estimator of π based on T and R_T. [4]

   (c) Check which of these two estimators (constructed in parts (a) and (b)) has smaller variance. [6]

2. Write down an algorithm for generating observations from the following distribution with the probability mass function given by

   (a) $f(x) \propto (1+x^2)^{-1}; \quad x = 0, \pm 1, \pm 2, \ldots$ [6]

   (b) $f(x,y) \propto \frac{p^x(1-p)^{-x}\{\lambda(1-p)\}^y}{x!(y-x)!} \quad y = 0,1,2,\ldots; \quad x = 0,1,\ldots,y$; where $\lambda > 0$ and $p \in (0,1)$ are known real constants. [6]

3. Prove or disprove the following statements.

   (a) Half-space depth of the half-space median of a distribution cannot be smaller than 0.5. [4]

   (b) Simplicial depth is invariant under affine transformation. [3]

   (c) If two points $\mathbf{x}_0$ and $\mathbf{x}_1$ have the same projection depth δ with respect to a distribution, for any $\lambda \in (0,1)$, the projection depth of $\mathbf{x}_\lambda = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_0$ cannot be smaller than δ. [3]

   (d) Spatial median of a bivariate continuous distribution cannot lie outside the support of the distribution. [3]

   (e) In the case of univariate continuous distribution, simplicial depth is a monotone function of half-space depth. [3]

   (f) In the case of univariate data, MVE and MCD estimates (based on 50% observations) of location coincide with LMS and LTS estimates of location, respectively. [4]

P.T.O

4. Consider a data set $\Omega_n = \left\{ \binom{x_1}{y_1}, \binom{x_2}{y_2}, \ldots, \binom{x_n}{y_n} \right\}$ consisting of n bivariate observations.

(a) Check whether $S_1(\beta) = \sum_{i=1}^{n} |y_i - \beta x_i|$ and $S_2(\beta) = \sum_{i=1}^{n} (y_i - \beta x_i)^2$ are convex functions of β. [3]

(b) Comment on the uniqueness of the minimizers of $S_1(\beta)$ and $S_2(\beta)$. [3]

(c) Let β_1 be a minimizer of $S_1(\beta)$ and β_2 be a minimizer of $S_2(\beta)$. For any $\lambda \in (0,1)$, let β_λ be a minimizer of $S_\lambda(\beta) = \lambda S_1(\beta) + (1-\lambda)S_2(\beta)$. Show that β_λ cannot lie outside the set $S = \{\alpha\beta_1 + (1-\alpha)\beta_2 : \ 0 \le \alpha \le 1\}$. [4]

5. Consider the following data set.

x	1	2	3	4	5	6	7	8	9	10
y	18.6	15.7	10.8	6.9	3.5	-0.9	-5.6	-8.7	-13.8	-16.5

(a) Find the equation of the line $y = \alpha + \beta x$ that passes through the point $(5,3)$ and minimizes $\sum_{i=1}^{10} |y_i - \alpha - \beta x_i|$, where $(x_i, y_i)'$ denotes the i-th $(i = 1, 2, \ldots, 10)$ observation. [6]

(b) Can this line (obtained in part (a)) be an LMS line? Justify your answer. [2]

(c) Find out all possible values of α that minimize $\psi(\alpha) = 4\sum_{i=1}^{10} |y_i - \alpha| + \sum_{i=1}^{10} (y_i - \alpha)$. [4]

Indian Statistical Institute

Mid-Semester Exam: 2016-17

M.Stat. II Year, Statistical Genomics

Max. Marks: 50 Time: 2 Hours 12.09.16

1. Consider the interval mapping for a backcross design. There are two markers M_1 and M_2. The recombination fractions between the two markers, M_1 and QTL , and M_2 and QTL are r, r_1, r_2 respectively. If the two markers are highly linked, show that $r \simeq r_1 + r_2$. [10]

2. Consider the QTL mapping problem for a Backcross design. The goal is to locate the QTL(s) controlling the observed trait. Suppose there are n subjects in the study and $Y_1, Y_2, \ldots, Y_n$ are response measures. Note that we dont know the QTL genotypes of the subjects under study.

   Define the variables $Z_1, Z_2, \ldots, Z_n$ as follows: Z_i=0, if the i-th subject belongs to the first genotype group ($N(\mu_1, \sigma^2)$) and 1 for the second genotype group ($N(\mu_2, \sigma^2)$). Let π be the proportion (unknown) of individual in the population belonging to the first genotype group.
   (a) Write down the incomplete data likelihood function.
   (b) Write down the complete data likelihood function.
   (c) Give the E-step of the EM algorithm and clearly give the expression for $E(Z_i|$data, initial estimates of the parameters).
   (d) Give the closed form expressions for π, μ_1, μ_2 and σ^2 from M-step after the first iteration. [3+3+3+6]

3. (a) Consider the QTL mapping based on Linkage Disequilibrium. Suppose we observe the marker data for n subjects where M and m are the marker alleles and Q, q are the QTL alleles. Let D be the coefficient of linkage disequilibrium between the marker and the QTL and $Y_1, Y_2, \ldots, Y_n$ be the observed responses from n subjects. Construct a bootstrap based test procedure for testing the existence of a potential QTL controlling the observed trait.
   (b) Show that the probability that the bootstrap resample does not contain the response from a particular subject will tend to e^{-1} as $n \to \infty$. [10+5]

4. (a) Consider a family where Della is the daughter of Alex and Brittany. Prince is a half-sib of Brittany, they share a common mother. Prince marries Jenn and they have a son Andrew. Derive the relationship between Della and Andrew in terms of the standard transition probability matrices.
   (b) Suppose now that Jenn is indeed double first cousin of Prince. Derive the relationship between Della and Jenn in terms of the transition probability matrices. [5+5]

Indian Statistical Institute

Mid-Semester Examination

Course Name: MSQE 1st Year & MStat 2nd Year 2016

Subject: Basic Economics

Full Marks -40 Time: 2 hrs

Answer all questions.

Date : 14.09.2016

1. The following data regarding a firm in a given year are specified below:
(All figures are in lakhs of rupees)

Revenue earned from the sale of (5/6)th of output (The remaining part was unsold)	600
Raw materials purchased from other firms	100
Unused part of raw materials	20
Interest paid to households	5
Payment made to a labour contractor for supplying labour	5
Travelling and hotel expenses of the officials of the company	1
Land purchased by the company for construction of an additional shed	100
Wages and salaries	2
Depreciation	1
Dividend paid	100
Tax on profit	1
Subsidy received by the firm	2
Donations made to the chief minister's relief fund	1

From the data given above compute the firm's contribution to

(i) NDP

(ii) National Income

(iii) Personal Income, Private income

(iv) Aggregate final expenditure [5x4 = 20]

2. (i) Is it possible for an economy to absorb more than what its purchasing power can command of the world NDP? Explain your answer. Also discuss the financial aspects.

(ii) Suppose in an economy in a given year the central bank had to sell foreign exchange worth Rs. 20,000 crore from its stock to hold the exchange rate at the target level. Some firms in the domestic economy borrowed from foreign financial institutions Rs. 30,000 crore, while some foreign firms borrowed Rs.12,000 crore from domestic financial institutions. In additions foreigners purchased shares of domestic companies worth Rs.8000 crore. Domestic residents purchased land abroad worth Rs.200 crore. Domestic government's budget deficit was Rs.500 crore. Domestic households' expenditure on produced goods and services was Rs.90,000 crore of which Rs.10,000 was spent on buying houses from construction companies. From the data given above, compute the economy's private disposable income in the given period. [8+12=20]

Note on Q2 (Firm investment in the domestic economy = Rs 60,000 cr
Depreciation in the same ec: = Rs 500 cr)

INDIAN STATISTICAL INSTITUTE

Mid-Semester of First Semester Examination : 2016 - 17

Course Name : M. Stat. II Year

Subject Name : Signal and Image Processing

Date: 15. 09. 2016 | Maximum Marks : 60 | Duration : 2 hours

Part - I

Answer any THREE questions: $3 \times 10 = 30$

1. (a) Consider the analog signal: $x_a(t) = 3\cos 100\pi t$.
   (i) Determine the minimum sampling rate required to avoid aliasing.
   (ii) Suppose that the signal is sampled at the rate $F_s = 200$ Hz. What is the discrete-time signal obtained after sampling?

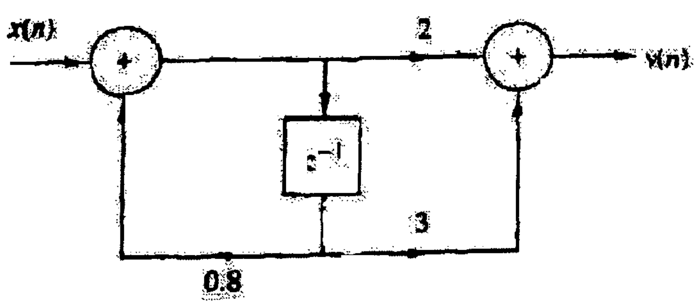
   (b) Show that the fundamental period N_p of the signals $s_k(n) = e^{j2\pi kn/N}$ for $k = 0, 1, 2,...$ is given by $Np = N/\ GCD(k, N)$, where GCD is the greatest common divisor of k and N. What is the fundamental period of this set for $N = 7$ and 16?

   (2+3)+(3+2) = 10

2. Compute and sketch the output response in discrete time domain for a system with unit impulse response $h(n) = [u(n+2) - u(n-3)].(3 - |n|)$ for the exciting input $x(n) = u(n+1) - u(n-4) - \delta(n-5)$. (10)

3. a) The discrete-time system $v(n) = ny(n-1) + x(n)$, $n > 0$ is initially relaxed [i.e., $x(-1) = 0$]. Check if the system is BIBO stable.

   b) Determine the impulse response for the system shown in the following figure:



(4+6 = 10)

P.T.O

4. a) Determine the causal signal with the z transform:

$$X(z) = \frac{1}{(1+z^{-1})(1-0.5z^{-1})^2}$$

b) The step response of an LTI system is $s(n) = \left(\frac{1}{3}\right)^{n-2} u(n+2)$. Find the system function H(z) and sketch the pole-zero plot. (5+5 = 10)

5. a) Derive the normalized autocorrelation sequence of the signal $x(n)$ given by:

$$x(n) = \begin{cases} 1, & \text{for } -N \le n \le N \\ 0, & \text{otherwise.} \end{cases}$$

b) Prove that for autocorrelation of a signal $x(n)$, $|r_{xx}(l)| \le r_{xx}(0)$ for any l. (5+5 = 10)

Part - II

Answer any FIVE questions: (5×6=30)

1. Given an image $I(r,c)$, how magnitude and phase of the frequency component of the image can be evaluated? What intuitive information are given by magnitude and phase data? (4+2)

2. Find matrix element values for Gaussian smoothing mask having 5 rows and 5 columns. Assume $\sigma^2 = 2$.

3. How area, centroid and perimeter of a binary shape in a discrete grid can be calculated from the projections of the binary shape? (1+2+3)

4. In a digital image of red blood cells (RBC) captured using a camera fitted on the microscope, how can the number of RBC be counted?

5. For a 3-bit image with pixel values 0 to 7, the number of pixels are 5, 3, 5, 0, 12, 18, 22 and 0 at pixel values 0, 1, 2, 3, 4, 5, 6 and 7 respectively. Find the histogram equalized values of the grey levels of the image.

6. It is known a priori that a binary image contains point of intersection of two straight lines. List steps of an algorithm to find and localize the point of intersection in the binary image.

7. How RLE can be used to compress grey level image? How CMY values can be obtained from RGB image values? (4+2)

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016-17

Course Name	:	**M. Stat.**
Subject Name	:	**Martingale theory**
Date	:	15/09/16
Maximum Marks	:	**35**
Duration	:	**2 hours**

Answer **any three** questions. Each questions carries 12 marks. The maximum you can score is 35.

1. Prove or disprove the following statement. If $((X_n, \mathcal{F}_n) : n \geq 1)$ is a martingale, such that $X_n \to X_\infty$ a.s. as $n \to \infty$ for some integrable random variable X_∞, then $\sup_{n\geq 1} \mathrm{E}|X_n| < \infty$.

2. Suppose that $X_1, X_2, \ldots$ are independent random variables with X_n following Normal$(0, \sigma_n^2)$ for all n.

   (a) If $\sum_{n=1}^\infty \sigma_n^2 < \infty$, show that
   $$S_n := \sum_{i=1}^n X_i,\ n \geq 1\,,$$
   converges a.s. and in L^p for every $p \in [1, \infty)$, as $n \to \infty$.

   (b) If $\sum_{n=1}^\infty \sigma_n^2 = \infty$, show that S_n doesn't converge weakly to any finite random variable.

3. Let $(X_n, \mathcal{F}_n)$ be a non-negative martingale with the following property. Given $\varepsilon > 0$ there exists $\delta > 0$ such that the following is true for every n. For all $A \in \mathcal{F}_n$ for which $P(A) \leq \delta$, it holds that $\int_A X_n dP \leq \varepsilon$. Show that there exists an integrable random variable X such that as $n \to \infty$,
   $$X_n \to X \text{ a.s. and in } L^1\,.$$

4. Consider a probability space $(\Omega, \mathcal{A}, P)$ and a measurable function $X : \Omega \to [0, \infty)$. Let $\mathcal{F}$ be a sub σ-field of $\mathcal{A}$.

   (a) Show that there exists a $\mathcal{F}$-measurable function $Y : \Omega \to [0, \infty]$ such that
   $$\int_A X dP = \int_A Y dP \text{ for all } A \in \mathcal{F}\,.$$

   (b) If $\mathrm{E}(X) = \infty$ and $\mathcal{F} = \{\emptyset, \Omega\}$, find a Y which satisfies the above.

   (c) Prove that in (a), Y can be chosen to be finite a.s. if and only if there exists $A_1, A_2, \ldots \in \mathcal{F}$ such that $\Omega = \bigcup_{n\geq 1} A_n$, and
   $$\int_{A_n} X dP < \infty \text{ for each } n \geq 1\,.$$

INDIAN STATISTICAL INSTITUTE
MID-SEMESTRAL EXAMINATION 2016

M.STAT 2nd year. Advanced Design of Experiments

September 15, 2016, Total marks 30 Duration: $1\frac{1}{2}$ hours
Answer all questions.
Keep your answers brief and to the point.

1. a) Define the criterion of D-optimality and explain the statistical significance of this criterion.

   b) State and prove the D-optimality of balanced incomplete block (BIB) designs for estimating full sets of orthonormal treatment contrasts under the usual additive model.

   c) Will a BIB design with parameters $v, b, k (< v)$ be universally optimal in the class of all block designs with v treatments in b blocks of size k each? Justify your answer.

   $[3 \times 4 = 12]$

2. a) Can there exist a Hadamard matrix of order 15? Justify your answer with a proof.

   b) Obtain the elements of GF(11).

   c) Obtain **only the first 4 columns** of the Hadamard matrix of order 12 which can be constructed using the elements obtained in (b) above.

   d) Define an orthogonal array. Will any $N \times k'$ subarray of an orthogonal array $OA(N, k, s, t)$, $(k > k')$, also be an orthogonal array.

   (e) Does an orthogonal array $OA(81, 10, 9, 2)$ exist? Justify your answer with a proof. (Actual construction of OA is not needed.)

   f) What is meant by a N-run plan in the set up of a 3^4 factorial experiment? $[6 \times 3 = 18]$

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2016 – 17

MStat (2nd Year)

Financial Econometrics

Date: 13 September 2016 Maximum Marks: 30 Duration: 2 Hours

This paper carries 33 marks. Attempt ALL questions. The maximum you can score is 30

1. (a) What is a dividend payment. What is its effect on the stock price?

   (b) As we know that stock price returns are thick tailed, Normal distribution is not a reasonable approximation. Does Stable distributions help?

[(3 + 2) + 4 = 9]

2. How are the Random Walk Hypotheses RW1, RW2 and RW3 related? Use a Venn diagram for your answer and provide specific examples. [7]

3. Suppose the trading process $\{\delta_{it}\}$ defined by

$$\delta_{it} = \begin{cases} 1\ (no\ trade) & with\ probability\ \pi_i \\ 0\ (trade) & with\ probability\ (1-\pi_i) \end{cases}$$

were not *iid*, but followed a two state Markov chain with transition probabilities

$$\begin{array}{ccc} & & \delta_{it} \\ & & \begin{array}{cc} 0 & 1 \end{array} \\ \delta_{it-1} & \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \pi_i & (1-\pi_i) \\ (1-\pi_i') & \pi_i' \end{pmatrix} \end{array}$$

Derive the unconditional mean, variance and first order auto covariance of δ_{it} as functions of π_i and π_i'. [2 + 4 + 4 = 10]

4. Distinguish between statistical and economic models in the context of event study analysis. Use a specific example for your answer. [7]

Indian Statistical Institute

Mid-semsetral examination : (2015-16)

M. Stat II year

Functional Analysis

Date : 13/9/16 Maximum marks : 45 Duration : 2 hours.

Answer any THREE questions. Each question carries 15 marks.

(1) Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Prove that any bounded linear functional f on $\mathcal{M}$ **has a unique** norm-preserving extension on $\mathcal{H}$, i.e. there is a unique bounded linear functional $\tilde{f}$ on $\mathcal{H}$ satisfying $\tilde{f}|_{\mathcal{M}} = f$ and $\|\tilde{f}\| = \|f\|$. Give a counterexample to show that the uniqueness of norm-preserving extension is no longer true if the Hilbert space is replaced by a general Banach space.

(2) Prove that the unit ball of an infinite dimensional Banach space is not compact.

(3) With the notation introduced in class, prove that $c_0^* \cong l^1$, where $\cong$ means isometric isomorphism of Banach spaces.

(3) Let $\{f_n\}_{n\geq 1}$ be a sequence of functions in $L^1(\mathbb{R})$. Prove that $\int f_n g dx \to 0$ for every $g \in L^\infty(\mathbb{R})$ if and only if $\sup_{n\geq 1} \|f_n\|_1 < \infty$ and $\int_E f_n dx \to 0$ for every Borel subset E of $\mathbb{R}$.

Date: 15.11.2016

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, First Semester, 2016-17

Semestral Examination

Time: $3\frac{1}{2}$ Hours | Statistical Computing | Full Marks: 100

Answer as many as you can. The maximum you can score is 100.

1. Using the random numbers given below generate an observation (without using the acceptance rejection principle) from each of the following probability distributions.

0.8468	0.2764	0.3269	0.9541	0.1703	0.6002	0.4945
0.1146	0.3531	0.7220	0.4359	0.0596	0.7813	0.1623

(a) $f(x,y,z) = \begin{cases} \frac{1}{y(1-y)}, & \text{if } 0 \le x \le y \le z \le 1 \\ 0, & \text{otherwise} \end{cases}$ [4]

(b) $f(x,y,z) = \begin{cases} \frac{3}{4}, & \text{if } |x|+|y|+|z| \le 1 \\ 0, & \text{otherwise} \end{cases}$ [6]

(c) $f(x,y,z) = Ce^{-\sqrt{x^2+y^2+z^2}}$, where C is a normalizing constant. [8]

2. Consider a data cloud of the form $\{(x_i, y_i);\ i = 0,1,2,\ldots,20\}$, where $x_i = i - 10$ and $y_i = x_i^2 - 21x_i + 20$.

(a) Find a linear fit which has the maximum regression depth with respect to this data cloud. Is this choice unique ? Justify your answer. [4+2]

(b) Let $\hat{f}_h(x)$ be the Nadaraya-Watson estimate of the regression function f based on a Gaussian kernel with bandwidth h. Show that irrespective of the choice of h, $\hat{f}_h(x)$ is monotonically decreasing. [10]

(c) Find the limiting value of $\hat{f}_h(2.5)$ when the bandwidth h shrinks to zero. [4]

3. (a) Let $t_0 < t_1 < \ldots < t_k$ be a set of points in R. Consider a function f defined on $(t_0, t_k]$, which is continuous and linear in $[t_{i-1}, t_i]$ for all $i = 1,2,\ldots,k$. Let C be the class of all such continuous functions. Define $f_0(t) = 1$, $f_1(t) = t$ and $f_i(t) = max\{0, t - t_{i-1}\}$ for $i = 2,3,\ldots,k$. Show that $f_0, f_1, \ldots, f_k$ form a basis for C. [6]

(b) Give an example to show that small reduction in impurity function should not be used as a single stopping criterion for the construction of a classification tree. [4]

(c) Give an example (with proper justification) of a regression problem, where

i. projection pursuit regression is expected to perform better than regression based on an additive model. [3]

ii. local linear regression is expected to perform better than kernel regression. [3]

4. (a) In the case of a spherically symmetric (about the origin) distribution F, show that spatial rank of an observation $\mathbf{x}$ has the same direction as $\mathbf{x}$. If $\delta(\mathbf{x}, F)$ denotes the spatial depth of $\mathbf{x}$ with respect to F, show that $\delta(\mathbf{x}, F)$ is a decreasing function of $||\mathbf{x}||$. [3+5]

   (b) Let G be an elliptically symmetric distribution with the location parameter $\boldsymbol{\mu}$ and the scatter matrix $\boldsymbol{\Sigma}$. If $\text{HD}(\mathbf{x}, G)$ denotes the half space depth of an observation $\mathbf{x}$ with respect to the distribution G, show that $\text{HD}(\mathbf{x}, G)$ is a decreasing function of the Mahanobis distance $\{(\mathbf{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\}^{1/2}$. [10]

5. (a) Assume that the life time of an electric bulb manufactured by a company follows an exponential distribution with the mean θ. In order to estimate θ, the manager of the company instructed a person to switch on n electric bulbs at the same time. Within the first T hours, only m of these n bulbs stopped working and their life times were noted.

      i. Write down the likelihood function and find the maximum likelihood estimate of θ. [4]

      ii. Considering the life times of other $n-m$ bulbs as missing data, write down an EM algorithm for finding the maximum likelihood estimate of θ. Does this algorithm converge? Does it lead to the same solution as obtained in (i)? Justify your answer. [4+2+2]

   (b) Show that the iteratively re-weighted least squares method used for LAD (least absolute deviations) regression can be viewed as a majorization-minimization algorithm. Prove the convergence of this algorithm. [5+3]

6. (a) Describe the Metropolis-Hastings algorithm for generating observations from a target distribution π. Show that if $X_t \sim \pi$, all subsequent observations generated by this algorithm will follow the same distribution. [2+4]

   (b) Let $x_1, x_2, \ldots, x_n$ be n independent observations from a normal distribution with the location parameter μ and the scale parameter σ^2. If the prior distribution of (μ, σ^2) is given by $\pi(\mu, \sigma^2) \propto 1/\sigma^2$, give an appropriate algorithm for generating observations from the posterior distribution of (μ, σ^2). [8]

INDIAN STATISTICAL INSTITUTE

End-Semester Examination: 2016-17

Course Name	:	**M. Stat. 2nd year**
Subject Name	:	**Martingale theory**
Date	:	17.11.2016
Maximum Marks	:	**100**
Duration	:	**3 hours**

Answer each of the following questions. Each questions carries 20 marks.

1. Let $(\Omega, \mathcal{A}, P)$ be a probability space, and ν be a finite measure on Ω. Suppose that $(\mathcal{F}_n)$ is a filtration such that $\nu \ll P$ on $\mathcal{F}_n$ for every $n \in \mathbb{N}$. Denote by X_n the Radon-Nykodym derivative of ν with respect to P on $\mathcal{F}_n$.

   (a) Show that there exists a random variable X_∞ such that
   $$X_n \to X_\infty \text{ a.s.},$$
   as $n \to \infty$.

   (b) If $\mathrm{E}(X_\infty) = \nu(\Omega)$, then show that $\nu \ll P$ on $\mathcal{F}_\infty$, where
   $$\mathcal{F}_\infty = \bigvee_{n=1}^{\infty} \mathcal{F}_n .$$

2. Suppose that $(X_1, X_2, \ldots)$ is an exchangeable sequence of random variables such that for each $n \geq 1$, $P(X_n \in \{0,1\}) = 1$. Show that $X_1, X_2, \ldots$ are i.i.d., if and only if, every event in the tail sigma field $\mathcal{T}$, defined by
   $$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots),$$
   has probability either 0 or 1.

3. Assume that $((X_n, \mathcal{F}_n) : n \geq 1)$ is a martingale such that
   $$\limsup_{n\to\infty} \mathrm{E}(e^{tX_n}) < \infty \text{ for all } t \in \mathbb{R}.$$
   Show that there exists a random variable X_∞ such that
   $$X_n \to X_\infty ,$$
   a.s. and in L^p for every $p \in [1, \infty)$.

P.T.O

4. Let $(Z_{i,j} : i \in \mathbb{N} \cup \{0\}, j \in \mathbb{Z})$ be a family of independent random variables taking values in $\mathbb{Z}$ such that

$$\begin{aligned} \mathrm{E}(Z_{i,j}) &= 0 \text{ for every } i, j\,, \\ \mathrm{E}(Z_{i,j}^2) &= 1 \text{ for every } i, j\,, \\ \text{and } \sup_{i,j} \mathrm{E}(|Z_{i,j}|^3) &< \infty\,. \end{aligned}$$

Define the random variables X_n inductively as follows:

$$\begin{aligned} X_0 &:= 0\,, \\ X_{n+1} &:= X_n + Z_{n,X_n},\ n \geq 0\,. \end{aligned}$$

Show that as $n \to \infty$,

$$X_n/\sqrt{n} \Rightarrow G\,,$$

where G follows standard Normal.

5. Suppose that $(X_n : n \geq 1)$ is a martingale with respect to its natural filtration such that $X_0 = 0$ and

$$|X_n - X_{n-1}| \leq \frac{1}{\sqrt{n}} \text{ for all } n \geq 1\,.$$

Show that $(X_n/\sqrt{\log n})$ is tight.

INDIAN STATISTICAL INSTITUTE
Semestral Examination
Advanced Design of Experiments
M.Stat- 2nd Year : 2016-17

Full Marks: 70 | Time : 3 hours | Date: 17.11.16

NOTE: Use separate Answer Booklets for Group A and Group B.
Answer all questions.

Group A

Q1.a) Show that for a response surface model, a first order orthogonal design is equivalent to a first order rotatable design.

b) State the necessary conditions for a second order rotatable design (SORD).

c) Suggest a useful second order rotatable design in 3 variables.

(8+2+4=14)

Q2.a) Define a strongly balanced uniform crossover design in t treatments, n units and p periods and state the necessary conditions for existence of such a design.

b) Show that the design defined in part a) is universally optimal for the estimation of direct effects of treatments. (You may assume the form of the C-matrix for direct effects for a general cross over design).

c) Construct a balanced cross over design which is uniform over periods, with 16 treatments, 6 periods and the minimum number of units.

((2+2)+ 7+ 5=16)

Q3. a) Construct a $1/8^{th}$ fraction of a 2^7 factorial experiments of maximum resolution. Justify why your design attains the maximum resolution.

b) Write down the alias sets for all main effects.

c) Can you view it as an orthogonal array? If so, write down the strength of the array, if not, justify your claim. (6+2+5+2=15)

Q4. Assignment. (5)

P.T.O

INDIAN STATISTICAL INSTITUTE
SEMESTRAL EXAMINATION 2016

M.STAT 2nd year. Advanced Design of Experiments

Group B

Answer all questions. Keep your answers brief and to the point.

1. (a) Explain the statistical significance of the E-optimality criterion.

   b) In an experimental situation, 7 treatments are to be compared and the available experimental material is such that a row-column design is appropriate with 3 rows and 7 columns. Identify a universally optimal design in this set-up. Justify your claim with a proof. [(3+7=10]

2. (a) Prove that if a Hadamard matrix of order 24 exists, then a symmetric balanced incomplete block (BIB) design with 23 treatments also exists. Determine the number of blocks and their sizes for this BIB design.

   (b) Does an orthogonal array OA(25, 4, 5,2) exist? Justify your answer.

   [5 + 5 = 10]

INDIAN STATISTICAL INSTITUTE

End-Semester of First Semester Examination : 2016 - 17

Course Name : M. Stat. II Year

Subject Name : Signal and Image Processing

Date: 18.11. 2016 Maximum Marks : 100 Duration : 3 hours

Use separate answer scripts for each part.

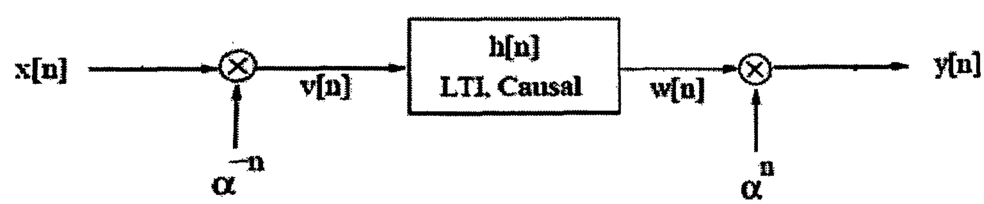
Part - I

Answer any FIVE questions: $5 \times 10 = 50$

1. Consider a causal linear time-invariant system with impulse response $h(n)$. The Z-transform of $h(n)$ is:

$$H(z) = \frac{1 - z^{-1}}{1 - 0.5z^{-1}}$$

Consider the cascade configuration given in the figure below. Here we assume that α is a real number.



(a) Draw the pole-zero plot of H(z) and also sketch the region of convergence. Is the system stable? Briefly justify.

(b) Assume that $\alpha = \frac{1}{3}$ and $x(n) = \left(\frac{1}{4}\right)^n u(n)$. Compute the output $y(n)$ for this particular choice of the input. (1+1+2)+6 = 10

2. Compute the eight-point DFT of the sequence:

$$x(n) = \begin{cases} 1, & \text{for } 0 \le n \le 7 \\ 0, & \text{otherwise.} \end{cases}$$

by using the decimation-in-time FFT algorithm. Show your steps and the signal-flow graph.

3. The Z - transform of the sequence $x(n) = u(n) - u(n-7)$ is sampled at five points on the unit circle as follows:

$$x(k) = X(z)\big|_{z=e^{j2\pi k/5}} \qquad k = 0,1,2,3,4$$

Determine the inverse DFT $x'(n)$ of $X(k)$. Compare it with $x(n)$ and explain the results.

4. Determine the circular convolution of the following two sequences and show your steps:

$$x_1(n) = \{\underset{\uparrow}{1},2,3,1\} \text{ and } x_2(n) = \{\underset{\uparrow}{4},3,2,2\}$$

5. A discrete-time system with input $x(n)$ and output $y(n)$ is described in the frequency domain by the relation:

$$Y(\omega) = e^{-j\omega}.X(\omega) + \frac{dX(\omega)}{d\omega}.$$

(a) Compute the response of the system to the input $x(n) = \delta(n)$.
(b) Check if the system is LTI and stable. 6+4 = 10

6. Consider the FIR filter:

$$y(n) = x(n) - x(n-4).$$

a) Compute and sketch the magnitude and phase of the frequency response of the filter's system function.
b) Compute its response to the input:

$$x(n) = \cos\left(\frac{\pi}{2}n\right) + \cos\left(\frac{\pi}{4}n\right), \qquad -\infty < n < \infty$$ 5+5 = 10

7. What do you mean by a white noise signal? Prove that for a zero mean white-noise input, the cross-correlation between input and output of an LTI system is proportional to the impulse response of the system. 2+8 = 10

8. a) Suppose a filter is given with the transfer function:

$$H(z) = 1 + z^{-1} + z^{-2} + \ldots + z^{-8}.$$

If the sampling frequency is $F_s = 1\,\text{kHz}$, determine the frequencies of the analog sinusoids that will be severely attenuated by the filter.

b) Show that any discrete-time signal $x(n)$ can be expressed as:

$$x(n) = \sum_{k=-\infty}^{\infty} [x(k) - x(k-1)]u(n-k)$$

where $u(n - k)$ is a unit step delayed by k units in time, that is:

$$u(n-k) = \begin{cases} 1, & n \geq k \\ 0, & \text{otherwise.} \end{cases}$$ 6+4=10

Part - II

Answer any FIVE questions from Q1 to Q6. Answer any TWO questions from Q7 to Q9.

1. Assuming RGB colour cube represents more than sixteen million colours, write an algorithm to find sixteen representative colours for the entire set of more than sixteen million colours. [6]

2. Derive 2D image coordinates for a 3D object captured by a camera. Assume that the coordinates for the 3D object, 2D image plane and the viewer are defined in a 3D world coordinate system. Define orthographic and perspective transformations from the above derivation. [4+2=6]

3. Derive an expression based on which one single image mask (a sub-image) can be designed that can smooth an image and subsequently detect the high frequency information in the image. [3+3=6]

4. A circle of diameter 20 pixels is drawn at the center of the computer screen having dimension 320×240 pixels. Give algorithmic steps to transform the circle into an ellipse having major and minor radii of 40 pixels and 15 pixels respectively. Note that the ellipse is to be displayed such that the center of the ellipse now lies in the screen coordinate (100, 100). [4+2=6]

5. Compare DCT with FFT in transforming a 2D spatial image to frequency domain. How a band-pass image filter can be designed using 2D FFT? [3+3=6]

6. Find the pixel values of the 3^{rd} row of the following two image matrices after application of Laplacian 2^{nd} derivative operator. [3+3=6]

2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8

2	2	2	2	2	5	8	8	8	8
2	2	2	2	2	5	8	8	8	8
2	2	2	2	2	5	8	8	8	8
2	2	2	2	2	5	8	8	8	8
2	2	2	2	2	5	8	8	8	8

7. A blob shape in an image needs to be identified using Canny edge detector. The blob shape is then approximated using piecewise linear segments. Write algorithmic steps for the above two operations. [5+5=10]

8. How JPEG compresses the information contained in an image? Give brief details of each step. [10]

9. State TRUE or FALSE [10×1=10]

   a. Smoothing by mean filtering uses localized intensity dependent weights.
   b. RLE can compress information of a gray level image.
   c. Projection of a binary matrix may separate objects present in the matrix.
   d. Colour can be represented as a vector.
   e. Laplacian of Gaussian finds maximum of the signal.
   f. Median filter sharpens the signal contrast.
   g. 3D pixels or voxels can have 6-neighborhood connectivity in a digital 3D grid.
   h. Retaining high frequency components of image intensity causes loss of image details.
   i. CMY color values of a pixel can be obtained from the RGB values of the pixel.
   j. Object and background pixels in a 2D binary image must have different pixel connectivity.

Indian Statistical Institute
First semester Examination 2016
Course Name: MSQE First Year and M - Stat Second Year
Subject Name: Basic Economics

Date of Examination: 18/11/2016 **Maximum Marks – 60** **Duration: 2.5 Hours**

Answer all questions

1. Consider the following information regarding a simple Keynesian model for an open economy with government: $mpc_{yd} = 0.8, m = 0.1, t = 0.5, i = 0.6, \rho = 0$ and $\overline{T} - \overline{R} = 40$.

(i) State the assumptions needed to regard $(Y - T + R)$ as personal disposable income.

(ii) Using the expression for personal disposable income given in (i), compute the required value of $(a + \overline{T} + \overline{G} + \overline{X} - \overline{M})$ that will make the equilibrium Y equal to 1600 units.

(All notations in the above question have their usual meanings and all relations in the model are linear) [6 + 16 = 22]

2. (i) Consider a Simple Keynesian Model for an open economy without government activities. Suppose 20 per cent and 60 per cent of aggregate planned consumption expenditure and aggregate planned investment expenditure, respectively, are spent on imported goods. Suppose the marginal propensity to consume and marginal propensity to invest (net) with respect to NDP (Y) are 0.8 and 0.3, respectively. Given the above information, answer the following questions under the assumption that all the relations in the model are linear:

a. Compute the marginal propensity to import with respect to Y and check if the equilibrium is stable.

b. Compute the autonomous expenditure multiplier in this model.

(ii) Suppose in a simple Keynesian model for a closed economy with government C = 0.75Y, G = 0 and I = 2000 – 100i, i_c= 8, m = 1 and m_d = 1 (where i_c denotes the repo rate and m_d and m are the relevant mark-ups; all other symbols have their usual meanings). Write down the equation of the aggregate demand function. Derive the equilibrium value of Y. To attain full employment, Y has to rise by 2000 units. By how much the central bank has to lower the repo rate to achieve full employment? **[7+3+12=22]**

3. (i) Consider a simple Keynesian model for a closed economy with government. Present the model in terms of saving and investment. Indicate the two components of aggregate planned saving.

(ii) The tax function is given by $T = \overline{T} + tY$. Suppose the government raises $\overline{T}$. What impact is it likely to have on private/personal saving, revenue deficit of the government and total saving of the economy?

(iii) Suppose the government intends to reduce revenue deficit by reducing $\overline{G}$. What kind of impact is it likely to have on Y and revenue deficit? In the light of the results derived, comment on the Fiscal Responsibility and Budget Management Act (2003) of Government of India.

(Assume government transfers to be a lump sum in answering the question). **[6+8+8=22]**

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2016 – 17

MStat (2^{nd} Year)

Financial Econometrics

Date: 21 November 2016 Maximum Marks: 100 Duration: 3 Hours

This paper carries 105 marks. Attempt ALL questions. The maximum you can score is 100. Be brief and to the point, use examples whenever applicable. You may use standard textbook notation automatically.

1. What is the Efficient Market Hypothesis? What are the different versions of this? Are these testable – answer with a specific example of model. [3 + 6 + 6 = 15]

2. (a) Define the Cowles – Jones statistic (CJ) for the sequences and reversals test for Random Walk. In the usual Markovian model, what are the situations when CJ = 1?

   (b) How is the variance ratio (VR) used in testing for Random Walk? Why do the weights decline linearly in
   $$VR(q) = \frac{Var[r_t(q)]}{q\ Var[r_t]} = 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho(k)?$$

   (c) What is long range dependence? Give an example of a test statistic for identifying long range dependence. [(3 + 2) + (3 + 3) = (2 + 2) = 15]

3. (a) What is the effect of non-synchronous trading on stock prices and observed returns?

   Under the non-trading process defined by
   $$\delta_{it} = \begin{cases} 1\ (no\ trade) & with\ probability\ \pi_i \\ 0\ (trade) & with\ probability\ (1-\pi_i) \end{cases}$$
   $$X_{it}(k) \equiv (1-\delta_{it})\delta_{it-1}\delta_{it-2}\dots\delta_{it-k}, \quad k > 0$$
   $$= \begin{cases} 1 & with\ probability\ (1-\pi_i)\pi_i^k \\ 0 & with\ probability\ 1-(1-\pi_i)\pi_i^k \end{cases}$$

   and assuming that virtual returns have a linear one factor structure

$$r_{it} = \mu_i + \beta_i f_t + \epsilon_{it} \quad i = 1, \dots, N,$$

show how non-trading affects the estimated beta of a typical security.

(b) What are the components of bid-ask spread? Using Glosten's model, show that this creates a negative serial correlation in stock returns.

[(3 + 5) + (3 + 4) = 15]

4. (a) For measuring normal performance of the market, explain the role of the following models:
   (i) Constant mean return model
   (ii) Market model
   (iii) Factor model

   (b) Describe a method for testing of Cumulative Abnormal Returns (CAR) across securities and through time; mention relevant assumptions carefully.

[(3 X 3) + 6 = 15]

5. (a) What is the mean-variance portfolio optimization problem?

   (b) Prove the following consequence of the mean-variance portfolio optimization exercise:

   For a multiple regression of the return on any asset or portfolio R_a on the return of any minimum-variance portfolio R_p (except the global minimum - variance portfolio) and the return of its associated orthogonal portfolio R_{op};

$$R_a = \beta_0 + \beta_p R_p + \beta_{op} R_{op} \, ;$$

   will satisfy (i) $\beta_0 = 0$ and (ii) $\beta_p + \beta_{op} = 1$.

   (c) Show that the intercept of the excess-return market model is zero if the market portfolio is the tangency portfolio.

[4 + 6 + 5 = 15]

6. What is the exact factor pricing model? What are the alternative versions of this model? How are they tested in practice?

[3 + 4 + 8 = 15]

7. (a) What is Brownian Motion? How do you show that this is a limiting version of a discrete time Random Walk?

   (b) Explain the Maximum likelihood and Generalised method of moments methods of parameter estimation for an option pricing model – use a specific example.

[(3 + 4) + (4 + 4) = 15]

M.Stat. II / Branching Processes

Final Exam. / Semester I 2016-17

Time - 3 hours

Maximum Score - 50

Date: 21-11-2016

NOTE : SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.

1. Let $M(\underline{s}, n) = ((m_{ij}(\underline{s}, n)))$, $1 \leq i, j \leq p$, where p is the number of types, $m_{ij}(\underline{s}, n) = \frac{\partial f_i(\underline{s}, n)}{\partial s_j}$, and $\underline{f} = (f_1, \ldots, f_p)$ is the generating function of the multi-type (p-type) non-singular and positive regular offspring distribution corresponding to a Branching chain.

   (a) (3+2=5 marks) Show that $M(\underline{s}, n) = M(\underline{f}(\underline{s}, n-1))M(\underline{s}, n-1)$, where $M(\underline{s})$ is used to denote $M(\underline{s}, 1)$. Thus, prove that $M(\underline{q}, n) = M^n(\underline{q})$, where q is the extinction probability vector.

   (b) (3+3=6 marks) Show that $M(\underline{s}, n) \to \mathbf{0}$ for $\underline{s} < \underline{1}$. Denote $\lambda(\underline{s})$ as the largest eigenvalue of $M(\underline{s})$ and ρ as the largest eigenvalue of $M(\underline{1})$. Show that $\lambda(\underline{q}) < 1$, when $\rho \neq 1$.

2. Let $\{X(n)\}$ be a p-type nonsingular and positive regular Branching chain with $E\|X(1)\|^2 < \infty$.

   (a) (4 marks) If $\rho = 1$, then show that $\lim_{n\to\infty}(P(X(n) \neq \underline{0} \mid X(0) = \underline{i}) = \frac{\underline{i} \cdot \underline{u}}{\underline{v} \cdot \mathbf{Q}(\underline{u})}$. where $\underline{u}$, $\underline{v}$ and $\mathbf{Q}[\cdot]$ are as defined in the class.

   (b) (5 marks) If $\rho > 1$, show that $\lim_{n\to\infty} \frac{X_j(n)}{X_1(n)+\cdots+X_p(n)}$ exists a.s. Identify the limit.

3. Let $\{X(t)\}$ be a one-dimensional continuous Branching chain with generator $u(s) = a(f(s) - s)$, where f is the p.g.f. of an offspring distribution $\{p_j\}$. Let q be its extinction probability.

   (a) (5 marks) If $u'(1) \neq 0$, then show that $\lim_{t\to\infty} e^{\beta t}[F(s,t) - q] = A(s)$, $0 \leq s < 1$, where the derivative of $A(s)$ is given by $\lim_{t\to\infty} e^{\beta t}\frac{\partial F(s,t)}{\partial s}$, with $\beta = u'(q)$.

   (b) (5 marks) If $u'(1) > 0$ and $\sum_j p_j j \log j < \infty$, show that $\lim_{s\uparrow 1}(1-s)^\alpha A(s)$ converges to a constant, Say K, where $\alpha = \beta/u'(1)$.
   [Hint: You may use $\frac{A(y)}{A(x)} = (\frac{1-x}{1-y})^\alpha \exp\{-\alpha\int_x^y [\frac{m-1}{f(s)-s} + \frac{1}{1-s}]ds\}$, for $q < x < y < 1$,]

P.T.O

(c) (5 marks) Let $\phi(y) = E(\exp\{-yW\})$, where $W = \lim_{t\to\infty} e^{-\lambda t} X_t$ a.s., with $\lambda = u'(1)$. Show that $\phi(u) = \int_0^\infty af[\phi(ue^{-\lambda y})]e^{-ay}dy$.

(d) (5 marks) Show that $(K/A(x))^{1/\alpha} = \phi^{-1}(x)$, by showing $F(s,t) \to \phi(\eta)$, as $t \to \infty$, with $s = \exp\{-\eta e^{-\lambda t}\}$ and using the relation $A(F(s,t)) = e^{-\beta t}A(s)$.

4. Let X_n be the number of offspring in the nth generation (of a Branching process) with $X_0 = 1$, where offspring distribution given by the probabilities $\{p_j\}$.

(a) (5 marks) If $p_j = c/((j+2)\log(j+2))^2$, for $j \geq 1$, where c is such that $\sum_{j\geq 1} p_j = 1$, does there exist a sequence of numbers $\{C_n\}$ going to infinity (as $n \to \infty$), such that, $X_n/C_n \to W$, where $P(W > 0) = 1$? Justify your answer.

(b) (5 marks) If $p_j = c/(j+1)^3$, for $j \geq 1$, where c is such that $\sum_{j\geq 1} p_j = 1$, show that $(X_n - m^n W)/\sqrt{X_n} \to N(0, \sigma^2/m(m-1))$, in distribution, where σ^2 and m are the variance and the mean corresponding to $\{p_j\}$.

(c) (5 marks) A family name existed for 500 years. Taking 25 years a generation, determine the probability (justifying your steps) that the family will stay another 1000 years, given that the expected no. of offspring who bears the name is 1 and has finite variance, say σ^2.

All the best.

Indian Statistical Institute

First Semester Exam: 2016-17

M.Stat. Second year, Statistical Genomics,

Maximum Marks: 100 Duration: 3 hours.

DATE: 21-11-2016

Answer all questions. Show your works to get full credit.

1. (a) What is meant by genomic imprinting? How does it affect the inference of genome-wide association studies (GWAS)? 3+2=5
(b) Why Fused Lasso is more effective for dimension reduction in GWAS compared to the ordinary Lasso? Explain the terms "Epistasis" and "Pleiotropy" in the context of GWAS. 5+5=10

2. (a) Sam and Liz are half-sibs sharing a common mother. Sam's father remarried Victoria and they had a son Matthew. Find the transition probability matrix of the relation between (i) Sam and Matthew, and (ii) Liz and Matthew. 3+3=6
(b) In the light of the Bioinformatics paper written by Loureno et al. (2011) propose a robust linear regression method for association studies. How will you perform a multiple SNP analysis using this method? Define "broad sense heritability". 4+3+2=9

3. (a) Consider genotypes with three alleles with respect to ABO blood groups. We have three alleles A, B, and O with frequencies p, q and r respectively. Derive the parent-child transition probability matrix in this setting. 10
(b) Suppose in a GWA study, there are 10 SNPs (each with two alleles) and the goal is to study the effects of all 2-factor and 5-factor interactions. Marginal effects of the SNPs are already studied. Propose a multifactor dimensionality reduction (MDR) algorithm for this study. 10

4. (a) "Functional GWAS is a more powerful statistical approach than the traditional GWAS", explain. Mention one strong limitation of functional GWAS. 5+5=10
(b) Suppose in a case-control based GWA study, rather than testing the gene-gene interactions, the researchers are interested in allowing interactions with the other genetic or environmental factors while testing for association at a given genetic locus.
(i) What is the most suitable statistical method for such study?
(ii) Considering 4 SNPs (each with two alleles) clearly write down the algorithm of the method.
(iii) How can you improve the predictive power of the method proposed in (i)? 5+10+5=20

5. The cardiovascular system of humans depend on the following 8 correlated response variables: (i) systolic blood pressure, (ii) diastolic blood pressure, (iii) blood sugar level, (iv) total cholesterol level, (v) Na level, (vi) K level, (vii) oxygen in the blood, and (viii) B-type natriuretic peptide (BNP).

Suppose these responses are measured for n subjects longitudinally at T evenly spaced time points. A marker genotype with two alleles, M and m, is observed for all subjects. The goal is to detect a possible QTL which control all these responses and linked to the observed marker gene. Consider an F_2 design.

(i) Write down the joint likelihood function for all subjects, responses and time points.
(ii) Test the existence of a QTL temporally controlling all the responses.
(iii) Test the significance of the linkage disequilibrium between the marker and the QTL. 5+10+5=20

Indian Statistical Institute

Semestral examination : (2016-17)

M. Stat. II year

Functional Analysis

Date : 23/11/16 Maximum marks : 60 Duration : 3 hours.

Answer ALL questions. The maximum you can score is 60.

Notations are as used in class.

(1) Let $1 \leq p < \infty$ and suppose that $\{a_{ij}, i, j = 1, \ldots, \infty\}$ are complex numbers such that $\sum_j a_{ij} f(j)$ converges for every element $f = (f(1), f(2), \ldots) \in l^p(I\!N)$. Prove that there is a bounded linear operator T from $l^p(I\!N)$ to $l^p(I\!N)$ satisfying $(Tf)(i) = \sum_j a_{ij} f(j)\ \forall f$ and $i \geq 1$. [10]

(2) Let $T \in \mathcal{B}(X)$ where X is a Banach space. Furthermore, suppose that there is some $\beta > 0$ such that $\mathrm{Re}(\lambda) > \beta$ for all $\lambda \in \sigma(T)$. Prove that $\|\exp(-tT)\| \to 0$ as $t \to \infty$. Hence or otherwise prove that for every nonzero $x \in X$, $\lim_{t \to \infty} \|\exp(tT)(x)\| = \infty$. [7+5=12]

(3) Let $T \in \mathcal{B}(X, Y)$ where X, Y are Banach spaces, with X reflexive and X^* separable. Suppose that $\|T(x_n) - T(x)\| \to 0$ whenever x_n, x in X such that $x_n \to x$ weakly, i.e. $\phi(x_n) \to \phi(x)$ for every $\phi \in X^*$. Prove that T is a compact operator. [Hint : Use Banach-Alaoglu Theorem] [15]

(4) Let A, B be two positive operators on a Hilbert space $\mathcal{H}$. Prove that the spectrum of AB must be contained in $[0, \infty]$. Is AB always positive? Justify your answer with argument and/or counterexample. [6+6=12]

(5) Suppose that U is a unitary operator on a Hilbert space $\mathcal{H}$, i.e. $U^*U = UU^* = I$. Moreover, assume that -1 is not in $\sigma(U)$. Define $T = i(U-I)(U+I)^{-1}$. Prove that T is self-adjoint. Using the spectral theorem for bounded self-adjoint operators, prove that there is a spectral measure P on $\sigma(U)$ such that $U = \int_{\sigma(U)} \lambda dP$. [5+12=17]

INDIAN STATISTICAL INSTITUTE
M. Stat. II Year (2016-17), I semester
Semestral Examination
PATTERN RECOGNITION

Date: 23·11·2016 Duration: 210 minutes Maximum Marks: 100

Note: This paper carries 105 marks. Answer as many questions as you can.

1. (a) Describe a density based clustering algorithm for data sets.
   (b) Describe any two feature selection algorithms when a criterion function for feature selection is given. [4+8=12]

2. Suppose you have two 2-dimensional normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$ where $\mu_1 = (0.0, 0.0), \mu_2 = (1.0, 2.0), \Sigma = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 2 \end{pmatrix}$. Let the prior probabilities of the populations be 0.5 and 0.5.

   (a) Find the Mahalanobis distance between the two populations.
   (b) Find the Bayes decision rule for separating the two populations. Also find its probability of misclassification in terms of standard normal probabilities. [5+(3+7)=15]

3. Suppose you are given a training set for a two class classification problem. It is also known that the classes are linearly separable. Describe perceptron algorithm for finding a separating hyperplane between the two classes. [12]
4. Suppose there is one hidden layer with J number of nodes in an MLP, and sigmoid function is used as transfer function. Suppose you are using online learning algorithm. Let a training dataset be given to you and let the number of classes be 3. Then
   (a) Write down the expression for the error for MLP.
   (b) Write down the expression for the change in the connection weight joining the i-th node in the hidden layer to the second node in the output layer. [3+7=10]
5. Let the prior probabilities in a c-class classification problem be $P_1, P_2, \ldots, P_c$. Show that the misclassification probability for the Bayes decision rule is $\leq 1 - \max\{P_1, P_2, \ldots, P_c\}$. [6]
6. Write short notes on the following.
   (a) Probabilistic separability measures for feature selection
   (b) VC dimension
   (c) Support vectors
   [5+5+5=15]

(P.T.O)

7. Assume that the sample dispersion matrix Σ is formed on the basis of n observations where the number of features is M, $M>>n$. Suggest a procedure for finding eigen values and eigen vectors of Σ, with the help of an $n \, X \, n$ matrix. [5]

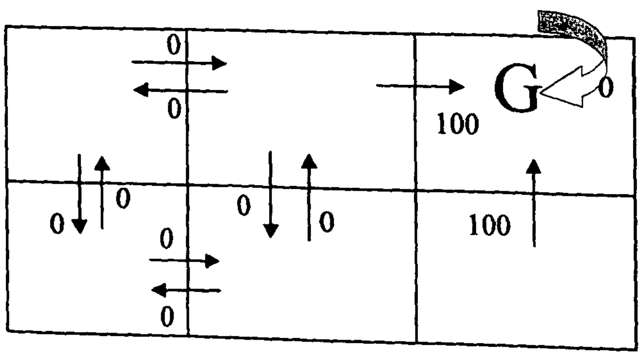
8. (a) Consider the following data table for 15 customers for whom the credit card application is either approved or rejected. Each customer has 4 attributes and belong to either the "accepted" class or the "rejected" class.

ID	Age	Has_Job	Own_House	Bank_Balance	Class (Credit card application status)
1.	Young	False	False	Medium	Rejected
2.	Young	False	False	High	Rejected
3.	Young	True	False	High	Accepted
4.	Young	True	True	Medium	Accepted
5.	Young	False	False	Medium	Rejected
6.	Middle	False	False	Low	Rejected
7.	Middle	True	True	High	Accepted
8.	Middle	False	True	High	Accepted
9.	MIddle	False	False	Medium	Rejected
10.	Middle	False	True	Low	Rejected
11.	Old	False	True	High	Accepted
12.	Old	True	False	Medium	Accepted
13.	Old	False	True	Medium	Accepted
14.	Old	False	False	Low	Rejected
15.	Old	True	False	High	Accepted

(i) Based on the information gain measure, which attribute is the best choice for being the root of the decision tree that can be built from this data? Explain your choice showing necessary calculations.

(ii) Draw the corresponding decision tree and explain how the class of a new customer with the following attributes can be decided on the basis of this tree:
Age = "Middle", Has_Job = "False", Own_house = "True", Bank_Balance = Medium". [6+(3+2)=11]

(b) Consider the following grid-map of the environment of a small robot which has to reach the goal state as has been shown by the letter G starting from any initial cell.



(P.T.O.)

Immediate rewards are marked beside each transition. Assume that the immediate reward for a transition from the goal state to itself is zero.

i) Assuming the discount factor $\gamma = 0.8$, show how the Q matrix changes for the first two episodes of a simple Q learning algorithm.

ii) Qualitatively sketch an optimal policy for the robot. [7+2=9]

9. (a) Describe the optimization criterion for Fuzzy C means algorithm.
   (b) Describe Fuzzy C-means algorithm. [5+5=10]

INDIAN STATISTICAL INSTITUTE
Semestral Examination: 2016-2017
M. STAT, II Year
Time Series Analysis

Date: 25 – *th* November, 2016 Maximum Marks 50

Duration 2 hours

All notations are self-explanatory. This question paper carries a total of 50 marks. Marks allotted to each question are given within parentheses.

1. Consider the following p th order autoregressive process:
$W_t = \rho_1 W_{t-1} + \rho_2 W_{t-2} + \cdots + \rho_p W_{t-p} + X_t, t = 1, 2 \ldots, T$, where $\{X_t, t = 1, 2 \ldots,\}$ is a sequence of $i.i.d$ Gaussian random variables with mean zero and with positive variance σ^2.
(a) Derive the stationary conditions for the above process.
(b) Write down the *exact likelihood* for the above process.
[5+5=10]

2. Let $\{X_t, t = 1, 2 \ldots,\}$ be a sequence of independent normal random variables with mean zero and with unit variance. Consider the following process:

$W_t = 10 + 1.5W_{t-1} - 0.56W_{t-2} + X_t.$

(a) Show that the process is stationary. (b) Find the mean and variance of the process W_t. **[5+ (3+7) =15]**

3. Consider the following autoregressive process:
$W_t = \rho W_{t-1} + X_t, \quad \rho < 1, t = 1, 2 \ldots, T$, where $\{X_t, t = 1, 2 \ldots,\}$ is a sequence of $i.i.d$ random variables with mean zero and with positive variance σ^2. Assume that $\{X_t, t = 1, 2 \ldots,\}$ has finite fourth moments.

(a) Assume that the above times series is stationary. Consider the OLS estimator of ρ. Show that the OLS estimator, $\hat{\rho}$, is biased but consistent.

(b) How will you test for $H_0 : \rho = 1$ against $H_0 : \rho < 1$? Derive analytically the asymptotic distribution of your test statistic.
[8+10 =18]

4. Consider the following time series process:
$Y_t = \sum_{i=1}^{k}\{A_i sin(t) + B_i cos(t) + \epsilon_{it}\}$, where A_i, B_i and ϵ_{it} are $i.i.d$

P.T.O

random variables with mean zero and with positive variance σ^2. All these random variables are independent to each other. Examine if the time series process is stationary or not. **[7]**

Indian Statistical Institute
Semestral Examination
First Semester, 2016-2017 Academic Year
M.Stat. 2nd Year
Statistical Inference II

Date: 28.11.2016 | Maximum Marks : 65 | Duration: $3\frac{1}{2}$ Hours

Answer all questions

1. (a) Define a highest posterior density (HPD) credible region for an unknown parameter.

   (b) Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent random samples, respectively, from $N(\mu, \sigma_1^2)$ and $N(\mu, \sigma_2^2)$, where both σ_1^2 and σ_2^2 are known. Construct a $100(1-\alpha)\%$ HPD credible interval for the common mean μ assuming a uniform prior. Compare this with the frequentist $100(1-\alpha)\%$ confidence interval for μ. [2+4=6]

2. Suppose $X_1, \ldots, X_n$ are iid $N(\theta, 1)$ where θ is unknown. Suppose we are interested in the model selection problem with two candidate models M_0 and M_1. Under M_0, θ equals 0 and under M_1, $\theta \in \mathcal{R}$. Suppose each model is given equal a-priori probability and under model M_1, one puts a $N(0, \tau^2)$ prior for θ where τ^2 is very large. Is this a reasonable specification of prior for this problem ? Justify your answer in concrete terms. [6]

3. (a) Define Arithmetic Intrinsic Bayes Factors (AIBF) and Fractional Bayes Factors (FBF). What is the reason for considering such Bayes Factors ? Given an example where the FBF is clearly unacceptable but the AIBF is not so. Justify your answer.

   (b) Consider a negative binomial experiment. Bernoulli trials, each having probability θ of success, are independently performed until a total of n successes are accumulated. You are given as data $X_1, \ldots, X_n$, where X_1 denotes the number of failures before the first success and for $i = 2, \ldots, n$, X_i denotes number of failures between the $(i-1)$th sucess and ith success. On the basis of this data, we want to select between two models $M_0 : \theta = \frac{1}{2}$ and $M_1 : \theta \in (0, 1)$. We consider the Jeffreys prior $\pi(\theta) = \theta^{-1/2}(1-\theta)^{-1}$, $0 < \theta < 1$ under M_1. Find the intrinsic prior in this problem. [(2+2+4)+6=14]

4. Consider p independent random samples, each of size n from p similar normal populations $N(\theta_j, \sigma^2)$, $j = 1, \ldots, p$ where σ^2 is known. Assume p is quite large. Our problem is to estimate $\theta_1, \ldots \theta_p$. We assume that $\theta_1, \ldots, \theta_p$ are iid $N(\eta_1, \eta_2)$, where (η_1, η_2) has a prior distribution $\pi(\eta_1, \eta_2)$. Argue in favour of modelling of this kind in such situations. Will the resulting Bayes estimators (with respect to squared error loss) be more appropriate compared to the MLE in this situation ? Explain mathematically. [2+5=7]

5. (a) Suppose we want to approximate $E_f(h(X)) = \int h(x)f(x)dx$ by importance sapmling, where f is a probability density. Consider the class of importance functions $u(x)$ such that $E_u\left[h^2(X)\frac{f^2(X)}{u^2(X)}\right] < \infty$. Argue mathematically that among such functions $u(x)$, the ones for which $u(x)/(|h(x)|f(x))$ is approximately constant are expected to give the best results in terms of precision of the resulting estimators of $E_f(h(X))$.

(b) Suppose one has to sample from the density $\frac{g(\theta)}{K}$ where $\theta \in [0,1]$ and $K > 0$ is a normalizing constant and $g(\theta)$ is given by

$$g(\theta) = \exp(-2\theta)\theta^{\alpha-1}(1-\theta)^{\beta-1}I\{0 \le \theta \le 1\},$$

$\alpha > 0$, $\beta > 0$ being known constants. Describe the Metropolis-Hastings algorithm for simulation from the described density. Explain why in independence sampler Metropolis-Hastings algorithm, having a proposal distribution which is lighter-tailed than the target distribution may not be ideal. [5+4+2=11]

6. (a) Let $X_1, \ldots, X_n$ be i.i.d. $\sim f(x|\theta)$, $\theta \in \mathcal{R}$ and $\pi(\theta)$ be a prior density of θ. Show that the Bernstein-von Mises Theorem about posterior distribution of $\sqrt{n}(\theta - \hat{\theta}_n)$ implies consistency of the posterior distribution of θ at θ_0, where θ_0 is the true value of the parameter. You may assume the regularity conditions needed to prove the Bernstein-von Mises Theorem.

(b) Show that under the set up described in part (a)

$$\begin{aligned}\log \int_R \prod_1^n f(X_i|\theta)\pi(\theta)d\theta &= \sum_{i=1}^n \log f(X_i|\hat{\theta}_n) - \frac{1}{2}\log n + \frac{1}{2}\log(2\pi) \\ &\quad -\frac{1}{2}\log I(\theta_0) + \log \pi(\theta_0) + \mathrm{R}_n,\end{aligned}$$

as $n \to \infty$, where $\hat{\theta}_n$ is the MLE and $I(\theta_0)$ is Fisher information number at θ_0 and R_n converges to zero almost surely under θ_0.

(c) State the condition used to handle the tail of the likelihood in the proof of the Bernstein-von Mises theorem. Is the condition satisfied if $X_1, \ldots, X_n$ are iid $N(\theta, 1)$ where $\theta \in \mathcal{R}$? Justify your answer.

(d) Give an example where the posterior mean of a parameter is not a consistent estimator of it. Prove your answer. [4+6+2+4+5=21]

INDIAN STATISTICAL INSTITUTE

First Semester Back paper Examination : 2016 – 17

MStat (2^{nd} Year)

Financial Econometrics

Date: 27 · 12. 2016 Maximum Marks: 100 Duration: 3 Hours

Attempt ALL questions

You may use all standard notation automatically

1. (a) What is a dividend payment. What is its effect on the stock price?

(b) As we know that stock price returns are thick tailed, Normal distribution is not a reasonable approximation. Does Stable distributions help?

[(5 + 3) + 9 = 17]

2. How are the Random Walk Hypotheses RW1, RW2 and RW3 related? Use a Venn diagram for your answer and provide specific examples. [16]

3. Suppose the trading process $\{\delta_{it}\}$ defined by

$$\delta_{it} = \begin{cases} 1 \; (no\ trade) & with\ probability\ \pi_i \\ 0 \; (trade) & with\ probability\ (1-\pi_i) \end{cases}$$

were not *iid*, but followed a two state Markov chain with transition probabilities

$$\begin{array}{cc} & \delta_{it} \\ & \begin{array}{cc} 0 & 1 \end{array} \\ \delta_{it-1} \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \pi_i & (1-\pi_i) \\ (1-\pi_i') & \pi_i' \end{pmatrix} \end{array}$$

Derive the unconditional mean, variance and first order auto covariance of δ_{it} as functions of π_i and π_i'. [4 + 6 + 7 = 17]

4. Distinguish between statistical and economic models in the context of event study analysis. Use a specific example for your answer. [16]

5. In the context of the mean-variance portfolio optimization problem, derive the efficient frontier. [17]

6. What is an Ito process? Explain how the Ito's lemma help in obtaining the price of a European Call option. State the relevant assumptions carefully.

[6 + 11 = 17]

Indian Statistical Institute
Back Paper Examination 2016
Course Name: MSQE First Year and M-Stat Second Year
Subject Name: Basic Economics

Date of Examination: 28/12/16 **Full Marks – 100** **Duration: 3 Hours**

Answer any four questions

1. Consider the following problem of national income accounting: For producing current period output firm B incurs the following costs: It purchases goods worth Rs.50,000 from firm A, holds half of it in inventory and uses the other half as raw material for current period production. It pays out Rs.40,000 in wages half of which is paid to a labour contractor. It pays Rs.35000 as interest to a bank, Rs.5,000 as interest on bonds sold to households, Rs.15,000 as rent to households, Rs.5,00,000 as dividend to households, Rs.20,000 in profit tax, Rs.75,000 in net indirect taxes and spends Rs10,000 to buy the prime minister's used pen. Its undistributed profit is Rs.2,00,000.

(i) Compute the intermediate input cost incurred by firm B, the value of output produced by firm B and the gross value added of the same firm.

(ii) Compute firm B's contribution to national income.

(iii) Compute firm B's contribution to personal income. [9+8+8=25]

2.a) Explain the concept of paradox of thrift in the simple Keynesian model.

b) Suppose in a simple Keynesian model planned consumption is a proportional function of NDP denoted Y. Marginal propensity to invest with respect to Y is 0.3. Start with an initial equilibrium situation. Suppose there takes place a parallel downward shift in the saving function by 3 units and following this, saving in the new equilibrium is found to increase by 9 units. Derive the new consumption function. (All relations are linear). [12+13=25]

3. Consider a simple Keynesian model for a closed economy without government. At NDP (Y) = 1000 units, producers have to sell 20 units from their stock to meet the customers' demand fully. It is given that the equilibrium level of GDP is 1040 units. Find out the impact of an increase in autonomous expenditure on the equilibrium level of Y, assuming the expenditure function to be linear. [25]

4. Suppose the tax function in a simple Keynesian model for a closed economy with government is given by $T = \overline{T} + tY$, $\overline{T} > 0$, $0 < t < 1$ and the transfer payment is a lump sum. Present the model in terms of the relevant equations. What can you say about the impact of an exogenous increase in $\overline{T}$ on the total tax collection? Explain your answer. **[25]**

5. Monetary policy of the government that seeks to revive the economy by reducing the policy rate of the central bank is unlikely to yield the desired impact during recession. It may even intensify the recessionary forces. Do you agree? Explain your answer. [25]

Indian Statistical Institute
Backpaper Examination
First Semester, 2016-2017 Academic Year
M.Stat. 2nd Year
Statistical Inference II

29 · 12 · 16

Full Marks : 100

Duration: 3 Hours

Answer all questions

1. Suppose, given $\theta \in (0,1)$, $X_1, \ldots, X_n$ are iid Bernoulli(θ) random variables for each $n \geq 1$. Assume that θ has a uniform prior distribution on the open interval $(0,1)$. Show that the probability of the event $\{X_{n+1} = 1\}$ given that $X_i = 1$ for all $1 \leq i \leq n$, increases to 1 as n increases to infinity. [10]

2. (a) Suppose $X_1, \ldots, X_n$ are iid with common density $f(x|\theta)$ where $\theta \in \mathcal{R}$. Suppose $\theta = \theta_0$ is the true value of θ. Assuming appropriate conditions and using an appropriate version of the asymptotic normality of posteriors under such conditions, prove that $\sqrt{n}(\tilde{\theta}_n - \hat{\theta}_n) \to 0$ with probability one (under P_{θ_0}), where $\tilde{\theta}_n$ and $\hat{\theta}_n$ denote respectively the posterior mean and MLE.

   (b) Suppose now that $X_1, \ldots, X_n$ are iid $N(\theta, 1)$ and $\theta \sim \pi(\theta)$. Suppose $\theta = \theta_0$ is the true value of θ. Prove that under appropriate conditions to be stated by you, the difference between the posterior mean and the MLE, upon proper rescaling, converges in distribution (under P_{θ_0}) to a non-degenerate limit. [10+16=26]

3. Define the Jeffreys prior. Describe in a few sentences the justification of thinking of Jeffreys prior as a noninformative/low information prior. [10]

4. (a) Argue that use of improper non-informative priors in model selection problems may lead to Bayes factors which are defined only upto arbitrary multiplicative constants. Can the use of diffuse proper prior be a solution to this problem ? Explain your answer.

(b) Consider the general model selection problem and derive the intrinsic prior determining equations starting with Arithmetic Intrinsic Bayes Factors. Argue that in the nested model selection problem, the solution suggested by Berger and Pericchi to these equations indeed satisfies the equations. [2+6+8+7=23]

5. Consider p independent random samples, each of size n from p normal populations $N(\theta_j, \sigma^2)$, $j = 1, \ldots, p$ where σ^2 is known. Our problem is to estimate $\theta_1, \ldots \theta_p$. We assume that $\theta_1, \ldots, \theta_p$ are iid $N(\eta_1, \eta_2)$, where $\eta_1 \in R$ and $\eta_2 > 0$ are unknown constants. Find the Parametric Empirical Bayes (PEB) estimate of the vector of θ's. Explain in what sense this estimator might be more appealing than the simple vector of sample means as the estimator of the vector of θ's in this setup. [9+4=13]

6. (a) Suppose our task is to evaluate $E_f(h(X)) = \int h(x)f(x)dx$ where f is a probability density such that it is very difficult to sample directly from f. Describe how the Importance Sampling technique can be useful to approximate $E_f(h(X))$. Suppose now we are given the class of importance functions $u(x)$ such that $E_u\left[h^2(X)\frac{f^2(X)}{u^2(X)}\right] < \infty$. Show that among this class, the importance function u which minimizes the variance of the importance sampling estimator (for any given value of m, the number of samples drawn for the approximation) is given by

$$u(x) = u^*(x) = \frac{|h(x)|f(x)}{\int |h(x)|f(x)dx}.$$

(b) Suppose we have a random observation X from a $N(\theta, 1)$ distribution and θ is assumed to have a Cauchy(μ, τ) distribution where both μ and τ are known constants. The problem is to approximate the posterior mean and variance of θ given X. Describe how one can approximate these quantities using the teachnique of Gibbs sampling. [2+8+8=18]

Indian Statistical Institute

Backpaper examination : (2016-17)

M. Stat. II year

Functional Analysis

Date : 30/12/16 Maximum marks : 100 Duration : 3 hours.

Answer ALL the questions. Each question carries 20 marks.

Notations are as used in class.

(1) Prove or disprove (with arguments in support of your answer):
The Banach space $C[0,1]$ cannot be isometrically isomorphic with a Hilbert space.

(2) With the notation introduced in class, prove that $c_0^* \cong l^1$, where $\cong$ means isometric isomorphism of Banach spaces.

(3) Let $A, B \in \mathcal{B}(\mathcal{H})$ satisfy $0 \leq A \leq B$, and moreover, assume that A, B are invertible. Show that $B^{-1} \leq A^{-1}$.

(4) Let $1 < p < \infty$. Prove that any bounded linear operator from l^p to l^1 is compact. You may assume the following fact : if a sequence f_n in l^1 converges weakly, then it also converges in the norm of l^1.

(5) Let $\mathcal{H}$ be a separable Hilbert space with an orthonormal basis $\{e_1, e_2, ...\}$. Consider the set Q consisting of vectors of the form $\sum_{n=1}^{\infty} c_n e_n$ with $|c_n| \leq \frac{1}{n}$. Prove that Q is a compact subset of $\mathcal{H}$.

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, First Semester, 2016-17

Semestral (Backpaper) Examination

30·12·2016

Time: $3\frac{1}{2}$ Hours | Statistical Computing | Full Marks: 100

1. (a) Let 0.3762, 0.7198, 0.9420 and 0.1505 be four random numbers independently generated from the $U(0,1)$ distribution. Using these random numbers, generate a random vector (X_1, X_2) from the bivariate distribution with cumulative distribution function given by

$$F(x_1, x_2) = 1 - \exp\{-x_1\} - \exp\{-x_2\} + \exp\{-x_1 - x_2 - 0.5x_1x_2\}, \quad \text{where } x_1, x_2 \geq 0.$$

[8]

(b) When a coin is tossed, 'Head' and 'Tail' appear with probability p and $1-p$, respectively $(0 < p < 1)$. Using this coin, is it possible to generate an unbiased estimate of $1/p$? If you think it is possible, write down the algorithm. If you think it is not possible, give reasons. [4]

2. (a) If $\mathbf{X} = (X_1, X_2, \ldots, X_d)'$ follows a spherically symmetric distribution, show that for any $\boldsymbol{\alpha} \in R^d$, $\boldsymbol{\alpha}'\mathbf{X}$ has the same distribution as $\|\boldsymbol{\alpha}\| X_1$. [4]

(b) Consider a uniform distribution on a unit sphere with centre at the origin. Calculate the mean and the covariance matrix of this distribution. Also calculate the α-trimmed versions of the mean and the covariance matrix. [3+5]

3. Consider a data cloud of the form $\{(x_i, y_i),\ i = 1, 2, \ldots, 21\}$, where $x_i = i - 11$ and $y_i = x_i^2 - 25x_i + 20$ for $i = 1, 2, \ldots, 21$.

(a) If $f(x) = 21x^2 + 11x + 8$ is the true regression function and $\hat{f}_h(\mathrm{x})$ is the Nadaraya-Watson estimate of f based on a Gaussian kernel with bandwidth h, show that there exists some $\mathrm{x}_0 \in [-10, 10]$ such that $\hat{f}_h(\mathrm{x}_0) = f(\mathrm{x}_0)$ [5]

(b) Derive the limiting value of $\hat{f}_h(4.5)$ when the bandwidth h shrinks to zero. [4]

(c) Find the limiting value of the local linear estimate of the regression function f at $x = 4.5$ when the kernel is Gaussian and the bandwidth h diverges to infinity. [5]

4. Prove or disprove the following statements.

(a) LMS (least median of squares) and LTS (least trimmed squares) estimators of location are one-dimensional analogs of MVE (minimum volume ellipsoid) (based on 50% observations) and MCD (minimum covariance determinant) estimators of location, respectively. [4]

(b) Consider a logistic linear regression problem involving a binary valued response variable Y and a p dimensional covariate $\mathbf{X}$. Let S_i be the convex hull formed by the $\mathbf{X}$-observations with $Y = i$ $(i = 0, 1)$. If S_0 and S_1 are disjoint, maximum likelihood estimate of the parameters of the logistic regression model will not exist. [6]

(c) Nearest neighbor estimate of a regression function is piecewise constant in nature. [4]

P.T.O

5. Consider a classification problem with J competing classes. If a node t of a classification tree contains p_j proportion observations from the j-th class ($j = 1, 2, \ldots, J$), its impurity function is given by $i(t) = \psi(p_1, p_2, \ldots, p_J)$, where ψ is concave and symmetric in its arguments.

(a) Show that (i) ψ is maximized when $p_j = 1/J$ for $j = 1, 2, \ldots, J$, (ii) it is minimized when $\max p_j = 1$ and (iii) for any split of a node t, the reduction in the impurity function, $\Delta i(t)$, is non-negative. [3+3+3]

(b) Consider a two-class problem (i.e., $J = 2$), where one needs to find the best split based on a categorical measurement variable C, which has three categories C_1, C_2 and C_3. The number of observations in these three categories are given below

Category	C_1	C_2	C_3	Total
No. of. observations from class-1	150	200	150	500
No. of. observations from class-2	150	0	350	500
Total	300	200	500	1000

Find the best split based on C. [5]

6. (a) Let $x_1, x_2, \ldots, x_n$ be n observations generated from a mixture of three univariate normal distributions $N(0,1)$, $N(1,1)$ and $N(-1,1)$ with mixing proportions p^2, $2pq$ and q^2 ($p+q=1$), respectively. Use the expectation-maximization (EM) algorithm (clearly state the choice for the initial value of p and the convergence criterion) to estimate p. [6]

(b) Show that the EM algorithm can be viewed as a special case of minorization-maximization algorithm. [4]

(c) The results of one-day international series played over a period two years are given below.

Home Team	AS	SR	IN	WI	NZ	AS	EN	SR	SA	EN	PK	WI	NZ	SA
Visiting Team	NZ	SA	EN	PK	SA	PK	WI	IN	EN	AS	SR	IN	PK	AS
Result	4-1	2-1	3-0	2-2	2-3	3-0	2-1	3-2	4-1	3-3	2-1	2-5	4-1	4-3

Describe how you will rank these 8 teams based on their performance in those two years using an appropriate Bradley-Terry model. [6]

7. (a) Suppose that there are $2m$ balls numbered $1, 2, \ldots, 2m$. These balls are randomly divided into two boxes. At any stage, we choose one of these $2m$ balls at random and move it to the other box. Let X_n ($n = 0, 1, 2, \ldots$) denotes the number of balls in the first box after the n-th stage.

i. Show that $\{X_n;\ n \geq 0\}$ is a Markov chain with state space $\{0, 1, \ldots, 2m\}$. [3]

ii. Show that $Bin(2m, 1/2)$ is the stationary distribution of the Markov chain. [6]

iii. Check whether $\sup_k |P(X_n \leq k) - \sum_{i=0}^{k} \binom{2m}{i} (\frac{1}{2})^k| \to 0$ as n diverges to infinity. [3]

(b) Use the Gibbs sampling algorithm to generate observations from the following distribution

$$f(x, y, z) = Cy^x(1-y)^{z-x}e^{-\lambda}\lambda^z/[x!\ (z-x)!], \quad z = 1, 2, \ldots;\ x = 0, 1, \ldots, z;\ 0 \leq y \leq 1,$$

where C is a normalizing constant. [6]

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, Second Semester, 2016-17

20 · 02 · 17

Statistical Computing II

Time : 3 Hours | Mid-Semetral Examination | Full Marks: 60

(Answer as many as you can. The maximum you can score is 60.)

1. Let $X_1, X_2, \ldots, X_n$ be independent observations with distribution function F with finite second moments. A bootstrap sample of size n is taken from this set of n observations. Let $\bar{X}_n^*$ be the mean of the bootstrap sample.

   (a) Show that $\bar{X}_n^*$ can be written as $\bar{X}_n^* = [\sum_{i=1}^n W_i X_i]/n$, where $(W_1, W_2, \ldots, W_n)$ follows a multinomial distribution with $\sum_{i=1}^n W_i = n$. [4]

   (b) Hence find the variance of $\bar{X}_n^*$ and check whether it is an unbiased estimator of $Var(\bar{X}_n)$, where $\bar{X}_n = \sum_{i=1}^n X_i/n$. [6]

   (c) Find the jackknife estimator for $Var(\bar{X}_n)$ and check whether it is unbiased. [6]

2. Let $X_1, X_2, \ldots, X_n$ be independent observations from a distribution with cumulative distribution function $F_\theta(t) = 1 - e^{-(t-\theta)}$, $t \geq \theta$. Define $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$. Let $X_1^*, X_2^*, \ldots, X_n^*$ and $X_1^{**}, X_2^{**}, \ldots, X_n^{**}$ be two independent bootstrap samples generated using parametric and nonparametric bootstrap methods, respectively. Define $X_{(1)}^* = \min\{X_1^*, X_2^*, \ldots, X_n^*\}$. and $X_{(1)}^{**} = \min\{X_1^{**}, X_2^{**}, \ldots, X_n^{**}\}$.

   (a) Let $G(\cdot)$ and $G^*(\cdot)$ denote the distribution functions of $X_{(1)}$ and $X_{(1)}^*$, respectively. Check whether $\sup|G(t) - G^*(t)|$ converges to zero almost surely as n tends to infinity. [6]

   (b) Let $H(\cdot)$, $H^*(\cdot)$ and $H^{**}(\cdot)$ denote the distribution of $n(X_{(1)} - \theta)$, $n(X_{(1)}^* - X_{(1)})$ and $n(X_{(1)}^{**} - X_{(1)})$, respectively. Check whether $\sup|H(t) - H^*(t)|$ and $\sup|H(t) - H^{**}(t)|$ converge to zero almost surely as n tends to infinity. [4+4]

3. Suppose that Y follows a distribution with probability density function $f(y)$ given by

$$f(y) = \exp\left\{\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right\}.$$

   Show that (i) $E(Y) = \frac{\partial b(\theta)}{\partial \theta}$, (ii) $Var(Y) = \phi \frac{\partial^2 b(\theta)}{\partial \theta^2}$ and (iii) $M_Y(t) = E(e^{tY}) = \exp\{\frac{b(\theta + t\phi) - b(\theta)}{\phi}\}$. [4+4+4]

4. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be independent and identically distributed as d-dimensional standard normal variates. Consider any fixed $\mathbf{x} \in R^d$ and any fixed $r > 0$. Let $B_r(\mathbf{x}) = \{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| \leq r\}$, a ball of radius r around $\mathbf{x}$ and $Y_n(\mathbf{x})$ be the nearest neighbor of $\mathbf{x}$ in the set $\{X_1, X_2, \ldots, X_n\}$. Show that

   (a) $P\{Y_n(\mathbf{x}) \in B_r(\mathbf{x})\} \to 1$ when d is fixed and n diverges to infinity. [6]

   (b) $P\{Y_n(\mathbf{x}) \in B_r(\mathbf{x})\} \to 0$ when n is fixed and d diverges to infinity. [6]

P.T.O

5. Some class-XII students of a school appeared at a state level mathematics examination. In order to prepare for this examination, some of the students attended some special training program organized by a private organisation, but others could not join it because of the high cost of the training program. The following data set shows the marks obtained by the students in the last mathematics examination conducted by the school, their genders, whether they attended the special program and the number of problems they correctly solved in the state level examination.

Serial Number	Gender	Marks in last exam	Whether attended the training program	Number of problems correctly solved
01	F	72	Attended	3
02	M	67	Not attended	2
03	M	55	Not attended	1
04	F	59	Attended	0
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
77	M	45	Not attended	1
78	M	95	Attended	4

Describe how you will fit a suitable generalized linear model to this data set. Describe how you will test whether the special training program was effective. [6+4]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: Semester-II
M. Stat. II Year: 2016-2017

Survival Analysis

Date: 20 February, 2017 Full Marks: 50 Duration: 2 hours

Note: Answer all questions

1. Suppose a continuous lifetime random variable T has hazard rate $\lambda(t)$, cumulative hazard function $\Lambda(t)$ and mean residual life function $m(t)$.
   (a) Find the distribution of $\Lambda(T)$.
   (b) Show that $m(t)$ fully specifies the distribution of T.

[4+8 =12]

2. Consider random right censoring with failure time $T \sim F$ and censoring time $C \sim G$ with $\bar{G} = \left[\bar{F}\right]^{\theta}$, $\theta > 0$. Assume T and C are indepndent and, both F and G are continuous. Prove that $X = \min(T, C)$ and $\delta = I\{T \leq C\}$ are independent.

[10]

3. Suppose that n identical items are put on a life test at time $t = 0$. The experiment is continued until a pre-fixed number $(r \leq n)$ of items have failed and the remaining items are censored at the r-th failure time. Suppose the lifetime follows exponential distribution with hazard rate λ.
   (a) Find the maximum likelihood estimate (MLE) of λ.
   (b) Verify whether the MLE of λ is unbiased or not.

[3+9=12]

4. Consider the following survival times in weeks of 21 patients.

$$1, 2, 2, 2, 6, 8, 8, 9, 13, 16, 17, 29, 34, 2+, 9+, 13+, 22+, 25+, 36+, 43+, 45+$$

   (a) Calculate the Kaplan-Meier estimate of survival function $S(t)$.
   (b) Is the survival function obtained in (a) a proper survival function? Justify your answer.

[7+3= 10]

5. Eighteen elderly individuals who had entered a nursing home in the past five years were asked when they experienced their first fall (post admitted). Some of the individuals only indicated that it occurred withinn a certain time period (in months), whereas others said they had never had a fall. The data (in months post-admitted) is as follows. Falls occurred in (6, 12], (40, 60], (24, 36], (12, 24], (18, 24], (9, 12], (36, 42], (12, 36]. Time since admittance for individuals who never had a fall: 23, 41, 13, 25, 59, 39, 22, 18, 49, 38. Assume that the lifetime follows exponential distribution with mean $1/\lambda$.
   (a) Write down the likelihood function to obtain the MLE of λ.
   (b) Assuming that the event occurs in the middle of the interval, find the MLE of mean lifetime.

[2+4 = 6]

INDIAN STATISTICAL INSTITUTE

M.Stat. II Year
Mid-Semester Examination : Semester II : 2016-2017
BROWNIAN MOTION AND DIFFUSIONS

Date : 20.02.2017 | **Maximum Score : 40** | Time : 2 Hours

Note : This paper carries questions worth a **total** of **48 marks**. Answer as much as you can. The **maximum** you can score is **40**.

1. Let T denote the set of rationals in $[0,1]$ and $(\Omega, \mathcal{A})$ denote the product space $(\mathbb{R}^T, \mathcal{B}^T)$. Also, let C denote the space of all real continuous functions on $[0,1]$ and $\mathcal{C}$ the σ-field on C generated by the evaluation maps.

   (a) Show that the set $A \subset \Omega$, consisting of all uniformly continuous functions on T, belongs to the σ-field $\mathcal{A}$.

   (b) We know that every $\omega \in A$ extends to an unique element $\overline{\omega} \in C$. Consider the map $\varphi : \Omega \to C$ defined as follows: $\varphi(\omega) = \overline{\omega}$ for $\omega \in A$, while $\varphi(\omega)$ equals the constant function 0 for $\omega \notin A$. Show that $\varphi : (\Omega, \mathcal{A}) \to (C, \mathcal{C})$ is measurable. (6+6)=**[12]**

2. Let $\{B_t, t \in [0,\infty)\}$ be a SBM.

   (a) Show that the process $X_t = \int_0^t B_u\, du, t \in [0,\infty)$, is a gaussian process with continuous paths. Find its mean function and covariance kernel.

   (b) For $c > 0$, let $\tau_c = \inf\{t > 0 : B_t > c\}$. Find a function $f : (0,\infty) \to (0,\infty)$ such that $\tau_c \stackrel{d}{=} f(c)\tau_1$ for each $c > 0$.

   (c) Let $\tau_+ = \inf\{t > 0 : B_t > 0\}$. Show that for every $\epsilon > 0$, $P(\tau_+ \leq \epsilon) \geq 1/2$ and use Blumenthal's $0-1$ law to deduce that $P(\tau_+ = 0) = 1$.

   (d) For $0 \leq a < b$, let $M_a^b = \sup\{B_t : a \leq t \leq b\}$. Show that for $0 \leq a < b \leq c < d$, $P(M_a^b = M_c^d) = 0$. [You may prove it for $b < c$ and then use (c).] (4× 6)= **[24]**

3. Let $\{B_t, t \in [0,\infty)\}$ be a SBM.

   (a) Show that for any $a < 0 < b$, $\tau = \inf\{t > 0 : B_t \notin [a,b]\}$ is an optional time.

   (b) Show that for any optional time $\tau < \infty$, the random variable B_τ is measurable with respect to the σ-field $\mathcal{A}_{\tau+}$. (6+6)=**[12]**

Statistical Methods in Epidemiology & Ecology
Mid-Semestral Examination
M.Stat.- II Year, 2016-2017

Date: 21.02.17

Total Marks - 75 Time: 3 hrs

Attempt all questions:

1. Let us consider the Blumbergs's growth equation

$$\frac{1}{x(t)}\frac{dx(t)}{dt} = rx^{\theta}(t)\left[1 - \left(\frac{x(t)}{k}\right)\right]^{\beta}, \quad (1)$$

where, x(t) be the size variable describing the growth profile of a specific species and (θ, β) are the non-negative growth curve parameters.

   (a) Interpret the above equation and its parameters through the concept of basic and two opposite growth pulses.
   (b) Find the approximate analytical solution of the growth curve assuming size to be bounded by the carrying capacity. Comment on the stability aspect of this approximate growth curve without using the stability theorem.
   (c) Evaluate the point of inflection of the growth curve and compare it with point of inflexion of the logistic growth using the definition of log, lag and stationary phase (you may use some limiting value of parameter(s)).
   (d) Find the condition for which the growth curve (1) is symmetric with respect to its point of inflexion.
   (e) Define Alleé effect and its critical size? Equation (1) can be used (with/without changing the model structure) to model the Allee effect phenomena with positive critical size/null critical size -discuss.
   (f) The relative growth rate (RGR) of the equation(1) can capture both monotonic and non-monotonic profiles while plotted against density- prove or disprove it.

[3+(4+3)+5+4+3+3=25]

2. Let us define the two species Lotka-Volterra model as

$$\begin{aligned} \frac{dx(t)}{dt} &= rx(t) - cx(t)y(t) \\ \frac{dy(t)}{dt} &= bx(t)y(t) - my(t), \end{aligned} \quad (2)$$

where, $x(t)$ and $y(t)$ are the prey and predator abundance at time point t. We are observing the interactions between a specific prey and predator based on the abundance data available for q time points. We have n sets of such data obtained from n territories.

Let us define, X and Y be the $(n \times q)$ data matrices in which the i^{th} row corresponds to the q variate abundance measurements of prey and predator respectively, available at q equispaced time points, for a specific territory. We have a total no of n territories.

Let us assume that i^{th} row $X_i = (X_i(1), ..., X_i(q))' \sim N_q(\theta, \Sigma)$ and $Y_i = (Y_i(1), ..., Y_i(q))' \sim N_q(\phi, \Sigma_1)$, where $X_i(t)$ and $Y_i(t)$ be the abundance of the prey and predator at time point t for the i^{th} territory and we also assume that $E(\bar{R}_{x_i}(t)/Y_i(t))$ and $E(\bar{R}_{y_i}(t)/X_i(t))$ have Lotka-Volterra mean structures, can be constructed from equation (2).

We define two commonly used estimates of "relative growth rate(RGR)" for any time interval (t, t+1) are as follows,

$$(i)\ \bar{R}_x(t) = \frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{X_i(t+1)}{X_i(t)}\right) \quad \text{and} \quad (ii)\ \bar{R}_x(t) = \ln\frac{\bar{X}(t+1)}{\bar{X}(t)} \quad \text{(prey)}$$

$$(i)\ \bar{R}_y(t) = \frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{Y_i(t+1)}{Y_i(t)}\right) \quad \text{and} \quad (ii)\ \bar{R}_y(t) = \ln\frac{\bar{Y}(t+1)}{\bar{Y}(t)} \quad \text{(predator)}$$

Show that each pair $(\bar{R}_x(t), \bar{Y}_t)$ and $(\bar{R}_y(t), \bar{X}_t)$ follows bivariate normal distribution under both the estimates of RGR.

[12]

3. (a) Define Fisher's Relative Growth Rate(RGR) based on the size measurement at two specific time points. Comments on its extension and growth law non-invariant form. Derive the expression for this extended RGR metric for logistic growth law based on the size measurement at three consecutive time points. How this extended metric is affected through reading/measurement error under logistic law.

(b) Let us define, $X(t)$ and $R(t) = \frac{1}{x(t)}\frac{dx(t)}{dt}$ be the size and relative growth rate(RGR) of a species measured at time point t. We assume $(R(1), ...R(q))' \sim N_q(\theta, \Sigma)$, where $E(R(t)) = \theta(t) = f(\phi, t)$, a suitable rate profile. Suppose we are interested in testing the hypothesis of extended Gompertz growth curve model(GGCM), i,e., to test

$$H_0 : \theta(t) = ae^{-bt}t^2 \quad \text{against} \quad H_1 : \text{not} \quad H_0.$$

Using the approximate expression for expectation and variance of the logarithm of the ratio of RGR for two consecutive time points, construct an asymptotic test for the null hypothesis of the extended GGCM. Also, suggest required modifications of the test statistic when errors are non-normal.

[(2+3+4+5)+(10+3)=27]

4. (a) Find the analytical solution of the growth curve governed by the following growth equation

$$\frac{1}{x(t)}\frac{dx(t)}{dt} = be^{-at^c}t^{c-1}, \tag{3}$$

where a,b and c are positive constants. Compare the point of inflections of the RGR for $c = 1$ and 2 respectively, with proper interpretations.

(b) Let us rewrite equation (3) as

$$R_t = be^{-at^c}t^{c-1} + \epsilon_t, \tag{4}$$

where, R_t is the empirical estimate of RGR at time point t and ϵ_t is the error of non linear regression. Show that nonlinear least square estimates exist and are consistent for the model (4).

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

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination 2016-17

M.Stat. - 2nd Year
Robust Statistics

21st February, 2017

Maximum Marks: 60 Time: 3 hours

[Note: Notations are as used in the class. Answer as much as you can. Best of Luck!]

1. What are the major objectives of 'Robust Statistics'? What are the advantages or disadvantages of 'Robust Statistics' over the outlier detection rules only? [2+3]=5

2. Suppose $X_1, \ldots, X_n$ are iid observations from a scale model $\{F_\theta(x) = F(x/\theta) : \theta \in \Theta \subseteq [0, \infty)\}$. An estimate of θ, known as the (unstandardized) α-Winsorized variance, can be defined as

$$T_n = \frac{1}{n}\left[[n\alpha]X^2_{([n\alpha]+1)} + \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X^2_{(i)} + [n\alpha]X^2_{(n-[n\alpha])}\right], \quad 0 \le \alpha \le 0.5,$$

where $X_{(1)} \le \cdots \le X_{(n)}$ denotes the order statistics.

(a) Argue that T_n is an L-Statistics and derive the corresponding statistical functional T_α.

(b) Derive the influence function of T_α. When is it B-robust?

(c) Derive the finite-sample breakdown point of T_n. Hence, or otherwise, derive its asymptotic breakdown point.

(d) Assuming that the required conditions hold, what can you say about the asymptotic distribution of T_n? (Hint: Use part(b)) [3+5+4+3]=15

3. Define the M-estimator and the One-step M-estimator. Prove that they are asymptotically equivalent (state the required conditions clearly). [3+7]=10

4. Consider a location model $\{F_\theta(x) = F(x - \theta) : \theta \in \Theta \subseteq \mathbb{R}\}$. Consider a location M-estimator $T(G)$ with $\psi(x, \theta) = \psi(x - \theta)$.

(a) Derive the max-bias and asymptotic breakdown point of T under gross-error neighborhood for monotone and bounded ψ. Are they the same if we consider the Levy neighborhood?

(b) When is T qualitatively robust? Justify.

(c) What are the most efficient and most B-robust M-estimators for this problem? Justify.

(d) Derive the optimum B-robust M-estimator for this problem. Clearly state all the results you will use. Simplify it for $F = \Phi$, the standard normal distribution.

[(4+4+2)+4+(1+4)+6]=20

5. Suppose $X_1, \ldots, X_n$ are independent observations from the location model of Question 4. The most common estimate of θ is $\widetilde{\theta}$ = the mean of the data, which can be seen as the solution to $\sum_{i=1}^{n}(X_i - \theta) = 0$. That method is famously non-robust. Consider an alternative estimate $\widehat{\theta}$ of θ defined as the solution to the estimating equation

$$\sum_{i=1}^{n} \tan^{-1}(X_i - \theta) = 0.$$

(a) Show that the estimator $\widehat{\theta}$ exists and is unique. Indicate why this method may be expected to be more robust than $\widetilde{\theta}$.

(b) Argue that $\widehat{\theta}$ is an M-estimator and check if it is B-robust?

(c) What can you say about the qualitative robustness or asymptotic breakdown point of $\widehat{\theta}$? (Hint: You may use question 4)

(d) Derive the asymptotic distribution of $\widehat{\theta}$. In particular, in the $N(\theta, 1)$ model, how much is the loss in efficiency by using $\widehat{\theta}$ over $\widetilde{\theta}$?

(e) One may 'work backwards from the estimating equation $\sum_{i=1}^{n}(X_i - \theta) = 0$ to deduce that $\widetilde{\theta}$ is the maximum likelihood estimator for the normal model. Attempt similarly to identify a model such that $\widehat{\theta}$ is its maximum likelihood estimator.

(f) Generalize the setup for a location-scale model with both parameters unknown.

[4+4+(2+2)+(2+2)+4+5]=25

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

Second semester

M.Stat Second Year, 2017

Random walk and electrical networks

Date: 22nd February, 2017 | Maximum Marks: 20

Duration: 2 hours

Anybody caught using unfair means will immediately get 0. Please try to explain every step. Only handwritten class notes are allowed in the exam.

(1) Show that if h is harmonic in a finite set A, then its maximum and minimum are attained at the boundary. What happens when A is infinite? [10]

(2) Let V_N be a box centered at origin with side length N in $\mathbb{Z}^4$. Let τ_{V_N} be the exit time of the simple random walk from V_N. Show that

$$\frac{\mathsf{E}[\tau_{V_N}]}{N^2} \to 1 \quad \text{as } N \to \infty.$$

Explain every step with details. [10]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: (2016–2017)

M. Stat Second Year

Inference for High Dimensional Data

Date: 23/02/2017 Marks: 40 Duration: $2\frac{1}{2}$ hours

Attempt all questions

1. (a) Prove that for large dimension p, most of the mass of the standard p-dimensional normal is concentrated in the tail.

   (b) Provide proper mathematical justification of the above phenomenon in terms of volume of balls in high-dimensional spaces.

[5+5=10]

2. Consider the linear model $\mathbf{Y} = \mathbf{f}^* + \epsilon = \mathbf{X}\beta^* + \epsilon$ in the co-ordinate sparse setting. Assume that the columns of $\mathbf{X}$ are orthogonal. Consider the family $\mathcal{M}$ and the models S_m of the co-ordinate sparse setting and consider the penalty $pen(m) = \lambda|m|$, for some $\lambda > 0$, where $|m|$ denotes the cardinality of m.

   (a) Show that with $\lambda = K\left(1 + \sqrt{2\log(p)}\right)$, where K is a constant, this penalty is approximately equal to the penalty

   $$pen(m) = K\left(\sqrt{d_m} + \sqrt{2\log\left(1/\pi_m\right)}\right)^2,$$

   when p is large. Here d_m denote the dimension of S_m and $\pi_m = \left(1+\frac{1}{p}\right)^{-p} p^{-|m|}$.

   (b) For $\lambda > 0$, we define $\hat{m}_\lambda$ as a minimizer of $\|\mathbf{Y}-\hat{\mathbf{f}}_m\|^2+\sigma^2 pen(m)$, with $pen(m) = \lambda|m|$. Prove that this minimization criterion is equivalent to $\|\mathbf{Y}\|^2 + \sum_{j\in m}\left(\lambda\sigma^2 - \left(\frac{\mathbf{X}_j^T\mathbf{Y}}{\|\mathbf{X}_j\|}\right)^2\right)$, and the minimizer is given by $\hat{m}_\lambda = \left\{j : \left(\mathbf{X}_j^T\mathbf{Y}\right)^2 > \lambda\|\mathbf{X}_j\|^2\sigma^2\right\}$. Here $\mathbf{X}_j$ denotes the j-th column of $\mathbf{X}$.

[3+7=10]

3. In question number 2. now assume that $\mathbf{f}^* = \mathbf{0}$.

   (a) Obtain the oracle model.

   (b) Obtain $E\left[|\hat{m}_\lambda|\right]$. Hence argue whether or not the Akaike Information Criterion (AIC) is appropriate for this context.

(c) Prove that for a standard normal variable Z,

$$P\left(Z^2 > x\right) \overset{x\to\infty}{\sim} \sqrt{\frac{2}{\pi x}} e^{-x/2}.$$

Hence, or otherwise, prove that for $K > 0$,

$$E\left[\left|\hat{m}_{2K\log(p)}\right|\right] \overset{p\to\infty}{\sim} \frac{p^{1-K}}{\sqrt{\pi K \log(p)}}.$$

Does this imply that the penalty $pen(m) = 2|m|\log(p)$ is minimal in this setting? Justify.

(d) The Bayes Information Criterion (BIC) is given by

$$\hat{m}_{BIC} \in \underset{m\in\mathcal{M}}{\arg\min} \left\{ \|\mathbf{Y} - \hat{f}_m\|^2 + \sigma^2 d_m \log(n) \right\},$$

where n is the dimension of $\mathbf{Y}$. Prove that for $p \sim n$,

$$E\left[|\hat{m}_{BIC}|\right] \overset{p\to\infty}{\sim} \sqrt{\frac{2p}{\pi \log(p)}}.$$

So, is the BIC suited for this context? Justify.

[5+5+5+5=20]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2016 – 17

MStat (2nd Year)

Quantitative Finance

Date: 24 February 2017 Maximum Marks: 30 Duration: 2 Hours

1. Define the following: [3 X 3 = 9]
   a) Dominant strategy
   b) Second order continuous parameter stochastic process
   c) Martingale

2. Assuming $V_0 > 0$, the discounted return is

$$R_n^* = [S_n^*(1) - S_n^*(0)]/S_n^*(0) \text{ for } n = 1,\dots,N$$

Show that

(a) $G^* = \sum_{n=1}^{N} H_n S_n^*(0) R_n^*$

(b) $R_n^* = \frac{R_n - R_0}{1+R_0}$ for n = 1,…,N.

(c) Q is a risk neutral probability if and only if $E_Q[R_n^*] = 0$ for n = 1,…,N.

[3+3 + 4 = 10]

3. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \frac{1}{\gamma} w^{\gamma}$ where $\gamma < 1$. Show that the Lagrange Multiplier

$$\lambda = v^{-(1-\gamma)} \{E[(L/B_1)^{-\frac{\gamma}{1-\gamma}}]\}^{(1-\gamma)}$$

the optimal attainable wealth

$$W = \frac{v(L/B_1)^{-1/1-\gamma}}{E[(L/B_1)^{-\gamma/1-\gamma}]}$$

and the optimal objective value is $E[u(W)] = \lambda v/\gamma$.

Compute the relevant expressions and solve for the optimal trading strategy when N = 1, K = 2, r = 1/9,

$S_0 = 5$, $S_1(\omega_1) = 20/3$, $S_1(\omega_2) = 40/9$ and $P(\omega_1) = 3/5$. [4 + 3 + 2 + 4 = 13]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016-17

Course Name	:	**M. Stat. 2nd year**
Subject Name	:	**Weak convergence and empirical processes**
Date	:	24.02.17
Maximum Marks	:	**40**
Duration	:	**2 hours**

Answer **any four** of the following questions. Each question carries 10 marks. Please state clearly **every** result you use. If you use a result not stated in this course, you will have to prove it.

1. Consider a casino where after every game, the rewards are i.i.d. with mean zero and known variance σ^2 which is finite and positive. A gambler wants to play at least n times in that casino without getting bankrupt. Find the initial capital α_n, which the gambler needs, in order to achieve his goal with asymptotic probability 0.95, as $n \to \infty$.

2. Let $\{X_n\}$ be a sequence of real valued random variables such that
$$\liminf_{n\to\infty} \mathrm{E}(X_n^2) < \infty .$$
Show that there exists an integrable random variable X and a subsequence $\{X_{n_k}\}$ of $\{X_n\}$ such that as $k \to \infty$,
$$X_{n_k} \Rightarrow X ,$$
and
$$\lim_{k\to\infty} \mathrm{E}(X_{n_k}) = \mathrm{E}(X) .$$

3. For each $n \in \mathbb{N}$, let Π_n be a tight family of probability measures on a metric space $\mathcal{S}$. Suppose that for each $n \geq 1$, $\{P_{n1}, P_{n2}, P_{n3}, \ldots\} \subset \Pi_n$. Show that there exist $1 \leq k_1 < k_2 < \ldots$ and probability measures $P_{1\infty}, P_{2\infty}, \ldots$ on $\mathcal{S}$ such that for every fixed $n \geq 1$,
$$P_{nk_r} \Rightarrow P_{n\infty} ,$$
as $r \to \infty$.

4. Let $\{x_n\}$ be a sequence in a metric space $\mathcal{S}$. For each $n \in \mathbb{N}$, let P_n be the probability measure which is degenerate at x_n. Show that the family $\{P_n\}$ is tight if and only if every subsequence of $\{x_n\}$ has a further subsequence which converges.

5. Suppose that for every $t \in [0,1]$, $X(t)$ is a random variable such that for all $k \geq 1$ and $t_1, \ldots, t_k \in [0,1]$, $(X(t_1), \ldots, X(t_k))$ follows

P.T.O

k-dimensional multivariate normal with mean zero. Furthermore, assume that

$$\mathrm{Cov}(X(s), X(t)) = \sqrt{s} + \sqrt{t} - \sqrt{|t-s|}, \; s, t \in [0, 1] \,.$$

Show that there exists a $C[0, 1]$-valued random element Y such that

$$P(X(t) = Y(t)) = 1 \text{ for all } t \in [0, 1] \,.$$

INDIAN STATISTICAL INSTITUTE

M.STAT Second Year
2016-17 Semester II

Computational Finance
Midterm Examination

27·02·17

Points for each question is in brackets. Total Points 100.
Students are allowed to bring 2 pages of hand-written notes
Duration: 3 hours

1. Consider the Black Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

   (a) Use a mesh of equal space steps of size δS and equal time steps of size δt, central differences for space derivatives, backward differences of time derivatives, to obtain the explicit finite difference equations

$$V_n^m = a_n V_{n-1}^{m+1} + b_n V_n^{m+1} + c_n V_n^{m+1}$$

   where V_n^m is the finite difference approximation to $V(n\delta S, m\delta t)$ and

$$\begin{aligned} a_n &= \frac{1}{2}(\sigma^2 n^2 - rn)\delta t \\ b_n &= 1 - (\sigma^2 n^2 + r)\delta t \\ c_n &= \frac{1}{2}(\sigma^2 n^2 + rn)\delta t \end{aligned}$$

   (b) Why is this an explicit method? What boundary and initial/final conditions are appropriate?

   (c) Derive the corresponding implicit finite difference equations.

2. (a) Assume that you have a random number generator from Unif(0,1). Briefly indicate an algorithm for generating a random variable from distribution with pdf $\frac{1}{16}\sqrt{x} + 1/6, 0 < x < 4$ using the following procedure:
   - Generate X from Uniform (0,4)
   - Generate Y from a distribution with pdf $\frac{3}{16}\sqrt{x}, 0 < x < 4$.
   - Combine X and Y suitably. Calculate actual realizations (one each) to demonstrate your method using the following values for the uniform random numbers: 0.6443 0.2077 0.3111

   (b) Suppose you want to evaluate the integral of the function $f(x) = e^x$ over $[0, 1]$ using Monte Carlo simulation from Unif(0,1). Use the control variable method with $g(x) = x$ and show that the variance is reduced by a factor of 60. Is there much additional improvement if you use a general quadratic function of x?

3. (a) Use the Ito formula to show that, for any integer $k \geq 2$,

$$\mathrm{E}W(t)^k = \frac{1}{2}k(k-1)\int_0^t \mathrm{E}W(s)^{k-2}ds,$$

   and use this to evaluate the sixth moment of the standard normal distribution.

P.T.O

(b) A Perpetual Option: Assume that the share prices of Stock follows a geometric Brownian motion with zero drift and volatility 1. Consider an option with no date of expiration that pays the owner $\exp \beta\tau_\alpha$ at the first time τ_α that the share price of Stock reaches α (if ever). Here β and α are positive real numbers, and $S_0 < \alpha$. Calculate the arbitrage price at time 0 of this option.

[Hint: Use $P(\tau_\alpha \leq t)=2P(S_t > \alpha)$, by the reflection principle.]

4. (a) Show that the Gaussian quadrature formula with n nodes is of order $2n-1$. That is, if we choose as nodes the n roots of a polynomial of order n, within a family of orthogonal polynomials, then for any polynomial f of order $2n-1$, we have $\int_a^b f(x)w(x)dx = \sum_{i=1}^n w_i f(x_i)$.

   (b) When approximating an integral by the trapezoid method, show that the error is bounded by $\frac{M}{12}(b-a)^3$ where $M = \sup f''(x)$.

Indian Statistical Institute
Mid semestral Examination: (2016–2017)
M. Stat.–II yr. (MSP)
Special topic: Random Matrices
Total marks=80

Date: 27.02.17 Time: 2 hours

Answer all questions. Questions carry equal marks.

1. Suppose $F_n(x, \omega), x \in R, \omega \in \Omega$ is a sequence of probability distribution functions with the natural measurability conditions. Suppose that $\int_{-\infty}^{\infty} x^k dF_n(x, \omega)$ converges *almost surely* to a non-random μ_k for each $k = 1, 2 \ldots$.

(i) Show that $\{\mu_k\}$ is a moment sequence.

(ii) Suppose $\{\mu_k\}$ determines a unique probability distribution F. Show that $F_n(x, \omega)$ converges to F weakly *almost surely.*

(iii) Show that the above statement remains true if "almost surely" is replaced by "in probability" in two places above.

2. State (no proof required) an asymptotic normality theorem for the ESD of the $n \times n$ Symmetric Circulant with i.i.d. input. Consider the $n \times n$ palindromic symmetric Toeplitz matrix. Find out its LSD using the above result.

3. Let m_n be the number of non-crossing pair partitions of $\{1, 2, \ldots, 2n\}$.

(i) Show that $m_n = \frac{(2n)!}{n!(n+1)!}$.

(ii) Find $\lim \frac{n}{\log m_n}$ as $n \to \infty$.

4. Suppose X is a $p \times p^2$ matrix with all i.i.d. entries all whose moments are finite. Consider the matrix $A_p = \sqrt{p}(XX' - I_p)$ where I_p is the identity matrix of order p. Show that as $p \to \infty$, the fourth expected moment of its ESD converges to 2 and the third expected moment of its ESD converges to zero.

$\frac{XX'}{p^2}$

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : Semester II (2016-17)

M. Stat. II Year

Clinical Trials

Date: 28.2.2017 Maximum marks: 50 Time: 120 minutes

(Total mark is 53. Calculator can be used.)

1. (a) Describe the different phases of a drug trial highlighting the objectives and some characteristics.

   (b) In the context of the different principles of scientific methods in a Randomized Clinical Trial (RCT), briefly describe three aspects of 'Purpose of the Trial'.

   (c) Mention two problems with a non-randomized trial.

   (d) Give one reason why Statistics in important in the conduct of an RCT.

[8+3+2+1=14]

2. (a) Describe 'selection bias' in the context of an RCT.

   (b) Consider an RCT with two groups T and C. The patients are classified as Serious (S) and Not Serious (N) with probability of Serious patients being α. The outcome variable X has mean μ_S for the Serious patients and μ_N for the Not Serious patients, with $\mu_N > \mu_S$, in the patient population before the trial. Given that a Serious patient is allocated to the T group, there is a probability p that the patient will not join the study. Similarly, a Not Serious patient allocated to the T group has a probability p' $(> p)$ of not joining the study.

   Assume additive effects θ_T or θ_C on the mean of X when the patient is under the trial in the T or C group, respectively.

   Work out if the difference of sample means of X in the two groups based on observations from the patients in the trial, as a measure of treatment difference, suffers from any bias or not.

[2+15=17]

3. An RCT is to be conducted on the treatment of asthma patients to investigate the efficacy of a new steroid inhaler against an existing preparation. The outcome variable is Forced Expiratory Volume in 1 Sec (FEV1) in litres, a measure of lung function. The aim is to be able to detect a change of 0.25 litres in mean FEV1 in the new inhalation group with at least 90% power. It is known that the distribution of FEV1 in the population is normal with standard deviation 0.5 litres.

   Find the minimum number of patients to be enrolled in each of the two groups with 5% nominal level. Consider the approximate power function after ignoring the left tail probability. It is given that the 90th and 97.5th percentiles for the standard normal distribution are 1.28 and 1.96, respectively.

P.T.O.

If it is decided that the number of patients in the new inhalation group should be m (known) times that in the existing preparation group, find an expression for the approximate power function with $2n$ number of patients in the trial. How does it behave with m?

[8+(2+1)=11]

4. (a) Describe Efron's Biased Coin Design. For this design, what is the probability of exact balance in the long run?

   (b) Consider a random allocation scheme in which the maximum allowable absolute imbalance is 4. Whenever the absolute imbalance is less than 4, next patient is allocated to one of the two groups by simple randomization. Describe this scheme formally.

[(4+1)+6=11]

INDIAN STATISTICAL INSTIUTE

MID-SEMESTRAL EXAMINATION

M-Stat. – II

Full Marks: 40 | Duration: 3 hours | Date: 28.02.2017

Note: Answer **ALL** questions

1. (a) Define a dominated statistical experiment. [2]

   (b) Let Ω be uncountable and $\{\omega\} \in \mathcal{A} \ \forall\, \omega$. Then show that $\{P_\theta : \theta \in T\}$ is dominated by a σ – finite measure on $(\Omega, \mathcal{A})$ if f $\exists$ a countable set A such that $P_\theta\ (A) = 1\ \forall_\theta$; here each P_θ is a discrete probability on $(\Omega, \mathcal{A})$. [3]

   (c) Give an example of a "discrete statistical experiment" which is undominated. Justify your answer. [1 + 1]

   (d) Let $P_\theta\ (\{\theta + 1\}) = (\{\theta - 1\}) = \frac{1}{2}\ \forall\, real\ \theta$; Here $\Omega = \mathbb{R}$, $\mathcal{A} =$ Borel σ – field on Ω. Find a minimal sufficient statistic; justify your answer. [The sample is random and of size n.] [2 + 1]

2. (a) Define a splitting set for $(\Omega, \mathcal{A}, \mathrm{M})$. [1]

   (b) Define connectedness of a statistical experiment. [1]

   (c) Show that the connectedness implies the nonexistence of a splitting set. Is the converse true? Justify your answer. [2 + 2]

   (d) Give an example of a connected experiment. [1]

   (e) Find cov (T_1, T_2) where $T_1 = \frac{X_1}{U}, T_2 = (3\,X_3 - 2\,min_{1 \le i \le 7}\, X_i)/\,U$, $U := max\,(X_1, \ldots, X_7)$ and $X_1, \ldots\ldots, X_7$ are iid $U\,(0\,;1)$. [3]

   [You may use, without proving, that
   $$E\ (U^\gamma) = \frac{7}{7+\gamma}, 0 < \gamma < \infty]$$
   [20]

3. Presentation of papers

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination

M. Stat 2nd year

Subject: Theory of Games and Statistical Decisions

Date: 1st March, 2017 **Full Marks**: 40 Duration: 1hour 30 minutes

Attempt all questions

1. a) Describe different types of Statistical Games related to set of states of Nature, Action space or, Sample space and randomisation or, non-randomisation.

   b) Discuss regarding solution of the types of games in 1a) – existence of value and optimal strategies, when set of states of Nature, Action space and Sample space are finite.

   [10 + 4]

2. a) For mixed extension of Matrix games, assuming existence of $Min_Y Max_X XAY^T$ and $Max_X Min_Y XAY^T$, show that for all values of matrix A, value of the game exists.

   b) Give example of a Statistical Game with finite states of Nature and infinite Action space for which value does not exist, even with randomisation.

   [14 + 2]

3. a) For identity matrix $I_{n\,x\,n}$ find value of the mixed extension of matrix game I. Also find all related optimal strategies.

   b) For all m, n (greater than 1) give example of a matrix $A_{m\,x\,n}$ such that the matrix game A does not possess an equilibrium situation.

   c) Let $A_{n\,x\,n}$ be a skew-symmetric matrix i.e. $A^T = -A$.
   Find value of mixed extension of the matrix game A.

   [(3 + 2) + 2 + 3]

---------- ***** ---------

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination 2016-17

MStatII Directional Data Analysis

16 March 2017 Maximum Marks: 20 Duration 2 hrs

Note: Attempt all questions

1. (8) (a) State and prove the result defining the relationship between the characteristic functions of a linear r.v. and its corresponding circular r.v.
   (b) Let the pdf of a circular r.v. Θ be
   $$f(\Theta) = K. [\,1 + A. \cos\Theta + B. \cos\Theta^2].$$
   Obtain K. Using the Fourier series representation of $f(\Theta)$ or otherwise, derive consistent estimators of A and B, when a random sample of n observations is obtained from $f(\Theta)$.

2. (6) (a) Using Möbius transformation, obtain the density of the linear r.v. corresponding to that of the circular r.v. having the pdf, $f(\Theta) = K. [\,1 + A. \cos\Theta\,]^n$, where K is the normalizing constant and $A \in (-1, 1)$ is a parameter.
   (b) Using the result in (a) above or otherwise, show how you can generate (linear) t random variates by generating only circular uniform random variates.

3. (6) (a) Establish that Watson's U^2 Goodness-of-Fit test is invariant w.r.t choice of origin for the data.
   (b) Show explicitly how you will implement the test in (a) above for $f(\Theta)$ given in 1(b).

****

14.3.2017.

INDIAN STATISTICAL INSTITUTE

Semestral Examination
Second semester

M.Stat Second Year, 2017

Random walk and electrical networks

Date: 24.4.17 Maximum Marks: 60

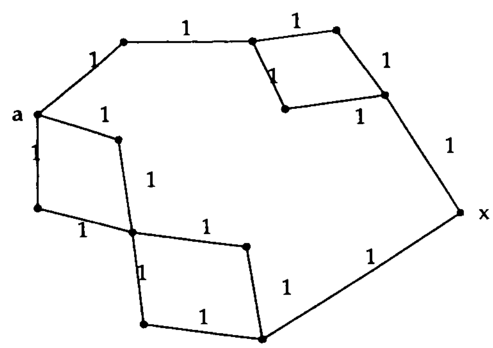
Duration: 3 hours

Anybody caught using unfair means will immediately get 0. Please try to explain every step. Only handwritten class notes are allowed in the exam.

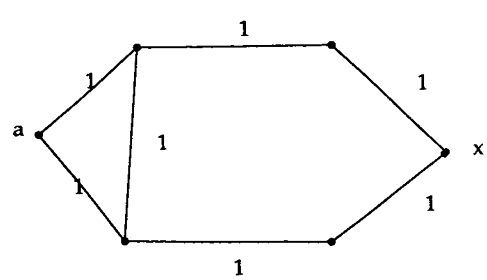
(1) Suppose that each edge in the following network has equal conductance 1. Recall that for a set A, hitting time $\tilde{H}_A = \inf\{n \geq 1, X_n \in A\}$ and entrance time $H_A = \inf\{n \geq 0 : X_n \in A\}$.

(a) Compute $\mathbf{P}_a[\tilde{H}_a > H_x]$ and $C_{\text{eff}}(a, x)$ for the following graphs.

(i) Network (a (i))

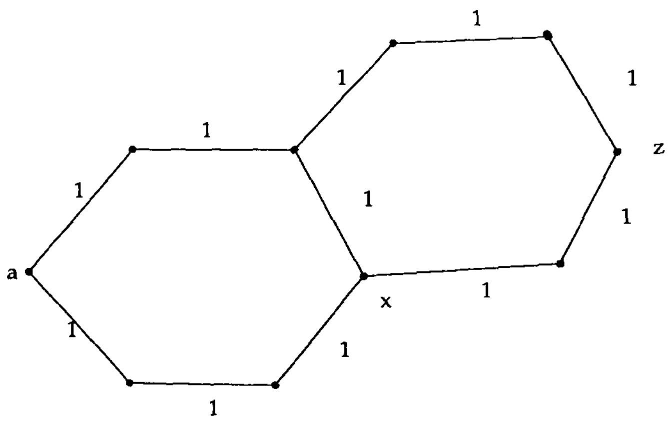


(ii) Network (a(ii))



[7+6=13 points]

(b) Compute $\mathbf{P}_x[H_a < H_z]$ for the following network.



Give explanation for each network reduction you use. [7 points]

(2) Let $\mathbb{T}$ be a rooted binary tree with root o. Let $\mathbb{T}_n$ be the set of 2^n vertices at distance n from the root o.

- Show that $\mathbb{T}$ with natural weights is transient.
- Suppose we allow weights on $\mathbb{T}$ to be given by

$$\mu_{x_n,x_{n+1}} = \frac{1}{r_n}, \qquad \text{for } x_n \in \mathbb{T}_n,\ x_{n+1} \in \mathbb{T}_{n+1}, x_n \sim x_{n+1},$$

where r_n is a positive sequence of numbers. Give with proper reasons, a condition on $\{r_n\}$ such that $\mathbb{T}$ is transient.

[10+5=15 points]

(3) Consider the following flow on $\mathbf{Z}^3$ as follows

$$I_{x,y} = -I_{y,x} = \begin{cases} \frac{x_1+1}{(S+1)(S+2)(S+3)} & \text{when } x = (x_1, x_2, x_3), y = (x_1+1, x_2, x_3), x_i \geq 0, \\ \frac{x_2+1}{(S+1)(S+2)(S+3)} & \text{when } x = (x_1, x_2, x_3), y = (x_1, x_2+1, x_3) \\ \frac{x_3+1}{(S+1)(S+2)(S+3)} & \text{when } x = (x_1, x_2, x_3), y = (x_1, x_2, x_3+1), \end{cases}$$

where $S = x_1 + x_2 + x_3$. Show that

(a) $\text{div} I(x) = 0$ when $x \in \mathbf{Z}^3 \setminus \{0\}$.

(b) Show that the flux I out of origin, $F(I, 0) = \frac{1}{2}$.

(c) Show that I has finite energy.

(d) Hence using the above information show that $\mathbf{Z}^3$ is transient.

(e) Show using above that $\mathbf{Z}^d$ is transient for $d \geq 4$.

[4+3+3+3+2=15 points]

(4) Let U be a proper (possibly infinite) subset of $\mathbf{Z}^2$. Let $T_U = \inf\{n \geq 0 : X_n \notin U\}$. Show that for fixed $x \in \mathbf{Z}^2$,

$$\mathsf{E}_x \left[\tilde{H}_x \mathbf{1}_{(\tilde{H}_x < T_U)} \right] < \infty.$$

[10 points]

INDIAN STATISTICAL INSTITUTE
Semester Examination 2016-17

M.Stat. - 2nd Year
Robust Statistics

24·4·17

Maximum Marks: 50 Time: 3 hours

[Note: Notations are as used in the class. State the results that you are using. Answer as much as you can. Best of Luck!]

1. Suppose $X_1, \ldots, X_n$ are iid observations from a scale model $\{F_\theta(x) = F(x/\theta) : \theta \in \Theta \subseteq [0, \infty)\}$. Consider a scale M-estimator $T(G)$ defined on the distribution G with the associated function $\psi(x, \theta) = \psi(x/\theta)$.

   (a) Derive the most B-robust estimator for this model.

   (b) Define the Change-in-variance function (CVF) and V-robustness of $T(G)$ with all the required assumptions.

   (c) What are the optimum V-robust estimators under this model? Compute their CVFs at the standard normal distribution $F = \Phi$.

   (d) What is the optimum V-robust redescending M-estimator for this problem. Draw its CVF at the standard normal distribution $F = \Phi$.

[3+3+3+(2+2)]=13

2. Consider a location-scale model $\left\{F_\theta(x) = F\left(\frac{x-\mu}{\sigma}\right) : \theta = (\mu, \sigma)^T \in \mathbb{R} \times \mathbb{R}^+\right\}$ with symmetric F. Consider an M-estimator $T(G) = (T_\mu(G), T_\sigma(G))^T$ for $\theta = (\mu, \sigma)^T$, defined based on the function $\psi(x, \theta)$.

   (a) What should be the form of ψ so that $T(G)$ is equivariant. Derive the asymptotic breakdown point of $T_\mu(G)$ under suitable conditions on ψ, assuming σ to be a nuisance parameter.

   (b) What is Huber's "proposal 2" (T^{P2}) for this problem? Derive the influence function and $\gamma_u^\star$ for the corresponding functional $T^{P2} = (T_\mu^{P2}, T_\sigma^{P2})^T$ at the standard normal distribution $F = \Phi$. Also derive the asymptotic breakdown point of its location component T_μ^{P2} at $F = \Phi$, and compare it with that for the known σ case (Huber's "proposal 1" for μ). Also, calculate the loss/gain in efficiency for the location estimator when σ is unknown over when σ is known.

[(1+4)+(1+3+3+3)]=15

3. Suppose $X_1, \ldots, X_n$ are iid observations from a population having true distribution G and true density g. We model it by a parametric family of distributions $\{F_\theta : \theta \in \Theta \subseteq \mathbb{R}^p\}$ with f_θ being the corresponding model density. Consider the problem of estimating θ based on the minimum distance approach with the density-based divergence measure

$$d_{A,B}(g, f_\theta) = \frac{1}{A}\int f_\theta^{1+\alpha} - \frac{1+\alpha}{AB}\int f_\theta^B g^A + \frac{1}{B}\int g^{1+\alpha},$$

where A and B are real numbers such that $A+B = 1+\alpha$ and $\alpha \in [0, 1]$. Note that, $d_{A,B}(g, f_\theta)$ is a proper statistical divergence for all $A, B \neq 0$.

(a) Complete this divergence family by defining $d_{A=0,B}(g, f_\theta)$ and $d_{A,B=0}(g, f_\theta)$ through continuous limits and show that they are adjoint to each-other. What are the self-adjoint members of this family?

(b) Show that this divergence family also contains the power divergence family and the density power divergence family as special cases.

(c) Derive the estimating equation for the corresponding minimum divergence functional $T_{A,B}(G)$ and verify when it becomes an M-estimator functional. How will you obtain the corresponding minimum divergence estimate of θ based on the observed data?

(d) Derive the influence function of $T_{A,B}(G)$ at the model $G = F_\theta$. When is it B-robust?

(e) Derive the asymptotic distribution of the corresponding MDE at the model $G = F_\theta$. When is it fully efficient?

[(2+1)+2+3+2+2]=12

4. Consider the linear regression set-up $y_i = x_i^T\theta + e_i$ for $i = 1, \ldots, n$, where e_is are iid standard normal random variables.

(a) Consider the Least Absolute deviation (LAD) estimator of θ and assume $p = 2$, $x_i = (1, z_i)^T$. Suppose z_1 is an outlier which is so far away that all the remaining z_i lies on the other side of their average value $\bar{z}$. Show that the LAD regression line will go through the leverage point (z_i, y_i) and hence is not robust with respect to the design-space outliers.

(b) Define the Huber's M-estimator for θ and verify if it is B-robust with respect to both y-space and x-space outliers. Derive its breakdown point, when (i) x_is are random and (ii) x_is are fixed as $x_i = i$ $(p = 1)$ for each $i = 1, \ldots, n$.

(c) Now let us assume x_is are iid with distribution $K(x)$ and are independent of e_is. Define the general M-estimators for θ and derive the most B_s-robust estimator among them.

(d) Find out the optimal B_u-robust estimators within the Mallows subclass of general M-estimators.

(e) What is the optimum V_u-robust estimator within the class of general M-estimators (you don't need to prove it)? Is it of Mallows type?

[3+(1+2+3+3)+(1+3)+4+2]=22

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, Second Semester, 2016-17

Semestral Examination

Statistical Computing II

Date: 26-04-17 Full Marks: 100 Time: 4 Hours

Answer as many as you can. The maximum you can score is 100.

Answers should be brief and to the point. If you use any standard result, mention it clearly.

1. Consider a moving average process $X_t = \mu + \varepsilon_t + \frac{1}{2}\varepsilon_{t-1}$ for $t = 1, 2, \ldots$, where the ε_t's are independent and identically distributed with the mean 0 and the variance σ^2.

   (a) Check whether $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ is a consistent estimator of μ. [6]

   (b) Find the asymptotic distribution of $Y_n = \sqrt{n}(\bar{X}_n - \mu)$. [4]

   (c) Now, consider a bootstrap sample $X_1^*, \ldots, X_n^*$ drawn from $\{X_1, \ldots, X_n\}$. Define $\bar{X}_n^* = \sum_{i=1}^{n} X_i^*/n$ and $Y_n^* = \sqrt{n}(\bar{X}_n^* - \bar{X}_n)$. Check whether the asymptotic distribution of Y_n^* matches with that of Y_n. [6]

2. (a) Consider a generalized linear model with binary response variable and probit link function. If there exist $\beta_0, \beta_1, \ldots, \beta_p$ such that $(y_i - 0.5)(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}) \geq 0$ for all $i = 1, 2, \ldots, n$, show that the maximum likelihood estimate of $(\beta_0, \beta_1, \ldots, \beta_p)$ will not exist. [6]

   (b) Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be independent multivariate observations coming from two elliptic distributions F_1 and F_2. For $i = 1, 2, \ldots, n$, define $Z_{1i} = HD(\mathbf{X}_i, F_1)$ and $Z_{2i} = HD(\mathbf{X}_i, F_2)$ for $i = 1, 2, \ldots, n$, where HD denotes for half-space depth. Also define $Y_i = 1$ if the i-th observation comes from F_1, and 0 otherwise. Show that the relation between Y and (Z_1, Z_2) can be expressed using an appropriate generalized additive model. [6]

3. (a) Define the Nadaraya Watson estimate of a regression function. [2]

   (b) Let $\hat{\psi}$ be the Nadaraya Watson estimate of a regression function $\psi : R \to R$ based on a kernel K and bandwidth h. Assume that ψ is Lipschitz continuous, K has bounded support and $h = O(n^{-\delta})$ for some $\delta \in (0, 1)$, where n is the sample size. Show that for any fixed x, $\hat{\psi}(x)$ converges to $\psi(x)$ in probability. [8]

   (c) Describe how Nadaraya Watson estimate of a regression function can be used to construct a classification rule when there are observations from two competing classes. [4]

4. (a) Describe how bagging (bootstrap aggregating) method can be used to choose the optimal bandwidth in kernel discriminant analysis. [4]

   (b) Can a random forest classifier be considered as a bagged version of classification tree? Justify your answer. [4]

   (c) Describe the discrete Adaboost algorithm for a two-class classification problem. Show that it can be viewed as an algorithm for fitting an additive logistic regression model. Describe how you will modify this algorithm when there are more than two classes. [4+8+4]

P.T.O

5. (a) Briefly describe the patient rule induction method (PRIM) for bump hunting. How does it differ from the beam search algorithm with beam width equal to 1? [4+2]

   (b) Write down different steps of genetic algorithm for finding the minimum covariance determinant (MCD) estimate of location from a data cloud $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{100}\}$ consisting of 100 observations. [8]

   (c) Can simulated annealing be used for finding the MCD estimate? If you think it can be used, briefly describe how you will use that algorithm in this problem. If you think it cannot be used, give reasons for your answer. [4]

6. (a) If $X_1, \ldots, X_d$ are independent and identically distributed $N(0, 1)$ variables, $Y = \sum_{i=1}^{n} X_i^2$ follows a chi-square distribution with d degrees of freedom. Using this result, compute the volume of a d dimensional unit sphere. [8]

   (b) Let $S_d = \{\mathbf{x} \in R^d : \|\mathbf{x}\| \leq 1\}$ be the d-dimensional unit sphere and C_d be the largest hypercube inscribed in it. If $\mathbf{X}$ follows the uniform distribution of S_d, show that the probability $P(\mathbf{X} \in C_d)$ converges to 0 as d tends to infinity. [6]

7. Consider a classification problem involving two d-dimensional normal distributions $N_d(\mathbf{0}_d, \mathbf{I}_d)$ and $N_d(\boldsymbol{\mu}_d, \sigma^2\mathbf{I}_d)$, where $\mathbf{0}_d = (0, 0, \ldots, 0)'$, $\boldsymbol{\mu}_d = (\mu, \mu, \ldots, \mu)'$ and $\mathbf{I}_d$ is the $d \times d$ identity matrix. Suppose that there are 40 observations from each of the two classes.

   (a) Find the training sample error rate of the best linear classifier when $\mu = 1$, $\sigma^2 = 4$ and $d = 100$. Justify your answer. [4]

   (b) Derive the asymptotic error rate (as the dimension d diverges to infinity) of the nearest neighbor classifier based on Euclidean distance when (i) $\mu = 1$ and $\sigma = 2$, (ii) $\mu = 2$ and $\sigma = 1$. [4+4]

Statistical Methods in Epidemiology & Ecology

Semestral Examination

M.Stat.- II Year. 2016-2017

Total Marks - 100

Time: 3 hrs 30 min.

Attempt all questions:

1. Let us define. $X(t)$ and $R(t) = \frac{1}{X(t)}\frac{dX(t)}{dt}$ be the size and relative growth rate(RGR) of a species measured at time point t. We assume. $(X(1),....X(q))' \sim N_q(\theta, \Sigma)$, where $E(X(t)) = \theta(t) = f(\varphi, t)$, a suitable growth profile. Here the vector $(X(1),....X(q))'$ represents q - longitudinal equispaced observations on the same individual. Suppose we are interested in testing the hypothesis of the exponential quadratic growth curve model (EQGCM), i.e., to test

$$H_0 : \theta(t) = e^{-b_0+b_1t+b_2t^2} \quad \text{against} \quad H_1 : \text{not } H_0.$$

   (a) Using the approximate expression for expectation and variance of the logarithm of the ratio of the size variables for two consecutive time points. construct an asymptotic test for the null hypothesis of the EQGCM based on the data $X_{n\times q}$. Note that, in the data matrix $X_{n\times q}$ any row corresponds to a q - variate vector of size measurements available at q time points on one of the n individuals. The rows represent independent measurements on n individuals.

   (b) Also, suggest required modifications of the test statistic when the time points are not equispaced.

   (c) Suggest two commonly used estimates of RGR based on the data matrix X and using any one of these two estimates show that the RGR vector asymptotically follows multivariate normal distribution.

[10 + 4 + (2+4) = 20]

2. (a) Show that the bias for Fisher's RGR is always negative under first order of approximation (where order is defined by the power of the "difference of relative errors" over the time points).

   (b) Also obtain the expression for the MSE under both first and second order of approximation.

   (c) Discuss Lyapunov and asymptotic stability. State and proof the stability theorem for single species growth dynamics.

[3 + 5 + (2+2+5) = 17]

3. (a) Define quasi-equilibrium probabilities in the context of a general birth-death process.

(b) Let us consider the following growth equation for describing the cooperation movement in single species dynamics :

$$\frac{1}{X(t)}\frac{dX(t)}{dt} = rX^{\gamma}(t)\left[1 - \left(\frac{X(t)}{k}\right)^{\theta}\right]. \tag{1}$$

where, $X(t)$ be the size variable describing the growth profile of a specific species and (r, γ, θ, K) are the non-negative growth curve parameters.

Derive the quasi-equilibrium probabilities (write necessary steps only) and determine the expressions for the approximate mean and variance, while a random perturbation term is incorporated in the above model.

[2 + 13 = 15]

4. (a) Find the analytical solution of the growth curve governed by the following RGR growth equation

$$\frac{1}{X(t)}\frac{dX(t)}{dt} = be^{-at}t^{c}, \tag{2}$$

where a, b and c are positive constants. Compare the point of inflections of the RGR curves for $c = 1$ and 2 respectively, with proper interpretations.

(b) Let us rewrite equation (2) as

$$R_t = be^{-at}t^{c} + \epsilon_t, \tag{3}$$

where, R_t is the empirical estimate of RGR at time point t and ϵ_t is the error of non linear regression. Show that the nonlinear least square estimates derived from (3) are consistent and asymptotically normal.

(c) Define profile likelihood. Illustrate its application in estimating the parameters of the growth curve model in ecology.

[(4+4) + (4+8) + (4+4) = 28]

5. (a) Write down a simple deterministic SIR epidemic model with proper assumptions. Reparameterize the model by introducing basic reproduction ratio and the ratio of the average life length to the average duration of infection.

(b) Discuss the case that corresponds to the (i) absence of infection and an (ii) endemic infection under this reparameterization.

(c) Construct the transition probabilities for the stochastic analogue of the deterministic model as proposed in 5(a).

(d) Find the distribution of the time to extinction of the disease.

[(3+3) + 3 + 3 + 8 = 20]

INDIAN STATISTICAL INSTITUTE

M.Stat. II Year
End-Semester Examination : Semester II : 2016-2017
BROWNIAN MOTION AND DIFFUSIONS

Date : 2 ..04.2017 **Maximum Score : 60** Time : 3 Hours

Note : This paper carries questions worth a **total** of **68 marks**. Answer as much as you can. The **maximum** you can score is **60**.

1. Let $\{B_t, t \in [0, \infty)\}$ be a SBM.

   (a) Denoting $M_t = \sup\{B_s : 0 \leq s \leq t\}$, show that $M_t \stackrel{d}{=} |B_t|$.

   (b) Using (a) or otherwise, show that, for any $a > 0$, $T_a = \inf\{s \geq 0 : B_s \geq a\}$ is an absolutely continuous random variable and find its density.

   (c) Consider the $[0, \infty)$-valued process $\{T_a, \ a \geq 0\}$. Show that it has non-decreasing, right-continuous paths almost surely and has independent, stationary increments.

   (d) Using (c) or otherwise, show that the process $\{T_a, \ a \geq 0\}$ has the Markov property and describe its transition density. $(5 \times 4) = $ **[20]**

2. With usual notations, let $(\Omega, \mathcal{A}, \mathcal{A}_t, \{X_t, t \geq 0\}, \{P_x : x \in S\})$ denote a Markov process.

   (a) Define the Resolvent operators for the Markov process and derive the Resolvent equation.

   (b) State what is meant by the Feller property and show that if the Markov process satisfies the Feller property, then it has the strong Markov property.

   (c) Let A denote the generator of the Markov process. Show that if the process has the strong Markov property, then for every $u \in \text{dom}(A)$ and every bounded optional time τ, $\mathbf{E}_x \left[\int_0^\tau Au(X_s)ds\right] = \mathbf{E}_x[u(X_\tau)] - u(x)$, for all $x \in S$.

   (d) Show that for every $u \in \text{dom}(A)$, the process $\{M_t = u(X_t) - \int_0^t Au(X_s)ds, \ t \geq 0\}$ is a martingale under P_x, for all $x \in S$. $(5 \times 4) = $ **[20]**

3. Let $(\Omega, \mathcal{A}, \{\mathcal{A}_t, \ t \geq 0\}, P)$ be a filtered complete probability space with $\mathcal{A}_0$ containing all P-null sets and let $\{B_t, t \in [0, \infty)\}$ be a SBM with respect to $\{\mathcal{A}_t, \ t \geq 0\}$.

   (a) Let $f : [0, \infty) \times \Omega \to \mathbb{R}$ be a measurable function, such that, for all t, the function $f(t, \cdot)$ is $\mathcal{A}_t$-measurable and $\int_0^t |f(s, \omega)| ds < \infty$. Show that for every t, the map $\omega \mapsto \int_0^t f(s, \omega) ds$ is $\mathcal{A}_t$-measurable.

   (b) Let $f, g \in \mathcal{L}_2$ (assume usual notation) and let $M_t = \int_0^t f_s \, dB_s, N_t = \int_0^t g_s \, dB_s, \ t \geq 0$.

   (i) Show that $\int_0^t f_s g_s \, ds, \ t \geq 0$ is an $(\mathcal{A}_t)$-adapted process.

   (ii) Starting from the definition of stochastic integrals, show that $\{M_t N_t - \int_0^t f_s g_s \, ds, \ t \geq 0\}$ is a martingale. $(5 + (2 + 5)) = $ **[12]**

4. Assume the set-up of the previous problem. For $n \geq 0$, let $\{Z_n(t), t \geq 0\}$ be defined recursively as follows: $Z_0(t) \equiv 1$; and for each $n \geq 1$, $Z_n(t) = \int_0^t Z_{n-1}(s) dB(s), t \geq 0$.

   (a) Show that, for each $n \geq 0$, $E[Z_n^2(t)] \leq t^n/(n!)$ for all t and hence conclude that, for each $n \geq 0$, $\{Z_n(t)\}$ is a continuous square-integrable martingale.

   (b) Show that, with probability 1, the series $\sum_{n \geq 0} Z_n(t)$ converges uniformly on every compact interval and hence $X(t) = \sum_{n \geq 0} Z_n(t)$ defines a continuous adapted process.

   (c) Show that $X \in \mathcal{L}_2$ and satisfies the SDE: $dX(t) = X(t)dB(t), \ X(0) = 1$.

   (d) Conclude that $\sum_{n \geq 0} Z_n(t) = \exp[B(t) - (t/2)], \ t \geq 0$. $(5 + 5 + 3 + 3) = $ **[16]**

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2016 – 17

MStat (2^{nd} Year)

Quantitative Finance

Date: 28 April 2017 Maximum Marks: 100 Duration: 3 Hours

1. Let

$$A_{(K+1)X(K+2N)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ \Delta S_1^*(\omega_1) & -\Delta S_1^*(\omega_1) & \Delta S_2^*(\omega_1) & \cdots & -\Delta S_N^*(\omega_1) & -1 & 0 & \cdots & 0 \\ \Delta S_1^*(\omega_2) & -\Delta S_1^*(\omega_2) & \Delta S_2^*(\omega_2) & \cdots & -\Delta S_N^*(\omega_2) & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta S_1^*(\omega_K) & -\Delta S_1^*(\omega_K) & \Delta S_2^*(\omega_K) & \cdots & -\Delta S_N^*(\omega_K) & 0 & 0 & \ldots & -1 \end{bmatrix}$$

and $b_{(K+1)} = (1, 0, \ldots, 0)'$. Show that

$$Ax = b, \ \ x \geq 0, x \in \mathbb{R}^{K+2N}$$

has a solution if and only if there exists an arbitrage opportunity in the securities market with N securities S_i ($i = 1,\ldots,N$) and K states of nature ω_j ($j = 1,\ldots, K$). S_i^*'s are discounted (by the bank process) values. [20]

2. Prove the Put – Call parity of European option for the multi-period market. Is the same relation true for American options? – Prove or refute logically. [6 + 6 = 12]

3. Define the following option contracts:

   (i) Lookback

   (ii) Barrier

   (iii) Chooser

   For each of them, state the payoff function carefully, explaining all notation. [3 X 6 = 18]

P.T.O

4. Prove directly from the definition of Ito integrals that

$$\int_0^t s\,dW_s = tW_t - \int_0^t W_s\,ds$$

[15]

5. (a) Use Ito's formula to write the following stochastic processes X_t in the standard form

$$dX_t = u(t,\omega)dt + v(t,\omega)dW_t$$

for suitable choices of $u \in R, v \in R$

(i) $X_t = W_t^2$

(ii) $X_t = 2 + t + e^{W_t}$

(b) Check whether $X_t = t^2 W_t - 2\int_0^t sW_s ds$ is a Martingale.

[(3 + 4) + 8 = 15]

6. (a) Define Duration, modified Duration and convexity of a Bond.

(b) Derive the formula $\Delta P = -D_M P \Delta\lambda$ where the symbols have their usual meanings. How do you think the formula should be modified if you want to incorporate convexity?

(c) A 10-year 8% coupon bond currently sells for \$90. A 10-year 4% coupon bond currently sells for \$80. What is the 10-year zero rate?

[6 + 6 +8 = 20]

INDIAN STATISTICAL INSTIUTE

Semestral II Examination

M. STAT : II YEAR

Statistical Inference - III

28·4·2017

Marks: 60 Time: 3 hours Date: ~~15/05/2017~~

Answer all questions

1 (a) Let $\mathcal{B}$ be sufficient. Then show that every event, which is independent of $\mathcal{B}$ under each $P \in M$, is ancillary iff there does not exist any splitting set.

(b) Let $X^{(m)} = (X_1, \ldots, X_m)$ be $iid\ N(\mu_1, \sigma_1^2)$ and $Y^{(n)} = (Y_1, \ldots, Y_n)$ $iid\ N(\mu_2, \sigma_2^2)$ and assume that $X^{(m)}, Y^{(n)}$ are independent. Compute covariance between

$$S_1^2/S_2^2 \ and \ (\bar{X}_m - \bar{Y}_n)^2 \ / [(m-1)\, S_1^2 + (n-1)\, S_2^2/\rho]$$

where $\bar{X}_m = (X_1 + \cdots + X_m)/m, \bar{Y}_n = (Y_1 + \cdots + Y_n)/n,$

$$(m-1)S_1^2 = (X_1 - \bar{X}_m)^2 + \cdots + (X_m - \bar{X}_m)^2, (n-1)\, S_2^2$$

$$= (Y_1 - \bar{Y}_n)^2 + \cdots + (Y_n - \bar{Y}_n)^2 \ and \ \rho = \sigma_2^2 / \sigma_1^2$$

[8 + 12 = 20]

2 (a) Let $(\Omega, \mathcal{A}, M)$ be dominated by the σ finite measure μ. Then show that $(\Omega, \mathcal{A}, M)$ is boundedly complete if the smallest closed linear space containing $\left\{\frac{dP}{d\mu} : P \in M\right\}$ is $L_1(\Omega, \mathcal{A}, M)$, the converse being true provided $\mathcal{N}_M = \mathcal{N}_\mu$. Also give an application of this result.

(b) Show that the Rao – Blackwell property implies pairwise sufficiency.

[(5 + 7) + 8 = 20]

3. (a) Explain clearly the Stopping Rule Paradox of Stein. Give detailed and critical explanation of it.

(b) Let $(\Omega, \mathcal{A}, M)$ be dominated by the σ finite measure μ, and $\mathcal{B}$ a sufficient σ field. If there exists a regular conditional probability (on $\mathcal{A}$) given $\mathcal{B}$ with respect to μ, then show that there exists a regular conditional probability (on $\mathcal{A}$) given $\mathcal{B}$ which works for each $P \in M$.

[12 + 8 = 20]

INDIAN STATISTICAL INSTITUTE
Final Examination 2016-2017

MStatII DIRECTIONAL DATA ANALYSIS

Date: 2 May 2017 Maximum Marks: 50 Duration: 3 hrs.

Note: Show all your work. Marks are indicated in the margin.

Q. 1. (a) Let $\Psi = \Theta + \Phi$, (mod 2π), where Θ has probability density function $f(\theta)$ given by

$f(\theta) = K.[1 + 2\rho \quad \cos(\theta - \mu) + 2\rho^2 \cos 2(\theta - \mu)]$,

and K is the norming constant. Let Φ follow a Circular Uniform distribution, and Θ and Φ be independently distributed. Obtain the distribution of Ψ.

(b) Identify the wrapped Cauchy distribution as a member of the wrapped stable family of distributions. Show that its corresponding Fourier series form has an equivalent single term representation.

(c) Do EITHER (i) OR (ii):

(i) Derive the Locally Most Powerful (LMP) Invariant test for Isotropy against the density $f(\theta)$ defined in (a) above.

OR

(ii)Prove that the LMP test for Isotropy against the wrapped Cauchy distribution (with known mean direction) has a globally monotone power function. [16=4+5+7]

Q. 2. (a) (i) Establish the entries of the table along with their corresponding distributions for Analysis of Mean Directions (ANOMED) of k independent von Mises distributions. State your assumptions explicitly. (ii) Give a real-life example of the application of the corresponding test.

(b) (i) Starting with linear and circular probability distributions as members of exponential families, show how in general you can derive the distribution of the points in a unit disc. State precisely, without proof, the basic theorem you may need for this derivation.

(ii) Give a specific example of a distribution in (i). [18=8+10]

Q. 3. (a) (i) Define one measure of circular-circular correlation. Obtain a toroidal distribution for which independence of its component circular variables results iff the correlation so defined is zero.

Modified [illegible]

(ii) Derive one version of a bivariate spherical von Mises-Fisher distribution. Define a measure of spherical correlation and show how you will implement it in practice to test that the components of the bivariate spherical distribution you derived are spherically uncorrelated in that sense.

(iii) Describe some approaches of modeling and corresponding inference procedures for cylindrical regression with a circular dependent variable. Give a critical appraisal of these methods. [16=5+5+6]

INDIAN STATISTICAL INSTITUTE
Semester Examination: 2016-2017
M. Stat. II Year: Semester-II

Survival Analysis

Date: 02 May, 2017 Maximum Marks: 100 Duration: 3 hours

Note: Answer as many questions as you can but the maximum you can score is 100. Total mark is 113.

1. (a) Consider two independent lifetime random variables T_1 and T_2, where T_i follows gamma distribution with shape parameter p_i and scale parameter λ, $i = 1, 2$. Check whether the distribution of $T = T_1 + T_2$ is IFR or DFR, when $p_1 = 5/7$ and $p_2 = 3/7$.

   (b) Suppose the independent lifetimes T_1 and T_2 have proportional hazards

$$\lambda_i(t) = \lambda_0(t)\eta_i,$$

   for $i = 1, 2$, respectively, where $\eta_i > 0$. Find $P(T_1 < T_2)$.

[5+7 =12]

2. (a) Consider the life table data in the following table, where n_i is the number at risk at the beginning of ith interval, w_i is the number withdrawals and d_i is the number of failures in the ith interval. Fill up the d_i values. Derive the life table estimate of $S(20.5)$ assuming that the withdrawals occur at the end of each interval.

Interval (Months)	n_i	w_i	d_i
$[0, 5.5)$	112	1	
$[5.5, 10.5)$	93	1	
$[10.5, 15.5)$	76	3	
$[15.5, 20.5)$	55	0	
$[20.5, 25.5)$	45	0	
$[25.5, 30.5)$	34	1	
$[30.5, 40.5)$	25	2	
$[40.5, 50.5)$	10	3	
$[50.5, 60.5)$	3	2	
$[60.5, \infty)$	0	0	

   (b) Show that the Kaplan-Meier estimator becomes empirical survival function when there is no censoring.

   (c) Is the Kaplan-estimator a proper survival function if the largest observation is uncensored? Explain.

[(3+10)+4+3=20]

3. Suppose the survival function of two groups are related by $S_2(t) = S_1(t)^\gamma$, where $\gamma > 0$. Derive log-rank test for testing the equality of two survival functions. Suggest a test when the alternative hypothesis is having crossing hazard. [Hint. Mean and variance of X having $P[X = x] = \binom{M}{x}\binom{N-M}{n-x}/\binom{N}{n}$ are np and $\frac{N-n}{N-1}npq$, respectively, where $p = \frac{M}{N}$]

[8+2 = 10]

4. Let T be the lifetime with hazard rate $\lambda(t)$. Consider random right-censored data (X_i, δ_i), $i = 1, \ldots, n$. Suppose $N(t) = \sum_{i=1}^{n} N_i(t)$ where $N_i(t) = I(X_i \le t, \delta_i = 1)$, and $Y(t)$ is the number at risk just prior to time t. Let $M(t)$ be the corresponding counting process martingale.

   (a) Show that the predictable variation process $\langle M \rangle(t) = \int_0^t \lambda(s)Y(s)ds$.

(ii) Derive one version of a bivariate spherical von Mises-Fisher distribution. Define a measure of spherical correlation and show how you will implement it in practice to test that the components of the bivariate spherical distribution you derived are spherically uncorrelated in that sense.

(iii) Describe some approaches of modeling and corresponding inference procedures for cylindrical regression with a circular dependent variable. Give a critical appraisal of these methods. [16=5+5+6]

Modified [signature]

INDIAN STATISTICAL INSTITUTE
Semester Examination: 2016-2017
M. Stat. II Year: Semester-II

Survival Analysis

Date: 02 May, 2017 | Maximum Marks: 100 | Duration: 3 hours

Note: Answer as many questions as you can but the maximum you can score is 100. Total mark is 113.

1. (a) Consider two independent lifetime random variables T_1 and T_2, where T_i follows gamma distribution with shape parameter p_i and scale parameter λ, $i = 1, 2$. Check whether the distribution of $T = T_1 + T_2$ is IFR or DFR, when $p_1 = 5/7$ and $p_2 = 3/7$.

   (b) Suppose the independent lifetimes T_1 and T_2 have proportional hazards

$$\lambda_i(t) = \lambda_0(t)\eta_i,$$

   for $i = 1, 2$, respectively, where $\eta_i > 0$. Find $P(T_1 < T_2)$.

[5+7 =12]

2. (a) Consider the life table data in the following table, where n_i is the number at risk at the beginning of ith interval, w_i is the number withdrawals and d_i is the number of failures in the ith interval. Fill up the d_i values. Derive the life table estimate of $S(20.5)$ assuming that the withdrawals occur at the end of each interval.

Interval (Months)	n_i	w_i	d_i
$[0, 5.5)$	112	1	
$[5.5, 10.5)$	93	1	
$[10.5, 15.5)$	76	3	
$[15.5, 20.5)$	55	0	
$[20.5, 25.5)$	45	0	
$[25.5, 30.5)$	34	1	
$[30.5, 40.5)$	25	2	
$[40.5, 50.5)$	10	3	
$[50.5, 60.5)$	3	2	
$[60.5, \infty)$	0	0	

   (b) Show that the Kaplan-Meier estimator becomes empirical survival function when there is no censoring.

   (c) Is the Kaplan-estimator a proper survival function if the largest observation is uncensored? Explain.

[(3+10)+4+3=20]

3. Suppose the survival function of two groups are related by $S_2(t) = S_1(t)^\gamma$, where $\gamma > 0$. Derive log-rank test for testing the equality of two survival functions. Suggest a test when the alternative hypothesis is having crossing hazard. [Hint. Mean and variance of X having $P[X = x] = \binom{M}{x}\binom{N-M}{n-x}/\binom{N}{n}$ are np and $\frac{N-n}{N-1}npq$, respectively, where $p = \frac{M}{N}$]

[8+2 = 10]

4. Let T be the lifetime with hazard rate $\lambda(t)$. Consider random right-censored data (X_i, δ_i), $i = 1, \ldots, n$. Suppose $N(t) = \sum_{i=1}^n N_i(t)$ where $N_i(t) = I(X_i \le t, \delta_i = 1)$, and $Y(t)$ is the number at risk just prior to time t. Let $M(t)$ be the corresponding counting process martingale.

   (a) Show that the predictable variation process $\langle M \rangle (t) = \int_0^t \lambda(s)Y(s)ds$.

(b) Derive the expression of Nelson-Aalen estimator using counting process notations. Give the expression of the estimator in terms of the data.

(c) Show that the Nelson-Aalen estimator has negative bias, when $P(Y(u) > 0) > 0$. Write down the condition when the bias tends to 0 as $n \to \infty$.

[9+ (4+2) + (5+2) =22]

5. Let T be a continuous survival time variate. Given a covariate Z, suppose that $Y = \ln T$ follows a linear model

$$Y = \mu + \beta^T Z + \sigma W,$$

where W is error random variable with survival function $S_W(w) = \exp(-\exp(w))$.

(a) Show that it is a proportional hazard model.

(b) Find the relationship between the median survival time under covariate effect with the baseline median survival time

(c) Write down the likelihood function based on random right-censored data.

[6+5+3=14]

6. Consider the proportional hazard model $\lambda(t; Z) = \lambda_0(t)\exp(\beta^T Z)$, where $\lambda_0(t)$ is unspecified and arbitrary. .Derive Breslow's estimator for $\Lambda_0(t)$, the cumulative baseline hazard function. Show that Breslow's method also gives partial likelihood for estimating β.

[$8 + 2 = 10$]

7. (a) Suppose T is the failure time and J is the failure mode in a competing risks set up. Assume that T is continuous. Show that the joint distribution of (T, J) is completely specified by the mode-specific hazards.

(b) What is subdistribution function in competing risk model? Is it a proper distribution function? Explain.

(c) Suppose that the joint survival function of the latent failure times T_1 and T_2 for two competing risks is

$$S(t_1, t_2) = \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \lambda t_1 t_2), \quad 0 < t_1 < \infty, \quad 0 < t_2 < \infty.$$

Construct a joint survival function S^* such that both S and S^* have the same cause specific hazards.

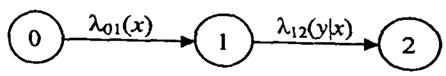
[5+(2+2) + $6 = 15$]

8. Consider a simple illness-death model as shown in the figure, where 0, 1 and 2 represents healthy, illness and death states, respectively. Let T_{01} be the sojourn time in state 0 and T_{12} the same in state 1 before death. Let C be the censoring random variable independent of both T_{01} and T_{12}. Suppose that there are n individuals in a study with state 0.

(a) Describe the data with appropriate notation when time to occurrence of illness state is observed for all the individuals.

(b) Derive the maximum likelihood estimate of $P[T_{01} + T_{12} \le t]$, when $\lambda_{01}(x) = \lambda_{01}$ and $\lambda_{12}(y|x) = \lambda_{12}$, based on data described in (a).

[3+7=10]



INDIAN STATISTICAL INSTIUTE

M Stat. (2nd Year) 2016 – 17

Semestral Examination

Subject: Theory of Games and Statistical Decisions

Date: 03.05.2017 Full Marks: 60 Duration: 2 ½ hours

Attempt all questions

1. Let X be a γ - dimensional random vector.
   $\theta \in R^{\gamma}$ and θ is unknown.
   X follows $N_{\gamma}(\theta, I_{\gamma})$
   Consider the problem of estimating θ.

   a) For $\gamma \geq 3$, discuss Stein phenomenon in this context.

   b) What is $\delta_{JS}(X)$, James – Stein estimate of θ in this context?

   c) Show that with loss $L(\delta, \theta) = \frac{1}{\gamma}\sum_{i=1}^{\gamma}(\delta_i - \theta_i)^2$, $\delta_{JS}(X)$ is uniformly better than X.

   [2 + 3 + 13]

2. a) Define Bi-matrix Non-cooperative Game.

   b) For $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, consider the Bi-matrix Game with A, B.
   Using general solution of 2 X 2 Bi-matrix games, find admissible situations (for each player) and equilibrium situations for the above.

   Relating the game in b) as the famous game of "Battle of sex" discuss the relevance of equilibrium situations.

   [2 + (8 + 6)]

3. a) In a problem of estimation with parameter space R, show that if an estimator with constant risk function is admissible, then it is minimax. Give an example.

   b) Consider a matrix game with matrix $A: m\ x\ n$.
   Discuss that if an equilibrium situation exists and value of the game is V, then to play the game is equivalent to paying amount V from 2nd player to 1st one, provided both the players are rational?

   [6 + 8]

4. For Matrix $A: m\ x\ n$,
   Let $v\ (A)$ denote value of the mixed extension of the matrix game A.
   $v\ (A)$ is a function from $R^{m\ x\ n} \rightarrow R$.

   a) Show that $v\ (A)$ in non-decreasing in each coordinate.

   b) Show that as function of mn real variables $v\ (A)$ is continuous.

   [4 + 8]

****

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2016–2017)

M. Stat Second Year

Inference for High Dimensional Data

Date: 05/5/17 Marks: ..60.. Duration: .3 hours.

Attempt all questions

1. (a) Prove that the volume $V_p(r)$ of a p-dimensional ball of radius $r > 0$ is given by

$$V_p(r) = \frac{\pi^{p/2}}{\Gamma\left(\frac{p}{2}+1\right)} r^p,$$

where for $x > 0$, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

(b) Consider n p-dimensional observations $X^{(1)}, \ldots, X^{(n)}$. For large p, compute a reasonable lower bound on the number n of points needed in order to fill the hypercube $[0,1]^p$ in such a way that at any $x \in [0,1]^p$, there exists at least one point at distance less than 1 from x.

[10+10=20]

2. Consider the linear model $\mathbf{Y} = \mathbf{f}^* + \epsilon = \mathbf{X}\beta^* + \epsilon$ in the co-ordinate sparse setting. Assume that the p columns of $\mathbf{X}$ are orthogonal, and that the components of ϵ are *iid* $N\left(0, \sigma^2\right)$. Consider the family $\mathcal{M}$ and the models S_m of the co-ordinate sparse setting and consider the penalty $pen(m) = \lambda|m|$, for some $\lambda > 0$, where $|m|$ denotes the cardinality of m. Letting $\mathbf{X}_j$ denote the j-th column of $\mathbf{X}$, set $Z_j = \mathbf{X}_j^T \epsilon / (\|\mathbf{X}_j\| \sigma)$. Assume henceforth that $f^* \in S_{m^*}$ with $|m^*| = D^*$.

(a) Writing $\mathbf{P}_j$ for the projection on the line spanned by $\mathbf{X}_j$, prove that

$$\begin{aligned}\|\hat{\mathbf{f}}_{\hat{m}_\lambda} - \mathbf{f}^*\|^2 &= \|\mathbf{f}^* - \sum_{j \in \hat{m}_\lambda} \mathbf{P}_j \mathbf{f}^*\|^2 + \sum_{j \in \hat{m}_\lambda} Z_j^2 \sigma^2 \\ &\geq \sum_{j \in \hat{m}_\lambda \setminus m^*} Z_j^2 \sigma^2 \geq \left(|\hat{m}_\lambda| - D^*\right) \lambda \sigma^2.\end{aligned}$$

(b) Prove that for $a, x \in \mathbb{R}^+$, we have

$$\int_{x-a}^{x} e^{-z^2/2} dz \geq \int_x^{x+a} e^{-z^2/2} dz.$$

(c) Prove that $E\left[|\hat{m}_\lambda|\right] \geq pP\left(\tilde{Z}^2 \geq \lambda\right)$, where $\tilde{Z}$ is a standard Gaussian random variable.

(d) For $K < 1$ and $D^* \ll p^{1-K} (\log p)^{-1/2}$, prove that

$$\begin{aligned} & E\left[\|\hat{\mathbf{f}}_{\hat{m}_{2K \log p}} - \mathbf{f}^*\|^2\right] \\ & \geq \left(E\left[|\hat{m}_{2K \log p}|\right] - D^*\right) 2K\sigma^2 \log p \\ & \overset{p\to\infty}{\sim} p^{1-K}\sigma^2 \sqrt{\frac{4K\log p}{\pi}}. \end{aligned}$$

(e) For $D \in \{1, \ldots, p\}$, define $V_D(\mathbf{X}) = \{\mathbf{X}\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p, |\boldsymbol{\beta}|_0 = D\}$, where $|\boldsymbol{\beta}|_0 = \text{Card}\{j : \beta_j \neq 0\}$. Also consider the minimax risk $R[\mathbf{X}, D] = \inf_{\hat{\mathbf{f}}} \sup_{\mathbf{f}^* \in V_D(\mathbf{X})} E_{\mathbf{f}^*}\left[\|\hat{\mathbf{f}} - \mathbf{f}^*\|^2\right]$, where the infimum is taken over all the estimators. For $K < 1$ and $D^* \ll p^{1-K} (\log p)^{-\frac{1}{2}}$, prove that for any $\mathbf{f}^* \in V_{D^*}(\mathbf{X})$,

$$E\left[\|\hat{\mathbf{f}}_{\hat{m}_{2K \log p}} - \mathbf{f}^*\|^2\right] \geq R[\mathbf{X}, D^*], \quad \text{as} \quad p \to \infty.$$

[4+4+4+4+4=20]

3. In a normal linear regression setting with error variance σ^2 consider a collection $\{S_m : m \in \mathcal{M}\}$ of linear subspaces of $\mathbb{R}^n$ and let $\hat{\mathbf{f}}_m = \text{Proj}_{S_m}^{\mathbf{Y}_n}$ be the projection of the response data $\mathbf{Y}_n = (y_1, \ldots, y_n)^T$ on S_m. Let $\hat{r}_m = \|\mathbf{Y} - \hat{\mathbf{f}}_m\|^2 + 2d_m\sigma^2 - n\sigma^2$, with $d_m = \dim(S_m)$. Define $\hat{\mathbf{f}} = \sum_{m\in\mathcal{M}} w_m \hat{\mathbf{f}}_m$, with $w_m = \pi_m \exp\left(-\frac{\beta \hat{r}_m}{\sigma^2}\right)/\mathcal{L}$, where $\mathcal{L} = \sum_{m\in\mathcal{M}} \pi_m \exp\left(-\frac{\beta\hat{r}_m}{\sigma^2}\right)$, and $\pi_m \propto \frac{e^{-|m|}}{\binom{p}{|m|}}$. Propose, with proper justification, a Metropolis-Hastings algorithm for computing $\hat{\mathbf{f}}$. [10]

4. (a) Show geometrically how Lasso selects variables.

   (b) Prove that for $p \in [0, \infty]$ the sub-differential of the ℓ^p-th norm of $\mathbf{x} \in \mathbb{R}^n$ is given by

   $$\partial|\mathbf{x}|_p = \{\boldsymbol{\phi} \in \mathbb{R}^n : \langle\boldsymbol{\phi}, \mathbf{x}\rangle = |\mathbf{x}|_p \text{ and } |\boldsymbol{\phi}|_q \leq 1\},$$

   where $1/p + 1/q = 1$ (you may directly use the extremal equality $|\mathbf{x}|_p = \sup\{\langle\boldsymbol{\phi}, \mathbf{x}\rangle : |\boldsymbol{\phi}|_q \leq 1\}$). Hence obtain the sub-differential of the ℓ^1-norm for any n-dimensional vector.

   (c) Discuss mathematically when the Lasso solution is unique.

   (d) Propose a practical method for checking uniqueness of the Lasso solution.

[2+3+3+2=10]

INDIAN STATISTICAL INSTITUTE

Final Examination: 2016-17

Course Name	:	**M. Stat. 2nd year**
Subject Name	:	**Weak convergence and empirical processes**
Date	:	15.5.2017
Maximum Marks	:	**60**
Duration	:	**3 hours**

Answer **any five** of the following questions. Each question carries 12 marks. Please state clearly **every** result you use. If you use a result not stated in this course, you will have to prove it.

1. Let $\mathcal{S}$ be a complete separable metric space. Show that a sequence of probability measures $P_1, P_2, \ldots$ on $\mathcal{S}$ is tight if and only if there exist compact subsets $K_1, K_2, \ldots$ of $\mathcal{S}$ such that
$$\lim_{m\to\infty} \liminf_{n\to\infty} P_n(K_m) = 1 .$$

2. For each $n \in \mathbb{N}$, suppose that X_n is a $C[0,1]$-valued random element having the following properties. The family $\{X_n(0) : n \geq 1\}$ is a tight family of real valued random variables. Furthermore, for every $n \geq 1$, X_n has continuously differentiable paths, and the family $\{\max_{t\in[0,1]} |X_n'(t)| : n \geq 1\}$ is also a tight family of real valued random variables. Show that $\{X_n : n \geq 1\}$ is tight in $C[0,1]$.

3. Let $(B_t : 0 \leq t \leq 1)$ be a Brownian motion. Show that there does not exist a random variable Z such that
$$\frac{B_t}{\sqrt{t}} \to Z \text{ a.s.},$$
as $t \downarrow 0$.

4. Let $P_1, P_2, \ldots, P_\infty$ be probability measures on a complete separable metric space $\mathcal{S}$ such that as $n \to \infty$,
$$P_n \Rightarrow P_\infty .$$
Show that there exist $\mathcal{S}$-valued random elements $X_1, \ldots, X_\infty$ defined on some probability space $(\Omega, \mathcal{F}, P)$ such that
$$P \circ X_n^{-1} = P_n,\ 1 \leq n \leq \infty,$$
and for every continuous function $f : \mathcal{S} \to \mathbb{R}$ such that $\{f(X_n) : 1 \leq n < \infty\}$ is uniformly integrable, it holds that as $n \to \infty$,
$$f(X_n) \to f(X_\infty) \text{ in } L^1 .$$

P.T.O

5. Consider $D[0,1]$ equipped with the Skorohod topology. Show that $C[0,1]$ is a Borel subset of $D[0,1]$ in this topology.

6. Let $X_1, X_2, \dots$ be i.i.d. real valued random variables taking finitely many values $\alpha_1, \dots, \alpha_k$ with respective probabilities $p_1, \dots, p_k$. Show that there exists an universal finite constant C such that for every $n \geq 1$,

$$\mathrm{E}\left[\max_{j=1,\dots,k}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i = \alpha_j) - p_j\right|\right] \leq C\sqrt{\frac{2+\log k}{n}}.$$

Indian Statistical Institute
Semestral Examination: (2016–2017)
M. Stat.–II year
Special topic: Random Matrices
Total marks=120

Date: 05/05/2017 Time: 3 hours and 30 minutes

Answer all questions. Questions carry equal marks.

1. Suppose T_n is the $n \times n$ tri-diagonal symmetric Toeplitz matrix with the input $x_0 = 0$ and $x_1 = 1$. Find the limiting spectral distribution of T_n. What happens if $x_0 = 1$?

2. Suppose C_1 and C_2 are $n \times n$ independent circulant matrices. Show that $C_1C_2 = C_2C_1$ and they are simultaneously diagonalisable. Now suppose further that the inputs are i.i.d. Gaussian. Show that C_1 and C_2 converge jointly (in the $*$ sense) to independent complex normal random variables in the state $\frac{1}{n}Tr \otimes E$ as $n \to \infty$.

3. Suppose s_1 and s_2 are free standard semi-circle variables. Find the Cauchy transform of $s_1s_2s_1$ in terms of a functional equation.

4. Consider the set $N_n = \{1, 2, \ldots, 2n\}$. Suppose π is a non-crossing partition of N_n. Define its Kreweras complement K_π. Explain how a partition can be considered as a permutation (maintaining the natural ordering). Let γ be the cyclic permutation of N_n. If π is also a pair-partition, show that $K_\pi = \pi\gamma$.

5. Suppose in a non-commutative probability space $(\mathcal{A}, \phi)$, a, b, s_1, s_2 are self adjoint and $\{a, b\}$ is free of $\{s_1, s_2\}$. Further, s_1 and s_2 are free standard semicircular and $\phi(a) = \phi(b) = 0$. Are s_1as_1 and s_2bs_2 free?

6. Consider the probability measure μ which puts mass 1/2 at -1 and 1. Describe the three fold additive free convolution of μ as explicitly as possible.

7. Define the Möbius function. Using the moment-(free) cumulant relations, derive formulae for the moments and (free) cumulants of a Poisson and a free Poisson variable.

INDIAN STATISTICAL INSTITUTE
M.STAT Second Year
2016-17 Semester II

Computational Finance
Final Examination

15.05.2017

Points for each question is in brackets. Total Points 100.
Students are allowed to bring 4 pages (one-sided) of hand-written notes
Duration: 3 hours

1. (15) We have three interpolation points $(x_i, f(x_i)), i = 0, 1, 2$ with $x_i = x + (i-1)h$ and want to approximate $f'(x)$ using Lagrange interpolation. Show that this leads to the second-order centered approximation of the first-derivative, that is,
$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{1}{6} f'''(\xi) h^2$$
where $\xi \in (\min\{x, x_0, x_1, x_2\}, \max\{x, x_0, x_1, x_2\})$.

2. (20) Suppose you want to evaluate the integral of the function $f(x) = e^x$ over $[0, 1]$ using Monte Carlo simulation from Unif(0,1). Use the control variable method with $g(x) = x$ and show that the variance is reduced by a factor of 60. Is there much additional improvement if you use a general quadratic function of x?

3. (10) Describe the Mersenne twister random number generator algorithm.

4. (2+6+6+6) Consider the case of a binomial lattice for which the probability of an up move has the same value p at all nodes.
   (a) What is the distribution of the total number of up moves N through an m-step lattice?
   (b) How can you generate stratified samples from this distribution with equiprobable strata?
   (c) Show that given N, all paths through the lattice with N up moves are equally likely.
   (d) Hence outline a procedure to generate a path conditional on N.

5. (15) Let us assume that $F_1(x)$ and $F_2(x)$ are two cdf satisfying $F_1(x) \leq F_2(x)$ for all values of x. If these two distributions are proposed as models for the returns, which of these two distributions will give the larger Value-at-risk at level 0.01? Show that from the given information it is not possible to infer which distribution leads to higher expected shortfall.

6. (10+10) The process ϵ_t is said to be a strong ARCH(1) process, if $E[\epsilon_t | F_{t-1}] = 0$, $Var(\epsilon_t | F_{t-1}) = \sigma_t^2$ and $Z_t = \epsilon_t / \sigma_t$ is i.i.d. with $\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2$. Assuming ϵ_t to be a strong ARCH(1) process,
   (a) Show that ϵ_t is white noise, but not independent. In particular show that the autocorrelations of ϵ_t^2 are non-zero.
   (b) If Z_t has normal distribution, show that the distribution of ϵ_t is leptokurtic.

INDIAN STATISTICAL INSTITUTE

Final Examination : Semester II (2016-17)

M. Stat. II Year

Clinical Trials

Date: 15.05.2017 Maximum marks: 100 Time: 3 hours

(Total mark is 105.)

1. Answer the following questions briefly.

   (a) Mention two ethical issues in the conduct of a Randomized Control Trial (RCT).

   (b) Explain assessment bias.

   (c) Describe Clinically Important Difference in treatment effect.

   (d) In Randomized Permuted Block (RPB) design with two treatment groups, the allocation of patients are independent and identically distributed (IID). Prove or disprove.

   (e) Suggest a multi-treatment generalization of the Play-the-Winner (PW) rule.

   (f) Consider a clinical trial with binary responses and one continuous prognostic factor. However, the allocation of patients was not randomized. What kind of analysis do you suggest and why?

   (g) Describe type I error spending function specifying its advantage over other methods of obtaining group sequential boundaries.

   [3+2+2+4+3+3+5=22]

2. An *RCT* is to be conducted with patients in two groups T and C. The outcome variable is binary with probability of success being p_T and p_C, respectively. Assume that p_C is known and the objective is to be able to detect an increase of τ in the success probability in the T-group through a size α test with the power being at least $1-\beta$. Determine the minimum number of patients to be enrolled in each of the two groups using approximate normality and the approximate expression for the power function. Give details. [15]

3. Describe the urn design $UD(r,s)$ for allocating patients in two treatment groups. Explain its advantage over RPB and Efron's Biased Coin design. Argue that this design is almost close to simple randomization in the long run. [8+3+3=14]

4. Consider a Randomized Play-the-Winner rule RPW(α,β) for allocating patients in two treatment groups T and C. Prove that the limiting probability of allocation to the T-group is $q_C/(q_T+q_C)$, where $p_T=1-q_T$ and $p_C=1-q_C$ are the success probabilities in the two groups, respectively. Noting that this limiting probability is same as that of the PW rule, what is the advantage of the RPW rule over the PW rule?

   Let $S_T(n)$ denote the number of successes in the T-group out of the first n patients. Find $\lim_{n\to\infty}\frac{1}{n}E[S_T(n)]$. [14+1+5=20]

5. (a) Describe a modification of the RPW rule for allocating patients in two treatment groups T and C in the presence of a categorical prognostic factor with $(K+1)$ ordered levels $0, 1, \cdots, K$.

   (b) Describe a modification of the RPW rule for allocating patients in two treatment groups T and C when the response variable is categorical with ordered grades $0, 1, \cdots, M$.

   (c) In both (*a*) and (*b*) above, derive π_{i+1}, the conditional probability of allocating the $(i+1)$st patient into the T-group, given the allocation and response history up to the ith patient.

[4+4+(3+3)=14]

6. In the context of analyzing binary response data from an RCT with two treatment groups, what is odds-ratio? How is it estimated? Derive an expression for the corresponding variance estimate. [3+2+5=10]

7. Consider group sequential trial with two treatment groups T and C, and N planned interim analyses. The response variables in the two groups are normal with unknown means and known common variance. In each period, it is planned to consider $2n$ and n patients in T and C groups, respectively. Develop the group sequential method with a fixed overall type I error α by specifying how to obtain the nominal levels to be used at the successive testing. Consider Pocock's choice for this purpose. [10]

INDIAN STATISTICAL INSTITUTE
Final Examination - Back Paper 2016-2017

MStat II DIRECTIONAL DATA ANALYSIS

Date: ~~XX May 2017~~ 10·07·2017 Maximum Marks: 50 Duration: 3 hrs.

Q. 1. [25=4+8+7+6] Consider the circular distribution with probability density function given by

$f(\theta) = K.[1 + 2\rho\cos(\theta - \mu)]$.

(a) Obtain K. (b) Obtain the mean direction and the circular variance. (c) Derive the characteristic function for $f(\theta)$. (d) Verify whether $f(\theta)$ is unimodal.

Q. 2. [25=7+7+11] (a) Let $\Psi = \Theta + \Phi$, where Θ has distribution $f(\theta)$ given in Q. 1 above and Φ has a Circular Uniform distribution, and Θ and Φ are independently distributed. Obtain the distribution of Ψ.

(b) Identify the wrapped Cauchy distribution as a member of the wrapped stable family of distributions. Show that its corresponding Fourier series form has an equivalent single term representation.

(c) Derive the Locally Most Powerful Invariant test for Isotropy against the distribution $f(\theta)$ defined in Q. 1 above.

Q. 3. [25=12+(9+4)(a) Establish the entries of the table for Analysis of Mean Directions (ANOMED) for k independent von Mises distributions, stating your assumptions explicitly. Give a real-life example of the application of the corresponding test.

(b) (i) Starting with linear and circular probability distributions as members of exponential families, show how in general you can derive the distribution of the points on a torus. State precisely, without proof, the basic theorem you may need for this derivation. (ii) Give a specific example of a distribution in (i).

Q. 4. [25=15+10] (a) Derive the joint distribution of C and S for a sample on n observations from the von Mises distribution vM (μ, κ).

(b) Derive a test for independence between [redacted] two circular random variables distributed on [redacted] a torus.

INDIAN STATISTICAL INSTITUTE

M.STAT Second Year

2016-17 Semester II

Computational Finance

Back Paper Examination

12.07.17

Points for each question is in brackets. Total Points 100.
Students are allowed to bring 4 pages (one-sided) of hand-written notes
Duration: 3 hours

1. (12+3)Consider the Black Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

   (a) Use a mesh of equal space steps of size δS and equal time steps of size δt, central differences for space derivatives, backward differences of time derivatives, to obtain the explicit finite difference equations

$$V_n^m = a_n V_{n-1}^{m+1} + b_n V_n^{m+1} + c_n V_{n+1}^{m+1}$$

   where V_n^m is the finite difference approximation to $V(n\delta S, m\delta t)$ and

$$\begin{aligned} a_n &= \frac{1}{2}(\sigma^2 n^2 - rn)\delta t \\ b_n &= 1 - (\sigma^2 n^2 + r)\delta t \\ c_n &= \frac{1}{2}(\sigma^2 n^2 + rn)\delta t \end{aligned}$$

   (b) Why is this an explicit method? What boundary and initial/final conditions are appropriate?

2. (5+15) The objective of this problem is to show that Binomial Option pricing formula with

$$u = e^{\sigma\sqrt{\Delta t}}, d = 1/u \quad \text{and} \quad p = \frac{e^{r\Delta t} - d}{u - d}$$

   approximately solves the BS equation as $\Delta t \to 0$.

   (a) Write down the relationship induced by no arbitrage, between prices at time t and $t + \Delta t$, in the single period Binomial model.

   (b) Use Taylor expansion and take limit at $\Delta t \to 0$ to show that this leads to the Black Scholes PDE.

3. (10) Show that for estimating $E[f(Z)]$ based on an antithetic pair $(Z, -Z)$, where $Z \sim \mathcal{N}(0, I)$, antithetic sampling eliminates all variance if f is antisymmetric.

4. (10+7+8)

   (a) For a given level α compute the Value at Risk and Expected shortfall under the following distributional assumption in the loss distribution of the portfolio:

P.T.O

i. Exponential with rate parameter λ.

ii. Standard Pareto distribution with shape parameter ξ, location parameter 0 and scale parameter 1, that is, the pdf is

$$f_\xi(x) = (1+\xi x)^{-(1+1/\xi)} \quad \text{if} \quad x > 0$$

(b) For each $\alpha \in (0,1)$, derive an equation that the rate parameter λ and the shape parameter ξ must satisfy in order for the values of $\mathrm{VaR}(\alpha)$ for the two distributions to be the same.

(c) Assuming that the parameters λ and ξ satisfy the relationship in (b) above, compare the corresponding values of the expected short fall in the two models and comment on the differences.

5. (10+10) Let C denotes the copula of the two random variables X and Y. Assume that the marginal cdfs are continuous and strictly increasing.

(a) Show that $P[\max(X,Y) \leq t] = C(F_X(t), F_Y(t))$

(b) Prove that the Spearman correlation coefficient $\rho(X,Y)$ is given by the formula

$$\rho(X,Y) = 12\int_0^1\int_0^1 uvC(u,v)dudv - 3$$

6. (10) Explain the problem of cointegration of two time series.

INDIAN STATISTICAL INSTITUTE
Second Semester Backpaper Examination: 2016-17

M. Stat. II Year
Theory of Games and Statistical Decisions

Date: 12.07.17 **Maximum Marks: 100** **Duration: 3 Hours**

Attempt all questions.

1.a) State and prove Nash's Theorem in context of equilibrium situation of finite non-cooperative games in mixed extension.

b) Show that if there are more than one equilibrium situations, for same fixed player pay off may differ from one equilibrium situation to another, provided number of players is greater than 2.

c) Show that in a zero-sum, two-player non-cooperative game, pay off for the same player is constant for each equilibrium situation.

d) Prove rectangular property of the set of equilibrium situations in a two player, zero-sum non-cooperative game.

e) Can you expect similar property as in d), in general non-cooperative games with more than 2 players.

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

2. Let $f(x,y)$ be a function from $R \times R \to R$.

a) Define where (x_0, y_0) is to be called a saddle point of $f(x,y)$.

b) Show that a saddle point of $f(x,y)$ exists iff $Max_x \inf_y f(x,y)$ and $Min_y \sup_x f(x,y)$ both exists and are equal.

c) How does saddle points of a bivariate function relate to existence of equilibrium situations in two - player non-cooperative games.

d) Show that in mixed extension of matrix game A: $m \times n$, $\max_X \min_Y XAY^T$ and $Min_Y \max_X XAY^T$ always exist, where X, Y are mixed strategies.

[2+18+4+18]

3. a) Let $A = ((a_{ij}))_{n \times n}$ with $a_{ii} = 4$ for all i and $a_{ij} = 2 \;\; \forall\, i \neq j$
Derive value of the mixed extension of the matrix game A.

b) Let $A: 3 \times 3$ be a diagonal matrix with each diagonal entries positive.

Find optimal strategies and value of the mixed extension of the matrix game A.

[10+10]

*****

INDIAN STATISTICAL INSTITUTE

M.Stat. II Year

Backpaper Examination : Semester II : 2016-2017

BROWNIAN MOTION AND DIFFUSIONS

Date : 12.07.2017 **Maximum Score : 45** Time : 3 Hours

Note : This paper carries questions worth a **total** of **100 marks**. Answer as much as you can. The **maximum** you can score is **45**.

1. (a) State clearly Kolmogorov Consistency Theorem and Kolmogorov Continuity Criterion.
   (b) Show that there is a zero-mean Gaussian process $\{X_t, t \geq 0\}$ with continuous paths and $\mathrm{Cov}(X_s, X_t) = \exp(-|s-t|)$.
   (c) Show that it is not possible to have a process $\{X_t, t \geq 0\}$ with continuous paths where the X_t, $t \geq 0$, are i.i.d. $N(0,1)$ random variables. $[(2+2)+8+8] = [20]$

2. (a) Let $\{B_t, t \in [0,1]\}$ be a SBM. Show that $X = \int_0^1 B_s^2\, ds$ is a random variable and find its mean and variance.
   (b) Let $\{B_1(t)\}, \ldots, \{B_k(t)\}$ be independent standard Brownian motions. Show that for every $\mathbf{x} \in \mathbb{R}^k$ with $\|\mathbf{x}\| = 1$, $\{B(t) = \langle \mathbf{x}, (B_1(t), \ldots, B_k(t))\rangle\}_{t \geq 0}$ is a standard Brownian motion. Here, $\|\cdot\|$ denotes the usual euclidean norm on $\mathbb{R}^k$.
   (c) Let $\{B_t, t \in [0,\infty)\}$ $\{\beta_t, t \in [0,\infty)\}$ be two SBMs idependent of each other. Denoting $\tau_a = \inf\{t \geq 0 : B_t \geq a\}$ for $a > 0$, find the distribution of the random variable β_{τ_a}.
   (d) Let $\{B_t, t \in [0,\infty)\}$ be a SBM.
   (i) Denote $\alpha = \frac{m+1}{2m-2}$, where $m \geq 2$ is an integer. For integers $j \geq 1, k \geq 1$, denoting $D_{j,k}$ to be the event $\bigcap_{n>(m+1)k} \bigcup_{1 \leq i < n-m} \bigcap_{l=1}^{m} \{\omega : |B(\frac{i+l}{n}, \omega) - B(\frac{i+l-1}{n}, \omega)| \leq j[(l+1)^\alpha + l^\alpha]/n^\alpha\}$, show that $P(D_{j,k}) = 0$.
   (ii) Deduce that, there is a P-null set N such that, if $\omega \notin N$, then for all $t \geq 0$ and for all $\alpha > \frac{1}{2}$, the set $\left\{\frac{B_s(\omega) - B_t(\omega)}{|s-t|^\alpha},\ s \neq t\right\}$ remains unbounded, as $s \to t$. $[4+4+4+(6+6)] = [24]$

3. Let C denote the set of all real continuous functions on $[0,\infty)$ and $\{X_t,\ t \geq 0\}$ denote the co-ordinate process on C. Denoting $\{\mathcal{C}_t\}_{t \geq 0}$ to be the natural filtration of the process $\{X_t\}$ and $\mathcal{C} = \vee\{\mathcal{C}_t, t \geq 0\}$, show the following:
   (a) τ is a $\{\mathcal{C}_t\}$-stopping time if and only if for every $\omega, \omega' \in \mathsf{C}$, $\tau(\omega) \leq t$ and $X_s(\omega) = X_s(\omega')$ for $s \leq t$ imply $\tau(\omega) = \tau(\omega')$.
   (b) For any open set $G \subset \mathbb{R}$, the random variable $\tau(\omega) = \inf\{t \geq 0 : X_t(\omega) \in G\}$ is a $\{\mathcal{C}_t\}$-optional time, but need not be a $\{\mathcal{C}_t\}$-stopping time.
   (c) A real random variable Z on $(\mathsf{C}, \mathcal{C})$ is $\mathcal{C}_\tau$-measurable if and only if for every $\omega \in \mathsf{C}$, $Z(\omega) = Z(\omega^\tau)$, where $\omega^\tau(t) = \omega(t \wedge \tau(\omega))$. $[8+6+8] = [22]$

4. Let $S = [0,\infty)$ and, for each $x \in S$, denote P_x to be the distribution of the process $\{X_t = |x + B_t|, t \geq 0\}$, where $\{B_t, t \in [0,\infty)\}$ be a SBM.
   (a) Show that the family $\{P_x,\ x \in S\}$ of probabilities on the appropriate path space is a Markov process with state space $S = [0,\infty)$ and find its transition probabilities.
   (b) Show that the above Markov process has the Feller property.
   (c) Taking $C_b(S)$ as the underlying Banach space, find the generator of the above Markov process. [May use: $\int_0^\infty e^{-(\alpha u - \beta/u)^2} du = \frac{\sqrt{2\pi}}{2\alpha}$, for any $\alpha > 0, \beta > 0$.] $[8+4+8] = [20]$

5. Let $(\Omega, \mathcal{A}, \{\mathcal{A}_t,\ t \geq 0\}, P)$ be a filtered complete probability space with $\mathcal{A}_0$ containing all P-null sets and let $\{B_t, t \in [0,\infty)\}$ be a SBM with respect to $\{\mathcal{A}_t,\ t \geq 0\}$.
   (a) Directly from the definiton of stochastic uintgeral and using properties of SBM, find the unique continuous path process process $\{X_t\}_{t \geq 0}$ such that $X_t = \int_0^t B_s\, dB_s\ \forall\, t \geq 0$.
   (b) Let $h \in \mathcal{L}$ and let $\{X_t\}_{t \geq 0}$ be the a.s. unique adapted process with continuous paths such that $X_t = \int_0^t h_s\, dB_s\ \forall\ t \geq 0$. Show that if τ is any finite stopping time and if $\tilde{h}_s = h_s 1_{\{\tau \geq s\}}$, $s \geq 0$, then $\tilde{h} \in \mathcal{L}$ and, almost surely, $\int_0^t \tilde{h}_s\, dB_s = X_{\tau \wedge t}\ \forall\, t$. $[6+8] = [14]$

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, Second Semester, 2016-17

Semestral (Back paper) Examination

Statistical Computing II

Date: 11.07.2017 Full Marks: 100 Time: 4 Hours

[Answers should be brief and to the point. If you use any standard result, mention it clearly.]

1. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed $U(0, \theta)$ variables. Let $X_1^*, X_2^*, \ldots, X_n^*$ and $X_1^{**}, X_2^{**}, \ldots, X_n^{**}$ be two samples of size n generated using the parametric bootstrap method and the nonparametric bootstrap method, respectively. Let $G_n(\cdot)$, $G_n^*(\cdot)$ and $G_n^{**}(\cdot)$ be the distribution functions of $M_n = \max\{X_1, X_2, \ldots, X_n\}$, $M_n^* = \max\{X_1^*, X_2^*, \ldots, X_n^*\}$ and $M_n^{**} = \max\{X_1^{**}, X_2^{**}, \ldots, X_n^{**}\}$, respectively.

   (a) Check whether $\sup_t |G_n(t) - G_n^*(t)|$ and $\sup_t |G_n(t) - G_n^{**}(t)|$ converge to zero almost surely as n tends to infinity. [6+4]

   (b) Let $H(\cdot)$, $H^*(\cdot)$ and $H^{**}(\cdot)$ denote the distribution of $n(\theta - M_n)/\theta$, $n(M_n - M_n^*)/M_n$ and $n(M_n - M_n^{**})/M_n$, respectively. Check whether $\sup|H(t) - H^*(t)|$ and $\sup|H(t) - H^{**}(t)|$ converge to zero almost surely as n tends to infinity. [4+4]

2. (a) Give an example of a generalized linear model with canonical like function. Show that in this case, the Fisher's scoring method for finding the maximum likelihood estimate follows the same steps as used by the Newton Raphson algorithm for finding the solution of the likelihood equation. [3+6]

   (b) Describe how the local scoring method and the backfitting algorithm can be used for fitting a generalized additive model involving multiple regressor variables. [6]

   (c) Give an example of a regression problem where additive model is not appropriate, but the projection pursuit regression model is supposed to perform well. Justify your answer. [3]

3. (a) Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be n independent copies of a d-dimensional random variable $\mathbf{X}$ following a distribution with a continuous density function f. Let $\hat{f}_{k,n}(\mathbf{x})$ be the k-nearest neighbor estimator of f based on $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$. If $k = O(n^\delta)$ for some $\delta \in (0,1)$, show that for any fixed $\mathbf{x} \in R^d$, $\hat{f}_{k,n}(\mathbf{x})$ converges to $f(\mathbf{x})$ in probability. [8]

   (b) Consider the following data set

X	1	2	3	4	5	6	7	8	9	10
Y	22.6	26.1	28.4	32.3	35.0	37.6	40.9	44.5	47.0	50.6

   Let $\hat{\psi}_h$ be the Nadaraya Watson estimate of the regression function $\psi : R \to R$ constructed based on this data set when the Gaussian kernel with bandwidth h is used.

   (i) Show that irrespective of the choice of h, $\hat{\psi}_h$ is monotonically increasing. [8]

   (ii) Find the limiting value of $\hat{\psi}_h(5.5)$ when (a) $h \to \infty$ and (b) $h \to 0$. [4+4]

P.T.O

4. (a) Consider a two-class classification problem, where the response variable Y takes two values 1 and -1. Let $\delta(\mathbf{x}) = sign(F(\mathbf{x}))$ be a classifier based on the function $F: R^d \to R$.

   (i) Show that $E[e^{-YF(\mathbf{X})}]$ gives an upper bound of the misclassification rate of δ. [3]

   (ii) Find the function F that minimizes $E[e^{-YF(\mathbf{X})}]$. [3]

   (iii) If F takes only two values β and $-\beta$, find the value of β and hence the function F that minimizes $E[e^{-Y(\alpha+F(\mathbf{X}))}]$, where α is a known constant. [6]

   (iv) Using the results in (ii) and (iii), show that the discrete Adaboost algorithm can be viewed as a method for fitting an additive logistic regression model. [4]

(b) For a J-class ($J > 2$) classification problem, define the J-dimensional vector $\mathbf{Y} = (Y_1, \ldots, Y_J)'$ with $Y_j = 1$ if the observation comes from the j-th ($j = 1, \ldots, J$) class, and $Y_j = -1/(J-1)$, otherwise. Let $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_J(\mathbf{x}))'$ be a J-dimensional function with $\sum_{j=1}^{J} f_j(\mathbf{x}) = 0$ for all $\mathbf{x} \in R^d$. Define $L(y, f(\mathbf{x})) = \exp\{-\frac{1}{J} y' f(\mathbf{x})\}$. Find the function f that minimizes $E(L(Y, f(\mathbf{X})))$ and relate them to the conditional probabilities of different classes. [6]

5. (a) Consider a set of $m+1$ points $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m$ in R^d, where $\mathbf{x}_0 = \mathbf{0}$. Let S_1 be a subset containing m_1 observations including $\mathbf{x}_0$ and S_2 be the subset of other m_2 observations ($m_1 + m_2 = m + 1$). Let C_i be convex hull formed by the observations in S_i ($i = 1, 2$).

   (i) Show that C_1 and C_2 are linearly separable if and only if the m vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are linearly independent. [5]

   (ii) Hence compute the Vapnik-Chervonenkis dimension (VC dimension) of the class of d-dimensional hyperplanes. [2]

   (iii) Give an example to show that the class of hyperplanes cannot shatter more than three points in two dimension. [3]

(b) Let $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ be independent and identically distributed d-dimensional $N(\mathbf{0}, \mathbf{I})$ vectors.

   (i) Show that $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ tend to lie on the vertices of an equilateral triangle as the dimension d tends to infinity. [4]

   (ii) Find the limiting value of angle between $\mathbf{X}_1$ and $\mathbf{X}_2$ as the dimension increases. [4]