INDIAN STATISTICAL INSTITUTE

M.Stat Second Year: First Semester, 2017-18

Full Marks: 60 | Pattern Recognition: Mid-Semestral Examination | Time : 3 Hours

[Answers should be brief and to the point. Answer as many as you can. The maximum you can score is 60]

1. Consider a classification problem between two bivariate normal distributions $N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_1)$ and $N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_2)$, where $\boldsymbol{\Sigma}_2 = 4\boldsymbol{\Sigma}_1$. Assume that the prior probabilities of the two classes are equal.

   (a) Find the Bayes risk (i.e., the average misclassification probability of the Bayes classifier). [6]

   (b) Find the average misclassification probability of the following classifier

   $$\delta(\mathbf{x}) = \begin{cases} 1 & \text{if } D(\mathbf{x}, F_1) > D(\mathbf{x}, F_2) \\ 2 & \text{otherwise,} \end{cases}$$

   where $D(\mathbf{x}, F_i)$ denotes the half-space depth of $\mathbf{x}$ with respect to the i-th distribution ($i = 1, 2$). [6]

2. Consider a classification problem where each of the two competing classes is an equal mixture of two bivariate normal distributions. While class-1 is a mixture of $N(2, 2, 1, 1, 0)$ and $N(-2, -2, 1, 1, 0)$, class-2 is a mixture of $N(-2, 2, 1, 1, 0)$ and $N(2, -2, 1, 1, 0)$. Assume that the prior probabilities of the two classes are equal. [Here $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ denotes the bivariate normal distribution with marginal means μ_1, μ_2, marginal variances σ_1^2, σ_2^2 and correlarion coefficient ρ.]

   (a) Find the Bayes classifier and its average misclassification probability. [4+4]

   (b) Compute the dispersion matrices of the two distributions. [4]

   (c) Suppose that we have large number of observations from each of the two classes. A person, who does not know the true underlying distributional structures of the two classes, assumes the underlying distributions to be normal and uses the quadratic discriminant analysis rule for classification. How will that classifier perform in this example? Give justification to your answer. [4]

3. (a) Consider a histogram density estimator $\hat{f}_{h_n}$ of a univariate density function f based on n observations and a common width h_n for all bins. Assume that $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$. If f has bounded first derivative, for any fixed x, show that $\hat{f}_{h_n}(x) \xrightarrow{P} f(x)$ as n tends to infinity. [6]

   (b) Consider a kernel discriminant analysis rule constructed based on a Gaussian kernel and a common bandwidth h for all J competing classes. Show that it behaves like the 1-nearest neighbor classifier when the bandwidth h shrinks to zero. Does it have the same behavior if different bandwidths $h_1, h_2, \ldots, h_J$ are used for J different classes and $\max\{h_1, \ldots, h_J\} \to 0$? Justify your answer. [4+4]

4. Suppose that we have following observations from two distributions.

   Class-1: $\mathbf{x}_{11}$=(0.27,-0.31), $\mathbf{x}_{12}$=(-0.43, -0.07), $\mathbf{x}_{13}$=(0.64, -0.85), $\mathbf{x}_{14}$=(0.09, -0.27), $\mathbf{x}_{15}$=(-0.13, -0.40)

   Class-2: $\mathbf{x}_{21}$=(0.27,0.31), $\mathbf{x}_{22}$=(-0.43, 0.07), $\mathbf{x}_{23}$=(0.64, 0.85), $\mathbf{x}_{24}$=(0.09, 0.27), $\mathbf{x}_{25}$=(-0.13, 0.40).

   (a) Let $MW(\boldsymbol{\alpha})$ and $KS(\boldsymbol{\alpha})$ be the value of the two-sided Mann-Whitney statistic and the two-sided Kolmogorov-Smirnov statistic computed from these observations when they are projected along the direction $\boldsymbol{\alpha} \in R^2$. Find $\sup_{\boldsymbol{\alpha}} MW(\boldsymbol{\alpha})$ and $\sup_{\boldsymbol{\alpha}} KS(\boldsymbol{\alpha})$. [4]

   (b) Define an indicator variable Y which takes the value 1 if the observation comes from the first class, and 0 otherwise. If we assume a logistic regression model for this two-class problem, show that the maximum likelihood estimate of the model parameters will not exist. [6]

(c) A person uses a kernel discriminant analysis rule based on a spherical kernel, where a common bandwidth h is used for both classes. If the prior probabilities of the two classes are equal, show that the class boundary estimated by the kernel discriminant analysis rule is linear and it does not depend on the choice of h [4]

5. Let $\mathbf{X}_n^*$ be the nearest neighbor of the origin in a set of n observations generated from a d-dimensional standard normal distribution.

(a) If d is fixed, for any fixed $\varepsilon > 0$, show that $P(\|\mathbf{X}_n^*\| \leq \varepsilon) \to 1$ as n tends to infinity. [5]

(b) If n is fixed, for any fixed $M > 0$, show that $P(\|\mathbf{X}_n^*\| \leq M) \to 0$ as d tends to infinity. [5]

Indian Statistical Institute

M.STAT - II Year

Mid-semester Examination 2017

Course name: **Introductory Economics**

Subject name: **Economics**

Date: **4 September 2017**

Maximum marks: **60**

Duration: **3 hours**

1. This question pertains to a situation in which a particular commodity is both available at a subsidised rate from a fair price shop (ration shop) and at a higher price from the open market. Suppose a consumer can buy a certain (fixed) quantity of rice at a lower price from the ration shop (that is, there is a ration quota). In addition, he can buy more of rice (assume a uniform quality of rice) from the open market at a higher price. (You may assume that consumer's preferences are represented by standard downward sloping, smooth, convex indifference curves.)

(i) Graphically depict the consumer's equilibrium and briefly describe it. **(5 points)**

(ii) What will happen to the quantity of rice purchased from the open market (over and above the ration quota) in equilibrium if there is a cut in the ration quota? Briefly explain. **(5 points)**

(iii) What will happen to the quantity purchased in the open market (over and above the ration quota) if the subsidised price (price at which the ration quota rice could be bought) is increased (but is still lower than the open market price)? Briefly explain. **(5 points)**

2. Find out the marginal revenue curve for a demand curve which is a rectangular hyperbola and graphically depict it. **(5 points)**

3. Answer "true", "false" or "uncertain" and give a defence for your answer for the following statement: "Negative income effect is sufficient but not necessary for Giffen's paradox". **(5 points)**

4. Graphically depict the kind of preferences that are represented by the following utility functions: **(1 x 5 = 5 points)**

(i) $u(x, y) = x + y$

(ii) $u(x, y) = \sqrt{x + y}$

(iii) $u(x, y) = 12x + 12y$

(iv) $u(x, y) = xy$

(v) $u(x, y) = x^2y^2$

5. Suppose Ankit says, "I like both tea and biscuits, but prefer to avoid eating them together."

(i) Draw an indifference map that illustrates the preferences of Ankit. **(3 points)**

(ii) Propose a utility function that could possibly depict the preferences you have drawn. **(3 points)**

(iii) Consider the utility function you propose in (ii). Suppose the price of a cup of tea is Rs. 5 and that of a biscuit is Rs. 2 and the money Ankit has to spend on these two goods is Rs. 30 . What will be Ankit's equilibrium choices of cups of tea and number of biscuits? **(4 points)**

6. Consider a two-good world while answering the following questions:

(i) Define and briefly describe "marginal rate of substitution" (MRS). What happens to MRS as we move downwards along a standard smooth, downward sloping, convex indifference curve? Briefly explain. **(3 points)**

(ii) "Diminishing marginal utilities of both the goods is neither necessary nor sufficient for diminishing MRS" - is this statement true? **(5 points)**

(iii) Assume that the utility function is such that the marginal utility of each of the good is independent of the units of the other good. What do you think about the statement in (ii) now? **(2 points)**

7. Suppose the consumer has a demand function for a particular kind of biscuit of the form:

$$x_1 = 5 + \frac{m}{10p_1}.$$

Initially his income is Rs.120 per week and the price of biscuit is Rs.3 per unit. Now suppose the price of biscuit falls to Rs.2 per unit. What is the *total* change in demand? Can you decompose the total change in demand into substitution and income effects? **(10 points)**

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : 2017-2018

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 5 September, 2017 Max. Marks: 60 Duration: 2 Hours

1. Compare the Bayesian approach and the classical frequentist approach in the context of estimation of a real parameter θ with a loss function $L(\theta, a)$, indicating the difference in the evaluation of performance of an estimator in these two approaches. [8]

2. Suppose that for a given set of data, for which the model involves an unknown real parameter θ, a classical 95% confidence interval and a 95% Bayesian credible interval for θ are both obtained as (2.7, 4.3). How will a frequentist and a Bayesian interpret this result? [5]

3. Let X_1, X_2 be i.i.d. with a common density belonging to a location parameter family of densities with a location parameter θ. Assume without loss of generality that $E_\theta X_1 = \theta$. One can find a frequentist 95% confidence interval of the form $(\bar{X} - c, \bar{X} + c)$. Suppose now that $X_1 - X_2$ is known and one calculates P(the interval $\bar{X} \mp c$ covers θ $|X_1 - X_2)$. What is Welch's paradox in this context? Give an example where Welch's paradox occur.

Can Welch's paradox occur if X_1, X_2 are i.i.d. $N(\theta, 1)$? (Explain.)

[5+3=8]

4. Let $X_1, \ldots, X_n$ be i.i.d. $N(\theta, \sigma^2)$ variables.

(a) Consider a standard noninformative prior for (θ, σ^2) and find the corresponding $100(1-\alpha)\%$ HPD credible set for θ.

(b) Assume that σ^2 is known and consider a conjugate prior for θ. Find the posterior distribution of θ and the posterior predictive distribution of a future observation X_{n+1}. [(7)+(4+7)=18]

5. 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 $\mu_1 - \mu_2$.

[16]

6. Let $X_1, \ldots, X_n$ be i.i.d. with a common density $f(x|\theta)$ where $\theta \in R$. State the result on asymptotic normality of posterior distribution of suitably normalized and centered θ under suitable conditions (to be stated by you) on the density $f(\cdot|\theta)$ and the prior distribution. [5]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: 2017-18 (First Semester)

M. MATH. II YEAR
Commutative Algebra I

Date : 6.09.17 | Maximum Marks : 30 | Duration : 2 Hours

GROUP A
ATTEMPT ANY TWO QUESTIONS.
Each question carries 7 marks.
R denotes a commutative ring with unity.

1. Prove that any finitely generated projective module over a local ring is free.
2. Let I be an ideal of R for which R/I is a projective R-module. Show that there exists an element $e \in I$ for which $e^2 = e$ and $I = Re$.
3. Let M be an R-module and $f_i : P_i \to M$ be surjective R-linear maps from projective R-modules P_i to M, $i = 1, 2$. Let K_i denote the kernel of f_i. Prove that $P_1 \oplus K_2$ is isomorphic to the fibre product $P_1 \times_M P_2 = \{(x, y) \in P_1 \oplus P_2 \mid f_1(x) = f_2(y)\}$.

GROUP B
ATTEMPT ANY THREE QUESTIONS.
Each question carries 6 marks.

1. Prove that any flat module over an integral domain is torsion-free.
2. Prove that the ring $\mathbb{Z}[X]/(2X - 3)$ is a PID.
3. Find a multiplicatively closed set S in the ring $B = \mathbb{C}[X, Y]/(XY)$ for which $S^{-1}B$ is an integral domain.
4. Find an element f in the ring $A = \mathbb{C}[X, Y, Z]/(XY - Z^2)$ such that $A/fA \cong \mathbb{C}[T, 1/T]$.
5. Let B be a subring of an integral domain A. Suppose that there exists an element t in B for which
(i) $B[1/t] = A[1/t]$ and
(ii) the induced ring homomorphism $B/tB \to A/tA$ is injective.
Prove that $B = A$.

INDIAN STATISTICAL INSTITUTE

M-Stat. (2nd year) 2017-18

MID-SEMESTRAL EXAMINATION

Subject: Time Series Analysis

Date: 06.09.2017 Full Marks: 40 Duration: 1 hour 30 minutes

Attempt **ALL** questions

1. Consider the equation for AR(1) process.
   $X_t = \emptyset X_{t-1} + Z_t,\ |\emptyset| \neq 1, \{Z_t\} \sim W\ N\ (0, \sigma^2)$
   a) For different values of $\emptyset$, write $\{X_t\}$ as linear processes in $\{Z_t\}$.
   b) For $|\emptyset| < 1$, show that $\{X_t\}$ is causal and show that the partial sums of the process converge to $\{X_t\}$
   (i) in L_2, (ii) in L_1, and (iii) almost everywhere.

[7 + 7]

2. Check if the following functions from Z ⟶ R are non-negative definite or not.
   a) $\cos wt\ (1 + \cos wt)$
   b) $\sin wt\ (1 + \sin wt)$
   Where $w \in \left(0, \frac{\pi}{2}\right), t \in Z$

[6 + 3]

3. a) Discuss why sometimes the product model is preferred in time series to the additive model.

   b) Discuss how seasonal component is estimated/ eliminated from the time series.

   c) Find all possible $(a, b, c) \in R^3$, such that the filter $aB^{-1} + b\ I + c\ B$ when applied on time series

      (i) removes any seasonal component of order 3 and allows to pass undistorted any linear trend

      (ii) removes any seasonal component of order 2 and allows to pass undistorted any linear trend.

   Justify your answer. B = backward shift operator

[2 + 5 + (5 + 5)]

******** ********

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: (2017–2018)

M. Stat 2nd Year

Statistical Computing

Date: 07/09/17 Marks: 30 Duration: 2 hours

Attempt all questions

1. Let the lifetimes of electric light bulbs of a certain type have uniform distribution in the interval $(0, \theta]$, where $\theta > 0$ is unknown. A total of $n + m$ bulbs are tested in two independent experiments. The observed data consists of $\mathbf{y} = (y_1, \ldots, y_n)'$ and $\mathbf{z}^* = (z^*_{n+1}, \ldots, z^*_{n+m})'$, where $\mathbf{y}$ are exact lifetimes of a random sample of n bulbs, and $\mathbf{z}^*$ are indicator observations on a random sample of m bulbs; that is, for $i = n + 1, \ldots, n + m$,

$$\begin{aligned} z_i^* &= 1 \text{ if bulb } i \text{ is still burning at a fixed time point } T > 0 \\ &= 0 \text{ if expired.} \end{aligned}$$

Let $\mathbf{z} = (z_1, \ldots, z_m)' = (y_{n+1}, \ldots, y_{n+m})'$ denote the missing data. Also, let $s = \sum_{i=n+1}^{n+m} z_i^* \geq 1$.

(a) Then directly obtain the maximum likelihood estimator (MLE) of θ.

(b) Derive an EM algorithm to obtain the MLE of θ.

(c) Does the result of the EM algorithm match the directly estimated MLE? Justify your answer.

Marks: 2+3+5=10

2. (i) For a sample $x_1, \ldots, x_n$ of real numbers and $q \in (0, 1)$, consider minimizing the function

$$f(\theta) = \sum_{i=1}^{n} \rho_q(x_i - \theta),$$

where

$$\rho_q(\theta) = \begin{cases} q\theta, & \theta \geq 0; \\ -(1-q)\theta, & \theta < 0. \end{cases}$$

Derive an MM algorithm for the minimization purpose, with suitable approximation such that the algorithm remains well-defined.

(ii) Consider the function

$$f(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2.$$

(a) Show that the function

$$g(x|x_n) = \frac{1}{4}x^4 + \frac{1}{2}x_n^2 - xx_n$$

majorizes $f(x)$ at x_n.

(b) Prove that your resulting MM algorithm does not find all the minimizers.

(c) Modify your MM algorithm such that the modified (perhaps, non-MM) algorithm leads to the same value of $f(\cdot)$ at each iteration as MM and has subsequences converging to the minimizers.

Marks: 2+2+6=10

3. For nodes $x_0 < x_1 < \cdots < x_n$ and function values $f_i = f(x_i)$, develop a quadratic interpolating spline $s(x)$ satisfying

(a) $s(x)$ is a quadratic polynomial on each interval $[x_i, x_{i+1}]$,

(b) $s(x_i) = f_i$ at each node x_i,

(c) the first derivative $s'(x)$ exists and is continuous throughout the entire interval $[x_0, x_n]$.

Do you require any additional information to completely determine the spline? Justify your answer. **Marks**: 10

Date: September 8, 2017
Time: 2 hours

Statistical Genomics
M-Stat (2nd Year)
Mid Semester Examination 2017-18

The paper carries 40 marks. This is an open notes examination.
Answer all questions.

1. Explain whether it is possible that each of two independent random samples of individuals conform to Hardy-Weinberg Equilibrium at a locus, but exhibit significant evidence of deviation from HWE when the two samples are combined. [10]

2. Consider two autosomal biallelic loci with alleles (*D,d*) and (*M,m*) respectively. The following are the genotype data on two nuclear families:

Family 1: father is *DM/dm*, mother is *Ddmm*, offspring are *DDMm*, *DdMm* and *DDmm*

Family 2: both parents are *Dm/dM*, offspring are *DdMM*, *DDmm* and *ddMm*

Test using the LOD score approach whether the two loci are linked. [12]

3. What is the minimum expected i.b.d. score of a pair of sibs, both heterozygous at an autosomal biallelic locus? [6]

4. The Kimura two-paremeter model is an extension of the Jukes-Cantor one-parameter model assuming that the rate of transitions is different from the rate of transversions. Explain whether a twelve-parameter model would be a significant improvement? [2]

5. Consider the alignment of an ancestral sequence S_0 and a descendent sequence S_1:

S_0: GGCAGTCGAAAAATCACACGGTTCCAACCTCCCGGCATA
S_1: GGGAGACAGATAATCCCACGGTTCTAGCTCCCCGGGATT

(a) Compute Jukes-Cantor distance between S_0 and S_1.
(b) What would the transition matrix be for Kimura two-parameter model?
(c) Compute the Kimura two-parameter distance between S_0 and S_1.
(d) Which of the above two distances is likely to be a more reasonable measure? Justify your answer. [2+2+4+2]

INDIAN STATISTICAL INSTITUTE

Mid-Semester of First Semester Examination : 2017 - 18

Course Name : M. Stat. II Year

Subject Name : Signal and Image Processing

Date: 11.09.2017 Full Marks : 50 Duration : 2 hours

Part - I

Answer any FIVE questions: $5 \times 5 = 25$

1. Show that the fundamental period N_p of the signals $s_k(n) = e^{j2\pi kn/N}$ for $k = 0, 1, 2, \ldots$ 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 24? 3+2 = 5

2. Consider the system described by the difference equation:

$$y(n) = ay(n-1) + bx(n).$$

Suppose the unit impulse response of the system is $h(n)$. Determine b in terms of a such that

$$\sum_{n=-\infty}^{\infty} h(n) = 1.$$

3. The step response of an LTI system is:

$$s(n) = \left(\tfrac{1}{3}\right)^{n-2} .u(n+2).$$

Find the system function $H(z)$ and sketch the pole-zero plot. Is the system stable? 4+1 = 5

4. By using appropriate properties of z-transforms find the corresponding $x(n)$ in case of the following relations:

$$X(z) = \log(1 - z^{-1}),\ \ |z| > \frac{1}{2}.$$

5. Determine the autocorrelation sequences of the following signals:

$$x(n) = \{\underset{\uparrow}{1},2,1,1\} \text{ and } y(n) = \{\underset{\uparrow}{1},1,2,1\}.$$

Is there any relation between the autocorrelation sequences that you observe? 2+2+1 = 5

6. When the input to an LTI system is $x(n) = \left(\tfrac{1}{2}\right)^n .u(n) + 2^n .u(-n-1)$ and the output is:

$y(n) = 6\left(\tfrac{1}{2}\right)^n .u(n) - 6.\left(\tfrac{3}{4}\right)^n .u(n)$, derive the difference equation that characterizes the system.

P.T.O

7. Derive and sketch the direct form I and direct form II realizations of the LTI system characterized by the following difference equation:

$$y(n) - 4y(n-1) - 2y(n-2) = 3x(n) + 2x(n-1) - 0.5x(n-2).$$

Part - II

1. Describe an algorithm for finding connected components (8-connected) in a binary image. [10]

2. Apply histogram equalization scheme on the following frequency distribution table of 8 gray values and show the output. [7]

Gray value	Frequency
0	1
1	20
2	2
3	30
4	10
5	20
6	10
7	7

3. Describe non-maxima suppression scheme in Canny edge detection method. [8]

MARTINGALE THEORY

Mid-Semestral Examination

M.Stat Second Year, 2017-18

Date: 11/09/2017 Maximum Marks: 30

Duration: 2 hours

Anybody caught using unfair means will immediately get 0. Please try to explain every step. NO NOTES ARE ALLOWED.

(1) Let $(\Omega, \mathcal{A}, P)$ be a probability space and $(\mathcal{F}_n)_{n\geq 0}$ be a filtration. Let $(X_n)_{n\geq 0}$ be an adapted process such that $X_n \in L^1$ for all $n \geq 0$. Show that $(X_n)_{n\geq 0}$ is a martingale if and only if $\mathsf{E}[X_T] = \mathsf{E}[X_0]$ for any bounded stopping time T. [10 marks]

(2) Let μ_1, μ_2, μ_3 be three probability measures on $(\Omega, \mathcal{F})$ such that $\mu_1 \ll \mu_2 \ll \mu_3$. Show that

$$\frac{d\mu_1}{d\mu_3} = \frac{d\mu_1}{d\mu_2} \times \frac{d\mu_2}{d\mu_3} \qquad a.s.$$

[4 marks]

(3) Consider an urn having black and white balls, the initial proportion of white balls being X_0. A ball is drawn at random and replaced. A ball of the same color is added to the urn. Let X_n be the proportion of white balls after n draws.

(a) Show that there exists X_∞ such that $X_n \to X_\infty$ as $n \to \infty$ almost surely and in L^1.

(b) Show that $0 \leq X_\infty \leq 1$ and $\mathsf{E}[X_\infty] = \mathsf{E}[X_0]$.

(c) When $0 < X_0 < 1$ then show that X_∞ is non-degenerate.

[2+2+2=6 marks]

(4) Let $(X_i)_{i\geq 1}$ be iid with $P(X_1 = +1) = 1/3$ and $P(X_1 = -1) = 2/3$ and $S_n = \sum_{i=1}^n X_i$. Find the probability that S_n hits b before $-a$ (with $a, b > 0$). [10 marks]

INDIAN STATISTICAL INSTITUTE, KOLKATA
MIDTERM EXAMINATION: FIRST SEMESTER 2017 - '18
M.STAT II YEAR

Subject	:	**Functional Analysis**
Time	:	2 hours 30 minutes
Maximum score	:	40

Attempt all the problems. Please use a new page to answer each question, making sure that the question number in the margin can be read, even after stapling. If you attempt the same problem several times, please strike out all the attempts except the final one before submitting your answer script. Justify every step in order to get full credit of your answers. All arguments should be clearly mentioned on the answer script. Points will be deducted for missing or incomplete arguments.

(1) Let $X = C^1([0,1])$ be the vector space of all complex valued functions on $[0,1]$ having continuous first order derivatives. Prove that $\|f\|_1 := \|f\|_\infty + \|f'\|_\infty$ defines a norm with respect to which X is complete . Show that $\|f\|_2 := |f(0)| + \|f'\|_\infty$ is another norm on X that is equivalent to the above norm.

[5+5 = 10 marks]

(2) (a) Prove that $\ell^\infty(\mathbb{N})$ is not separable. (Hint: Consider S to be the set of sequences all whose entries are either 0 or 1. What is the distance between two distinct elements of S?)
(b) Let x, y be vectors in S. Prove that if $x_k \neq y_k$ for at most finitely many k, then $x + c_0 = y + c_0$. Prove that if $x_k \neq y_k$ for infinitely many k, then $\|x - y + c_0\| = 1$.
(c) Conclude that ℓ^∞ / c_0 is not separable.

[3+4+3 = 10 marks]

(3) Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Show that there is a constant $c > 0$ such that $\|T(x)\| \geqslant c\|x\|$ for all $x \in X$ if and only if T is 1-1 and range of T is closed in Y.

[7 marks]

(4) Let $T : \ell^p(\mathbb{N}) \to \ell^p(\mathbb{N})$ (where $1 \leq p < \infty$) be the right shift operator; i.e., $T(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$. What is the adjoint of T?

[5 marks]

(5) Prove that $\ell^1(\mathbb{N})$ is not reflexive, using the following steps:
(i) For each $n \in \mathbb{N}$, define $\mu_n : \ell^\infty \to \mathbb{C}$ by

$$\mu_n(x) = \frac{x_1 + x_2 + \cdots + x_n}{n} \text{ for } x = (x_1, x_2, \ldots) \in \ell^\infty$$

Show that $\mu_n \in (\ell^\infty(\mathbb{N}))^*$ for every n and in fact, $\|\mu_n\| \leq 1$
(ii) Show that there is an element $\mu \in (\ell^\infty(\mathbb{N}))^*$ which is an accumulation point of the sequence μ_n.
(iii) Show that $\mu \neq \hat{x}$ for any $x \in \ell^1(\mathbb{N})$, where $\hat{x}$ is the image of x under the natural embedding of $\ell^1(\mathbb{N})$ into $(\ell^\infty(\mathbb{N}))^* \cong \ell^1(\mathbb{N})^{**}$.

[3+4+5 = 12 marks]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2017 – 18

MStat (2nd Year)

Quantitative Finance

Date: 12 September 2017 Maximum Marks: 30 Duration: 2 Hours

1. Critically explain the concepts: [2½ X 4 = 10]
    a) Mutual Fund Principle
    b) Law of One Price
    c) Barrier Options
    d) Reflection Principle

2. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \ln w$. 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 $P(\omega_1) = 3/5$, $Q(\omega_1) = 1/2$.

   [8 + 4 = 12]

3. 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. [5 + 3 = 8]

Brownian Motion and Diffusions

Compensatory exam

Time: 3 hours

Date: 11/10/2017

Answer **any five** of the following questions. Each question carries 20 marks.

In all the following questions, $(B_t : t \geq 0)$ is a standard Brownian motion, unless mentioned otherwise.

1. Calculate
$$P\left(\max_{0\leq t\leq 1} B_t \geq 1, B_1 \leq 0\right).$$

2. Let $(S_n : n \geq 0)$ be a simple symmetric random walk starting from zero. Show that there exists $\beta \in (0, \infty)$ such that as $n \to \infty$,
$$P\left(n^{-\beta} \max_{0\leq i\leq 2n} S_i \in \cdot \,\middle|\, S_{2n} = 0\right) \Rightarrow \mu(\cdot),$$
weakly, for some non-degenerate probability measure μ on $\mathbb{R}$.

3. Let
$$T_n := \min\{t \geq 0 : B_t = n\}, \, n \geq 1.$$
Show that for a fixed $n \geq 2$,
$$T_n \stackrel{d}{=} Z_1 + \ldots + Z_n,$$
where $Z_1, \ldots, Z_n$ are i.i.d. copies of T_1.

4. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a filtered probability space supporting a standard Brownian motion $(B_t)_{t\geq 0}$ with natural filtration $(\mathcal{F}_t)_{t\geq 0}$. Let $(h_t)_{t\in[0,T]}$ be an adapted bounded process and define the process $(Z_t)_{t\in[0,T]}$ as the unique solution to $dZ_t = -h_t Z_t dB_t$, starting from $Z_0 = 1$. For any $t \geq 0$, define the restriction to $\mathcal{F}_t$, a measure $\mathbf{Q}$,
$$d\mathbf{Q}_{|\mathcal{F}_t} := Z_t d\mathbf{P}_{|\mathcal{F}_t}.$$
Prove that
$$\mathrm{E}_{\mathbf{P}}\left[Z_t \log(Z_t)\right] = \mathrm{E}_{\mathbf{Q}}\left[\frac{1}{2}\int_0^T h_s^2 ds\right].$$

5. Let

$$X_t = \cos(B_t),\ t \geq 0\,.$$

(a) Show that there exists a function $f : [0, \infty) \to (0, \infty)$ such that $f(t)X_t$ is a martingale with respect to the canonical filtration.

(b) Show that there exists a process $(Z_t : t \geq 0)$ with differentiable paths such that $X_t - Z_t$ is a martingale with respect to the canonical filtration.

6. Let Z be a random set defined by

$$Z = \{0 \leq t < \infty : B_t = 0\}\,.$$

Show that almost surely, the set Z

(a) has Lebesgue measure zero,

(b) is closed and unbounded,

(c) has an accumulation point at $t = 0$,

(d) and has no isolated points in $(0, \infty)$.

7. Show that there exists $\beta \in \mathbb{R}$ such that as $n \to \infty$,

$$n^{\beta} \frac{\sum_{i=1}^{n} (B_{i/n} - B_{(i-1)/n})^3}{\sum_{i=1}^{n} (B_{i/n} - B_{(i-1)/n})^2} \Rightarrow Z\,,$$

for some random variable Z. Find out the distribution of Z.

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2017–2018)

M. Stat 2nd Year

Statistical Computing

Date: 20/11/17 Full Marks: 50 Duration: 3 hours

Attempt all questions

1. (i) Let $X_1, \ldots, X_m$ be an *iid* sample from a normal density with mean μ and variance σ^2. Suppose for each X_i we observe $Y_i = |X_i|$ rather than X_i. Formulate an EM algorithm for estimating μ and σ^2.

(ii) Let

$$X_1, \ldots, X_n \sim f(x \mid \theta) = \frac{1}{\pi\{1 + (x-\theta)^2\}}.$$

Derive an EM algorithm to obtain the MLE of θ. (*Hint: You may use the fact that the ratio of two independent standard normal random variables have the standard Cauchy distribution.*)

[5+5=10]

2. (i) Show that the EM algorithm is a special case of the MM algorithm.

(ii) In the context of maximum likelihood estimation in multinomial distribution show that your MM algorithm converges to the maximum likelihood estimate at a linear rate.

(iii) Develop an MM algorithm for minimizing the function

$$f(x_1, x_2) = \frac{1}{x_1^3} + \frac{3}{x_1 x_2^2} + x_1 x_2,$$

where $x_1, x_2 > 0$.

[2+3+5=10]

3. Let V be the collection of functions f with $f'' \in L^2[0,1]$. Consider the subspace

$$W_2^0 = \{f(x) \in V : f, f' \text{ absolutely continuous and } f(0) = f'(0) = 0\}.$$

Define an inner product on W_2^0 as

$$< f, g >= \int_0^1 f''(t) g''(t) dt.$$

(i) Show that for any $f \in W_2^0$, and for any s, $f(s)$ can be written as

$$f(s) = \int_0^1 (s-u)_+ f''(u) du,$$

where $(a)_+$ is a for $a > 0$ and 0 for $a \leq 0$.

(ii) Hence, obtain the reproducing kernel of W_2^0.

[5+5=10]

4. (i) Let $\mathbf{X}$ be a random vector with density f on $\mathbb{R}^d$, and let U be an independent uniform $[0,1]$ random variable.

(a) Prove that $(\mathbf{X}, cUf(\mathbf{X}))$ is uniformly distributed on $\mathbf{A} = \{(\mathbf{x}, u) : \mathbf{x} \in \mathbb{R}^d, 0 \leq u \leq cf(\mathbf{x})\}$, where $c > 0$ is an arbitrary constant.

(b) Prove the converse: if $(\mathbf{X}, U)$ is a random vector in $\mathbb{R}^{d+1}$ uniformly distributed on $\mathbf{A}$, then $\mathbf{X}$ has density f on $\mathbb{R}^d$.

(ii) Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be a sequence of *iid* random vectors taking values in $\mathbb{R}^d$, and let $\mathbf{A} \subseteq \mathbb{R}^d$ be a Borel set such that $P(\mathbf{X}_1 \in \mathbf{A}) = p > 0$. Let $\mathbf{Y}$ be the first $\mathbf{X}_i$ taking values in $\mathbf{A}$. Then prove the following:

(a) $\mathbf{Y}$ has a distribution that is determined by

$$P(\mathbf{Y} \in \mathbf{B}) = \frac{P(\mathbf{X}_1 \in \mathbf{A} \cap \mathbf{B})}{p}; \ \mathbf{B} \text{ for any Borel set } \mathbf{B} \text{ of } \mathbb{R}^d \ \mathbb{R}^d.$$

(b) If $\mathbf{X}_1$ is uniformly distributed in $\mathbf{A}_0$, where $\mathbf{A}_0 \supseteq \mathbf{A}$, then $\mathbf{Y}$ is uniformly distributed in $\mathbf{A}$.

(iii) The basic version of the rejection algorithm assumes the existence of a density g and the knowledge of a constant $c \geq 1$ such that, for all $\mathbf{x}$,

$$f(\mathbf{x}) \leq cg(\mathbf{x})$$

Using results proved in 4. (i) and 4. (ii) show that random variates with density f on $\mathbb{R}^d$ can be obtained as follows:

REPEAT

(a) Generate two independent random variates $\mathbf{X}$ (with density g on $\mathbb{R}^d$) and U (uniformly distributed on $[0,1]$).

(b) Set $T \longleftarrow c\frac{g(\mathbf{X})}{f(\mathbf{X})}$.

UNTIL $UT \leq 1$
RETURN $\mathbf{X}$.

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

5. (i) Assume that the Markov transition kernel $P(x, A)$ satisfies a minorization condition on a small set C, in addition to the regularity conditions necessary for convergence to the target distribution π. Assume that two different chains are run, one started from an arbitrary fixed starting value x_0, and another started by drawing the initial value y_0 from the invariant distribution.

(a) Derive the coupling inequality

$$\| P^n(x_0, \cdot) - \pi(\cdot) \| = \sup_{A \in \mathcal{B}} |P^n(x_0, A) - \pi(A)| \leq Pr(T > n)$$

where $P^n(x_0, A)$ is the probability of hitting the set A in n iterations, beginning at x_0, T is the time at which coupling occurs and $\mathcal{B}$ is the appropriate Borel σ-algebra.

(b) Using the coupling inequality, show that if the entire state space is small, then the Markov chain is uniformly ergodic.

(ii) Suppose that a posterior distribution is given by

$$\pi(x) \propto \frac{1}{1+x^2}; \quad |x| \leq k,$$

where k is a known constant.

Consider a Metropolis-Hastings algorithm for sampling from the above distribution, where the proposal distribution $q_{\sigma^2}(x^{(c)}, \cdot)$ ($x^{(c)}$ is the current value of the Markov chain) is a normal distribution with mean $x^{(c)}$ and known variance σ^2.

Prove that the Markov chain is geometrically ergodic.

[(3+2)+5=10]

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2017-2018

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 22 November, 2017 Max. Marks: 100 Duration: 3 Hours

Answer all questions

1. Consider the linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{y} = (y_1, \ldots, y_n)'$ is the vector of observations on the "dependent" variable, $\boldsymbol{X} = ((x_{ij}))_{n\times p}$ is of full rank, x_{ij} being the values of the nonstochastic regressor variables, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ is the vector of regression coefficients and the components of $\boldsymbol{\epsilon}$ are independent, each following $N(0, \sigma^2)$. Consider the noninformative prior $\pi(\boldsymbol{\beta}, \sigma^2) \propto \frac{1}{\sigma^2}$, $\boldsymbol{\beta} \in R^p$, $\sigma^2 > 0$. Find the following:

(a) The marginal posterior distribution of $\boldsymbol{\beta}$.
(b) The marginal posterior distribution of σ^2.
(c) The $100(1-\alpha)\%$ HPD credible set for $\boldsymbol{\beta}$.

[6+5+12=23]

2. (a) What are the difficulties with improper noninformative priors in Bayes testing? Describe the intrinsic Bayes factor (IBF) as a solution to this problem with improper priors.

(b) What is an intrinsic prior in the context of nonsubjective Bayes testing? Consider the general nested case and find the intrinsic prior determining equations corresponding to AIBF. Show that the solution suggested by Berger and Pericchi satisfies the intrinsic prior determining equations.

[(3+7)+15=25]

3. Consider p independent random samples, each of size n, from p normal populations $N(\theta_i, \sigma^2)$, $i = 1, \ldots, p$. Assume σ^2 to be known. Also assume that $\theta_1, \ldots, \theta_p$ are i.i.d. $N(\mu, \tau^2)$. Our problem is to estimate $\theta_1, \ldots, \theta_p$.

(a) A natural estimate of $(\theta_1, \ldots, \theta_p)$ is the vector of sample means. Why a suitable shrinkage estimate is expected to perform better than this estimate?

P.T.O

(b) Describe the Hierarchical Bayes and the parametric empirical Bayes (PEB) approaches in this context. Derive the James-Stein estimate as a PEB estimate.

[4+(15+7)=26]

4. Consider the set up of Question (3) with σ^2 unknown. Assume that σ^2 follows Inverse-Gamma (a_1, b_1) and is independent of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$. Consider the second stage priors:

$$\mu \sim N(\mu_0, \sigma_0^2) \text{ and } \tau^2 \sim \text{ Inverse-Gamma } (a_2, b_2)$$

where μ and τ^2 are independent. Assume that $a_1, b_1, a_2, b_2, \mu_0$ and σ_0^2 are specified constants.

Describe how Gibbs sampling can be used to sample from the posterior distribution of $\boldsymbol{\theta}$. Find only the conditional posterior distribution of σ^2 given $\boldsymbol{\theta}, \mu, \tau^2$ and that of μ given $\boldsymbol{\theta}, \sigma^2, \tau^2$.

[16]

5. Suppose we have observations $X_1, \ldots, X_n$. Consider the model selection problem with two candidate models M_0 and M_1. Under M_0, X_i are i.i.d. $N(0,1)$ and under M_1, X_i are i.i.d. $N(\theta, 1)$, $\theta \in R$. As there are difficulties with improper noninformative prior, one may like to use a conjugate $N(0, \tau^2)$ prior for θ under M_1 where τ^2 is very large. Is this a reasonable specification of prior for this problem? Justify your answer.

[10]

Indian Statistical Institute

Final Examination 2017

M. Stat. IInd year

Course name: **Introductory Economics**

Subject name: **Economics**

Date: 23.11.2017

Maximum marks: **80**

Duration: **3 hours**

1. Compute the cost function for the production function given by

$$f(x_1, x_2, x_3) = x_1 + \min\{x_2, x_3\},$$

where x_1, x_2, x_3 are the inputs used in producing the output. **(15 points)**

2. Suppose trade unions are successful in bargaining and increasing the salaries of engineers working at ONGC (Oil and Natural Gas Corporation) India. Does the share of labour income relative to capital income at ONGC, necessarily rise? Suppose, now, trade unions bargain and increase wage rate of workers working in a textile industry. Does the distribution of income move in favour of labour in this case? In general, trade unions often demand increase in wage rate of the workers presumably "because it leads to distribution of income moving in favour of labour." Do you agree with such a position? Why or why not? **(5+5+5=15 points)**

3. Let production of contracts be given by $Q = E^{1/2}L^{1/2}$ where Q is the number of contracts, E is the number of hours put in by economists and L is the number of hours put in by lawyers. Assume input prices are Rs. 4/hour for the use of economists and Rs. 1/hour for the use of lawyers. **(5x5=25 points)**

(a) What is the long-run cost-minimising number of hours of economist and lawyer time the firm will employ to produce 1 contract, 2 contracts, 3 contracts, 4 contracts and 5 contracts?

(b) Calculate the total cost, average cost, and marginal cost, for each of the five contract levels. (Note that these are long run costs.)

(c) Assume that the law firm has already hired on contract four hours of lawyer time and cannot alter this amount in the short run. Given this amount of lawyer time, what is the amount of economist time to be used to produce contract levels of 1, 2, 3, 4, 5?

P.T.O

(d) Calculate the short run total cost, average cost and marginal cost for each contract level.

(e) Can the total cost and average cost in (d) ever be less than the total cost and average cost determined in (b)? Explain briefly.

4. A discriminating monopolist faces two demand curves that have the same intercepts. What can you say about the prices he charges in the two markets? Now assume that the demand curves have the same slope but different intercepts. Now what can you say about the prices he charges in the two markets? **(5+5=10 points)**

5. Graphically depict a firm earning normal profits under monopoly and a firm earning normal profits under perfect competition. What do you think is the main difference between the two situations? **(15 points)**

INDIAN STATISTICAL INSTITUTE

M-Stat. (2nd year) 2017-18

SEMESTRAL EXAMINATION

Subject: Time Series Analysis

Date: 24.11.2017 Full Marks: 60 Duration: 3hrs

You are allowed to use calculator if needed

Attempt **ALL** questions

1. Let $\{X_t\}$ be a stationary time series with ACF $\rho(k)$, k= 0, 1, 2,…….
a) For $\underline{\rho}_k = (\rho(1), \rho(2),\ldots\ldots, \rho(k))$, state asymptomatic distribution of $\underline{\rho}_k$ using Bartlett`s formula. (Derivation not required).
b) The following is a series of 16 random digits (from 0 to 9)
4,6,5,2,9,0,3,1,2,0,3,0,0,6,4,1
Using these construct an MA (1) process $(with\ the\ constant\ \theta = 0.5), X_t, t = 1(1)\ 15$. Calculate $\hat{\rho}(1), \hat{\rho}(2), \hat{\rho}(3)$.
c) Using Bartlett`s formula and above MA(1) process calculate the asymptotic distributions of $\hat{\rho}(1), \hat{\rho}(2), \hat{\rho}(3)$
d) Hence check how much the values of X_t , t=1(1)15, support that it is from an MA(1) process [Given $P(X < 1.96) = 0.975, when\ X \sim N\ (0,1)$]

[5+7+3+3]

2. a) Consider a moving average process (assuming $\{Z_t\}$ follows WN $(0, \sigma^2)$)
$X_t = \sum_0^\infty \psi_i Z_{t-i}$ with $\Sigma|\psi_i| < \infty$ Find the spectral density of X_t.
b) Using this derive the spectral density of ARMA(p, q) process which is assumed to be causal.
c) Prove that if X_t is an ARMA(p, q) process which may not be causal or, invertible, there always exists a different representation of X_t with respect to some other white noise, for which it would be causal and invertible [write down what assumption we need]

[4+4+4]

3. a) Let $\{X_t\}$ be a stationary time series. Find $P_n(X_{n+h})$, best linear predictor of X_{n+h} (h>0) using $X_n, X_{n-1}, \ldots\ldots, X_2, X_1, 1$
b) Show that in the above prediction, Error is uncorrelated to the predictor variables.
c) Find the expression for expected squared Error.
d) Show that $P_n(X_{n+h})$ is unique.

[5+2+3+4]

b) Following gives equation which are used to define ARMA (p, q) processes.
a) $(1+B+B^2)\ X_t = (1-0.5B^2)\ Z_t$
b) $(1-0.7B+0.1B^2)\ X_t = (1+0.4B^5)Z_t$
Check which one gives a stationary process and which one is causal/ invertible.

[4+4]

c) Find the set of values of ρ for which the following defines an ACVF of some stationary time series $\{X_t\}$.
a) $\Upsilon(t) = 1$ if $t = 0$
$= \rho$ if $|t| = 1$
$= 0$ otherwise

b) $\Upsilon(t) = 1$ if $t=0$
$= \rho$ if $|t|=2$
$= 0.5$ if $|t| = 4$
$= 0$ otherwise

[3+5]

****

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2017 – 18

MStat (2nd Year)

Quantitative Finance

Date: 27 November 2017 Maximum Marks: 100 Duration: 3 Hours

1. Define the following: [6 X 3 = 18]
   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) The discounted gain, $G^* = \sum_{n=1}^{N} H_n S_n^*(0) R_n^*$

(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.

[6+6 + 8 = 20]

3. Consider a two period model with $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and asset with time 0 price $A_0 = 100$.

(a) If the bank process is deterministic with per period interest rate $r = 0.05$ and the asset can be sold short, then what is the time 0 forward price O_0 of the asset for delivery at time 2?

(b) If the bank process is random with $B_1 = 1.05, B_2(\omega_1) = B_2(\omega_2) = 1.12$ and $B_2(\omega_3) = B_2(\omega_4) = 1.1$, and if the asset can be sold short, then what is O_0 now? Give an expression in terms of the risk neutral probability $Q(\{\omega_1, \omega_2\})$.

(c) If the bank process is as in (b) but **short selling is not allowed,** then what is the largest value of O_0 consistent with no-arbitrage?

(d) What happens if in (c), the asset has a carrying cost of 5 per period?

[5 + 5 + 5 + 5 = 20]

4. Prove directly from the definition of Ito integrals that

$$\int_0^t B_s^2 dB_s = \frac{1}{3} B_t^3 - \int_0^t B_s ds,$$

where $\{B_t\}$ is the standard Brownian motion (or b.m.) [15]

5. Prove that $M_t = B_t^2 - t$ is a martingale where $\{B_t\}$ is a b.m. [12]

6. Let $x > 0$ be a constant and define $X_t = (x^{1/3} + \frac{1}{3} B_t)^3;\ t \geq 0$, where $\{B_t\}$ is a b.m. Show that $dX_t = \frac{1}{3} X_t^{1/3} dt + X_t^{2/3} dB_t; X_0 = x$. [15]

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2017-18

M Stat (2^{nd} Year)

Statistical Genomics

Date: November 27, 2017 | Maximum marks: 60 | Duration: 3 hours

The paper carries 65 marks. The maximum you can score is 60.
Answer all questions.

1. If the exact identity-by-descent scores at a marker locus are available for a random set of affected first cousin-pairs, develop a mean allele-sharing test for linkage between the marker locus and the disease locus. [6]

2. Given genotype data at two biallelic loci for an unrelated set of individuals, describe a suitable algorithm to estimate the coefficient of linkage disequilibrium between the loci. [10]

3. Show that the classical TDT is a test for both linkage and association under a dominant disease model. Explain whether the power of the TDT is susceptible to the presence of population stratification. [8 + 4]

4. Consider data on a quantitative trait on sib-pairs and genotypes at a marker locus for both parents as well as the sibs. Assuming the classical Haseman-Elston framework with no dominance, show that the correlation of the sib-pair trait values conditioned on their i.b.d. scores at the marker locus is a linear function of the i.b.d. scores. How would you test for linkage between a QTL and a marker locus based on the above property? [5+2]

5. (a) Why is raw sequence similarity an underestimated measure of evolutionary distance?
(b) Suppose in a BLAST search you got an E-value of about 2×10^{-5}. What does this E value mean? Name two parameters that determine the E-value. [4+3]

6. (a) What are the advantages and disadvantages of UPGMA and NJ method for phylogenetic tree construction?
(b) Generate a tree using the NJ method from the following distance matrix. Show each step. [3+10]

	A	B	C	D	E
A		5	4	7	6
B			7	10	9
C				7	6
D					5
E					

(Presentations carry 10 marks)

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2017-18

M.Stat Second Year

Pattern Recognition

Date: 29.11.2017 Maximum Marks: 100 Duration: $3\frac{1}{2}$ Hours

(All answers should be brief and to the point. Answer as much as you can. The maximum you can score is 100.)

1. Consider a two-class classification problem based on a single variable, where the densities of the two classes are given by

$$f_1(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 1+\cos(2\pi x) & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Assuming that the prior probabilities of the two classes are equal, compute the average misclassification probability of the Bayes classifier. [5]

   (b) Show that the minimax classification rule will be of the form

$$\delta(x) = \begin{cases} 1 & \text{if } c_0 \le x \le 1-c_0 \\ 2 & \text{otherwise,} \end{cases}$$

   where c_0 is a positive constant smaller than $1/4$. Show that the average misclassification probability of this classifier cannot be smaller than $(\pi-1)/2\pi$. [5+2]

2. (a) In order to construct a classification tree for a two-class classification problem, one needs to find the best split based on each of the measurement variables. Consider 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	30	40	30	100
No. of. observations from class-2	30	0	70	100
Total	60	40	100	200

   If the impurity function $\psi(p_1, p_2)$ is concave and symmetric in its arguments, find the best split based on the measurement variable C. [6]

   (b) Describe how you will construct a random forest classifier using the classification tree algorithm. Also describe how you will compute the out-of-bag error estimate of this classifier. [3+3]

3. (a) If two normal distributions differ only in their locations and have the same prior probability, show that the average misclassification probability of the best linear classifier is a decreasing function of the Mahalanobis distance between the two distributions. [6]

   (b) What is the Nadaraya Watson estimator of a regression function? How can this estimator be used to construct a nonparametric classifier? How does this classifier differ from the usual kernel discriminant analysis rule? [2+2+2]

4. (a) Formulate the minimization problem for a linear support vector machine when the observations from the two classes are not linearly separable. Describe how you will use support vector machine for nonlinear classification. [3+3]

   (b) What is semi-supervised classification? Briefly describe how EM algorithm can be used for model based semi-supervised classification. [2+4]

P.T.O

5. Prove or disprove the following statements. [14 × 3=42]

(a) It is not possible to have a two-class classification problem, where all linear classifiers have misclassification probability 0.5.

(b) If two elliptically symmetric distributions differ only in their locations and they have the same prior, the Bayes classifier will always be linear.

(c) Half-space median of a bivariate continuous distribution cannot lie outside the support of the distribution.

(d) Half-space depth of the half-space median of a bivariate distribution cannot be smaller than 0.5.

(e) In a two-class classification problem with equal priors, the average misclassification probability of the maximum half-space depth classifier cannot be 0.5.

(f) If $\mathbf{X}$ and $\mathbf{HX}$ have the same distribution for all orthogonal matrix $\mathbf{H}$, $Var(\mathbf{X})$, if it exits, will be of the form $C\mathbf{I}$, where $\mathbf{I}$ is the identity and C is a positive constant.

(g) L_1 distance between two density functions remains invariant over smooth monotone transformation of the random vector.

(h) A single layer perceptron model with sigmoid transformation is equivalent to a logistic regression model.

(i) If the impurity function used in a classification tree is concave, the reduction in impurity function due to a split is always non-negative.

(j) The integral of a k-nearest neighbor density estimate over the entire measurement space may not be finite.

(k) A kernel discriminant analysis rule that uses Gaussian kernel and a common bandwidth h for all classes behaves like the 1-nearest neighbor classifier when the the bandwidth h shrinks to zero.

(l) Popular k-means algorithm may fail to properly identify two elliptic clusters even when they are well separated.

(m) If the data set contains two well separated clusters having widely different spreads, single linkage algorithm can outperform the complete linkage algorithm.

(n) If an weighted graph has k connected components, the unnormalized Graph Laplacian used by spectral clustering algorithm has one eigenvalue with multiplicity k.

6. Computer Assignments [20]

INDIAN STATISTICAL INSTITUTE, KOLKATA

FINAL EXAMINATION: FIRST SEMESTER 2017 - '18

M.STAT II YEAR

Subject	:	**Functional Analysis**
Time	:	2 hours 30 minutes
Maximum score	:	50

Attempt all the problems. Please use a new page to answer each question, making sure that the question number in the margin can be read, even after stapling. If you attempt the same problem several times, please strike out all the attempts except the final one before submitting your answer script. Justify every step in order to get full credit of your answers. All arguments should be clearly mentioned on the answer script. Points will be deducted for missing or incomplete arguments.

(1) Let $\mathcal{H}$ be a Hilbert space, $x \in \mathcal{H}$ and $C \subset \mathcal{H}$ be a closed convex set. Show that the distance between x and C is attained at a unique point of C.

[8 marks]

(2) Let (X, Ω, μ) be a σ-finite measure space. Take $\varphi \in L^\infty(X, \Omega, \mu)$. Define M_φ $L^2(X, \Omega, \mu) \to L^2(X, \Omega, \mu)$ by $(M_\varphi f)(x) = \varphi(x)f(x)\ \forall x \in X$.

a) Prove that the spectrum of M_φ is the essential range of φ, namely,
$$\{\lambda \in \mathbb{C} : \mu(\varphi^{-1}(U)) > 0 \text{ for each open set } U \text{ containing } \lambda\}$$

b) When is M_φ a projection?

c) When is M_φ a unitary operator?

[(5+2+2)=9 marks]

(3) The Volterra operator $V : L^2[0,1] \to L^2[0,1]$ is defined by
$$Vf\ (x) = \int_0^x f(y)\ dy \text{ for all } f \in L^2[0,1].$$

a) V is a bounded operator.

b) Show that $V^n f\ (x) = \int_0^x f(y)\frac{(x-y)^{n-1}}{(n-1)!}dy$ for all $n \in \mathbb{N}$.

c) Show that $\sigma(V) = \{0\}$.

d) Show that V is a compact operator.

e) Find V^*.

[3+5+4+2+2 = 16 marks]

(4) Consider bounded linear maps S and $T : L^2[0,\infty) \to L^2[0,\infty)$ defined by
$$(Sf)(x) = f(x+1) \quad \text{and} \quad (Tf)(x) = f(x) + f(x+2)$$

(i) Show that $\sigma(S) = \{z \in \mathbb{C} : |z| \leq 1\}$ (Hint: For $|\lambda| < 1$, consider functions $f : [0, \infty) \to \mathbb{C}$ of the form $f(x+n) = \lambda^n f(x)$ for all $x \in [0,1), n \in \mathbb{N}$)

(ii) Find $\sigma(T)$ (You may use the spectral mapping theorem for polynomial functions.)

[8+5 = 13 marks]

(5) Consider the operator $T : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ defined by

$$l^2(\mathbb{N}) \ni (x_1, x_2, x_3, \ldots) \stackrel{T}{\mapsto} (x_2, x_4, x_6, \ldots) \in l^2(\mathbb{N}).$$

Find (a) $\sigma_p(T)$, (b) $\sigma(T)$.

[4+3 = 7 marks]

(6) Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\{e_n\}_{n=1}^{\infty}$. Define $T : \mathcal{H} \to \mathcal{H}$ by

$$T(x) = \sum_{n=1}^{\infty} \frac{1}{n+1} \langle x, e_{n+1} \rangle e_n$$

for $x \in \mathcal{H}$. Show that T is a compact operator and find T^*.

[7 marks]

INDIAN STATISTICAL INSTITUTE

M. Stat 2nd Year

Martingale Theory

Date: 01.12.17 Maximum Marks: 70 Duration: 3 hours

Anybody caught using unfair means will immediately get 0. Please try to explain every step. NO NOTES ARE ALLOWED.

(1) Let Θ be distributed as $U[0,1]$ and conditional on Θ, let $X_1, X_2 \cdots$ be iid Ber(Θ). Show that

$$\mathsf{E}[\Theta \,|X_1, \cdots X_n] \to \Theta \text{ almost surely .}$$

[10 marks]

(2) (a) Let $X_1, X_2, \cdots$ be independent with $\mathsf{E}[X_i] = 0$ and $\mathrm{Var}(X_m) = \sigma_m^2 < \infty$ and $s_n^2 = \sum_{m=1}^n \sigma_m^2$. Then show that $S_n^2 - s_n^2$ is a martingale, where $S_n = \sum_{i=1}^n X_i$.

(b) Suppose now additionally $|X_m| < K$, then show that

$$\mathbf{P}\left(\max_{1 \le m \le n} |S_m| \le x\right) \le \frac{(x+K)^2}{s_n^2}.$$

[5+5=10 marks]

(3) Let X_i be iid random variables. Then show that $\frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \sin(X_i X_j)$ converges to a constant.

[10 marks]

(4) Suppose $\mathbf{P}$ and $\mathbf{Q}$ are two probability measures on space $\Omega = \{0,1\}^{\mathbb{N}}$ of infinite binary sequences equipped with product σ-algebra generated by its cylinder sets, with $p_k = \mathbf{P}(\omega : \omega_k = 1) \in (0,1)$ and $q_k = \mathbf{Q}(\omega : \omega_k = 1) \in [0,1]$.

(a) Show that $\mathbf{Q} \ll \mathbf{P}$ if and only if

$$\sum_k \left(1 - \sqrt{p_k q_k} - \sqrt{(1-p_k)(1-q_k)}\right) < \infty.$$

(b) Show that if $\sum_k |p_k - q_k|$ is finite then $\mathbf{Q} \ll \mathbf{P}$.

[5+5=10 marks]

(5) Mr.Trump decides to post a random message on Facebook and he starts typing a random sequence of letters $\{U_k\}_{k \ge 1}$ such that they are chosen independently and uniformly from the 26 possible english alphabets. Find out the expected time of first appearance of the word COVFEFE.

P.T.O

[10 marks]

(6) Let $(X_n : n \geq 0)$ be a Markov chain with state space $\mathbb{Z}$ and transition matrix P. Let $f : \mathbb{Z} \to \mathbb{R}$ be a function such that the following hold for every $i \in \mathbb{Z}$:

$$\begin{aligned} \sum_{j \in \mathbb{Z}} P(i,j) f(j) &= f(i), \\ \sum_{j \in \mathbb{Z}} P(i,j) \left[f(j) - f(i)\right]^2 &= 1, \\ \text{and } \sum_{j \in \mathbb{Z}} P(i,j) \left|f(j) - f(i)\right|^{2+\delta} &< \infty, \end{aligned}$$

for some $\delta > 0$. Show that $f(X_n)/\sqrt{n} \Rightarrow N(0,1)$.

[10 marks]

(7) Suppose X and Y are integrable random variable and suppose that $\mathsf{E}[X|Y] = Y$ and $\mathsf{E}[Y|X] = X$ almost surely. Then show that $X = Y$ almost surely.

[10 marks]

Indian Statistical Institute
M. Stat. II year, First Semester, 2017-18
Semestral Examination
Signal and Image Processing

Date: 01.12.2017 **Duration: 210 minutes** **Maximum Marks: 100**

Note: (i) This paper carries 110 marks. Answer as much as you can.
(ii) Answer the two parts in two different answer scripts.

Part A: Signal Processing

1. a) For the sequences: $x_1(n) = \cos\left(\frac{2\pi}{N}n\right)$ and $x_2(n) = \sin\left(\frac{2\pi}{N}n\right)$, $0 \le n \le N-1$, determine the N-point circular convolution of $x_1(n)$ and $x_2(n)$.

b) Let $X(k)$ be the N-point DFT of the sequence $x(n)$ for $0 < n < N-1$. How does the N-point DFT of the sequence $s(n) = X(n)$ for $0 < n < N-1$ relate to $x(n)$ - explain.

6+4 = 10

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

$$x(n) = \left\{\underset{\uparrow}{\frac{1}{2}}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0\right\}$$

using the in-place radix-2 decimation-in-time algorithm. Show all the steps and the resulting signal flow graphs. 10

3. Consider a discrete time LTI system described by the following input-output relation:

$$y(n) = ay(n-1) + x(n) - \frac{1}{a}x(n-1).$$

a) Determine a closed-form expression for the impulse response $h(n)$ of this system (in terms of a).

b) Determine a closed-form expression for the autocorrelation sequence $r_{hh}(n)$ corresponding to the impulse response $h(n)$ of the system. Simplify as much as possible and show all the steps.

c) The signal below is input to the system considered above:

$$x(n) = \left\{ \frac{\sin\left(\frac{3\pi}{4}n\right)}{\pi n} \right\}$$

(i) Plot the energy density spectrum for the output $y(n)$.

P.T.O

(ii) Find the total energy in $y(n)$ (in terms of a).

3+3+(2+2) = 10

4. Consider a causal FIR filter of length $M = 3$ with the following impulse response starting at $n = 0$: $h(n) = \{1, 2, 1\}$.

a) Provide a closed-form expression for the 8-point DFT of $h(n)$, denoted by $H_8(k)$, as a function of k. Simplify as much as possible.

b) Consider the sequence $x(n)$ of length $L = 8$ below, equal to a sum of several finite-length sine waves.

$$x(n) = \left[2\cos\left(\frac{\pi}{4}n\right) + 3\cos\left(\frac{\pi}{2}n\right) + 4\cos(\pi n)\right]\{u(n) - u(n-8)\}$$

Suppose $y_8(n)$ is formed by computing $X_8(k)$ as an 8-point DFT of $x(n)$, $H_8(k)$ as an 8-point DFT of $h(n)$, and then $y_8(n)$ as the 8-pt inverse DFT of $Y_8(k) = X_8(k).H_8(k)$. Express the result $y_8(n)$ as a weighted sum of the finite-length sine waves similar to how $x(n)$ is written above.

4+6 = 10

5. a) Suppose $x(n)$ is the sample sequence of a wide-sense stationary discrete-time random process. $x(n)$ is applied as an input to a deterministic LTI system with impulse response $h(n)$. Derive a relation between the autocorrelation function of the output $y(n)$ with the autocorrelation function of the input $x(n)$. Show all the intermediate steps.

b) 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.

c) Suppose $a = [a_1, a_2, ..., a_M]$ and $b = [b_1, b_2, ..., b_N]$ be two discrete sequences. Also let $c = a * b$, where * denotes linear convolution. If A denotes the Toeplitz matrix defined on the sequence a, prove that the deconvolution of sequence b is given by:

$$b = \left(A^T A\right)^{-1} A^T c.$$

5+4+6 = 15

Part B: Image Processing

1. Describe Canny edge detection method for gray level images. [15]

2. Describe a method for finding line segments in a binary image using Hough Transform. [12]

3. Describe a region-based segmentation method for gray level images using quad tree. [8]

4. (a) Define skeleton of a region in $\Re^2$
   (b) Describe a thinning algorithm for an object in a binary image. [5+10]

5. Describe a method for introducing salt and pepper noise in a gray level image. [5]

INDIAN STATISTICAL INSTITUTE
First Semester Examination: 2017-18

M.Stat Second Year
Pattern Recognition (Backpaper)

Date: 14/12/2017 | Maximum Marks: 100 | Duration: 3 Hours

1. (a) Give an example of a classification problem, where densities of the two competing classes are not identical but the average misclassification probability of the Bayes classifier is equal to the smaller of the two prior probabilities. [4]

   (b) Consider a classification problem, where each of the two classes is an equal mixture of two bivariate normal distributions. While Class-1 is a mixture of $N_2(1,1,1/4,1/4,0)$ and $N_2(-1,-1,1/4,1/4,0)$, Class-2 is a mixture of $N_2(-1,1,1/4,1/4,0)$ and $N_2(1,-1,1/4,1/4,0)$. If the two classes have the same prior probability, find the Bayes classifier and the corresponding Bayes risk. [4+4]

2. Consider a two-class classification problem between two bivariate normal distributions $N_2(0,0,1,1,\rho)$ and $N_2(0,0,1,1,-\rho)$, where $\rho > 0$.

   (a) Assuming equal prior probabilities for the two classes, find the Bayes classifier. Is this a minimax rule? Justify your answer. [3+3]

   (b) Compute the Bayes risk and show that it is a decreasing function of ρ. [6]

3. Suppose that $\mathbf{X}_1$ and $\mathbf{X}_2$ follow two elliptically symmetric distributions, which have the same scatter matrix Σ, and differ only in their locations $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$. If $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent, show that

   (a) $\mathbf{X}_1 - \mathbf{X}_2$ follows an elliptically symmetric distribution. [6]

   (b) $P\{\boldsymbol{\alpha}'(\mathbf{X}_1 - \mathbf{X}_2) > 0\}$ is maximized when $\boldsymbol{\alpha} \propto \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$. [6]

4. Let $\mathbf{x}_1,\ldots,\mathbf{x}_m$ and $\mathbf{y}_1,\ldots,\mathbf{y}_n$ be two sets of independent observations from two multivariate continuous distributions F and G, respectively. Project these observations along a direction $\boldsymbol{\alpha}$ and define the Mann-Whitney statistic $MW(\boldsymbol{\alpha}) = \left|\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} I\{\boldsymbol{\alpha}'\mathbf{x}_i > \boldsymbol{\alpha}'\mathbf{y}_j\}\right|$ and the Kolmogorov-Smirnov statistic $KS(\boldsymbol{\alpha}) = \sup_\beta |F_m^{\boldsymbol{\alpha}}(\beta) - G_n^{\boldsymbol{\alpha}}(\beta)|$ along that direction, where $F_m^{\boldsymbol{\alpha}}(\beta) = \frac{1}{m}\sum_{i=1}^{m} I\{\boldsymbol{\alpha}'\mathbf{x}_i \leq \beta\}$ and $G_n^{\boldsymbol{\alpha}}(\beta) = \frac{1}{n}\sum_{i=1}^{n} I\{\boldsymbol{\alpha}'\mathbf{y}_i \leq \beta\}$. Show that

   (a) $\sup_{\boldsymbol{\alpha}} MW(\boldsymbol{\alpha}) = 1 - HD(\mathbf{0},\Omega_{mn})$, where $HD(\mathbf{0},\Omega_{mn})$ is the half-space depth of the origin w.r.t. the data cloud $\Omega_{mn} = \{\mathbf{z}_{ij} = \mathbf{x}_i - \mathbf{y}_j : 1 \leq i \leq m,\ 1 \leq j \leq n\}$. [6]

   (b) Maximization of $KS(\boldsymbol{\alpha})$ is equivalent to finding the best linear classifier (the linear classifier with the minimum training sample error rate) when the prior probabilities of the two classes are equal. [6]

5. (a) Consider a J class ($J > 2$) classification problem, where the competing classes have prior probabilities $\pi_1,\ldots,\pi_J$. If $\pi_1 > \pi_j$ for all $j > 1$, show that the error rate of the kernel discriminant analysis rule converges to 1 - π_1 as h tends to infinity. [4]

   (b) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed with a density $f(x)$, which is a decreasing function of $\|\mathbf{x}\|$. Let $\hat{f}_h$ be the kernel density estimate of f based on a spherically symmetric kernel. Show that $E[\hat{f}_h(x)]$ is also a decreasing function of $\|\mathbf{x}\|$. [6]

   (c) Consider a classification problem between two Gaussian distributions $N(\mathbf{0},\sigma_1^2\mathbf{I})$ and $N(\boldsymbol{\mu},\ \sigma_2^2\mathbf{I})$, where $\boldsymbol{\mu} = (\mu,\mu,\ldots,\mu)'$. If the sample size is fixed, study the high dimensional behavior of the 1-nearest neighbor rule when (*i*) $\mu^2 > |\sigma_1^2 - \sigma_2^2|$ and (*ii*) $\mu^2 < |\sigma_1^2 - \sigma_2^2|$. [3+3]

P.T.O

6. (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) Let $T_0 \supset T_1 \supset \ldots \supset T_k$ be a nested sequence of regression trees. Consider a cost function $RSS_\alpha(T) = RSS(T) + \alpha|\tilde{T}|$, where RSS stands for the residual sum of squares and $|\tilde{T}|$ is the cardinality of the leaf nodes in the tree T. For $i = 1, 2, \ldots, k$, define

$$\eta_i = (RSS(T_i) - RSS(T_0))/(|\tilde{T}_0| - |\tilde{T}_i|).$$

If $\eta_1 > \eta_2$, show that for all choices of $\alpha \geq \eta_2$, $RSS_\alpha(T_1)$ exceeds $RSS_\alpha(T_2)$. [4]

(c) Give an example of an impurity function, which can be used for constructing a classification tree when there are more than two classes. [2]

7. Suppose that observations from two competing classes are linearly separable.

(a) Show that the maximum likelihood estimates (MLE) of parameters of the logistic regression model will not exist. [6]

(b) Formulate a quadratic optimization problem for linear classification using support vector machine. How will you modify your objective function and associated constraints if the observations are not linearly separable? [3+3]

8. (a) Give a brief description of the spectral clustering algorithm based on an unnormalized Graph Laplacian L. If the similarity graph has k connected components, find the rank of the matrix L. [6+3]

(b) Describe how k-means algorithm can be used for image compression. [3]

Indian Statistical Institute

M.Stat. Second year, Mid Semester Exam: 2018 Date: 19.02.2018

Topic: Statistical Computing II

Maximum Marks: 50, Duration: 2 hours

1. Suppose data are collected on diastolic blood pressure, LDL (bad) cholesterol, and blood glucose level for N patients longitudinally. Different subjects are measured at different time points and suppose we get T_i measurements from the i-th patient. Thus the response from the i-th patient at the j-th time point on the k-th response feature can be denoted by $Y_{ik}(t_{ij})$, where $i = 1, \ldots, N$, $j = 1, \ldots, T_i$ and $k = 1, 2, 3$.

   Using a non-parametric approach of modeling mean trajectories and a parametric approach of modeling covariances, propose a semi-parametric modeling of the above dataset. Write down the joint likelihood function, and explain the parameter estimation method explicitly. [15]

2. In health economics, total medical expenditure is considered as an important measure of financial burden for the older people. Suppose data are collected on total medical expenditure and other related covariates for a particular season of the year from n older individuals (with age > 65 years). In the data, $n_1 (< n)$ responses are found to be exactly equal to zero and non-zero (continuous) responses are denoted by $Y_1, Y_2 \ldots, Y_{n_2}$, where $n_1 + n_2 = n$. Suppose there are p regressor variables $X_1, X_2, \ldots, X_p$ in the data. Develop a Bayesian Two-Part model for the above dataset with the goal of predicting the medical expenditure for $(n+1)$-th individual given the relevant information on the predictors for that subject. [15]

3. Suppose annualized data are collected from 6 different airlines from the United States for 2005-2014 (10 years) on several variables including output (revenue passenger miles), total cost (thousands of dollars), fuel price, and load factor (average capacity used). Suppose the goal is to predict the output (annualized) based on the other three variables. Develop a suitable Bayesian model for such prediction. Write down the data model, prior structure and the full posterior density explicitly. Use the following notations:
   Y_{jt}= output (annualized) for the j-th airline at the t-th year, $j = 1, \ldots, 6$; and $t = 1, \ldots, 10$.
   x_{1jt}= total cost (annualized) for the j-th airline at the t-th year.
   x_{2jt}= fuel price (annualized) for the j-th airline at the t-th year.
   x_{3jt}= load factor (annualized) for the j-th airline at the t-th year. [10]

4. Define the following terms:
   (i) Missing Completely at Random, (ii) Missing at Random, (iii) Missing Not at Random, (iv) Intermittent Missingness, (v) Monotone Missingness. [10]

MSTAT II - Stochastic Calculus for Finance

Midsem. Exam. / Semester II 2017-18

Date - February 20, 2018 / Time - 2 hours

Maximum Score - 30

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

1. (5+5+5=15 marks) Write TRUE or FALSE and justify your answer.

   (a) Let (X, d) be a complete separable metric space and $\mathcal{B}(X)$ be the Borel σ-field generated by the open sets of (X, d). Then any finite Borel measure μ defined on $\mathcal{B}(X)$ is tight. [A finite Borel measure is a map $\mu : \mathcal{B}(X) \mapsto [0, \infty)$ is defined as (i) $\mu(\phi) = 0$, (ii) for $\{A_i : i \geq 1\} \in \mathcal{B}(X)$ mutually disjoint, $\mu(\cup A_i) = \sum \mu(A_i)$, and (iii) $\mu(X) < \infty$.]

   (b) Let $\{M_n\}$ be a square integrable martingale with respect to the filtration $\{\mathcal{F}_n\}$. Then $\{M_n^2 - < M >_n\}$ is a martingale with respect to the filtration $\{\mathcal{F}_n\}$, where $< M >_n = \sum_{k=1}^{n} E(M_k - M_{k-1})^2$.

   (c) Let $\{B(t) : t \geq 0\}$ be a standard Brownian motion. Then $W(t) = \begin{cases} tB(1/t) & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}$ is a standard Brownian motion.

2. (5+4=9 marks) Let $S_n = X_1 + \cdots + X_n$ for $n \geq 0$, where $X_i \sim N(0, 1)$, i.i.d. random variables. Let $Y_n = \exp\{S_n - n/2\}$.

   (a) Show that, $P(max_{1 \leq k \leq n} Y_k \geq a) \leq 1/a$, for any real $a > 0$.

   (b) Show that, $E(max_{1 \leq k \leq n} Y_k)^2 \leq 4e^n$.

3. (5+4=9 marks)

   Let $\{B(t) : t \geq 0\}$ be a standard Brownian motion.

   (a) Show that, for any reals $a < b$, $\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2 \to (b - a)$ in L_2, as mesh of the partition, $a = t_0 < t_1 < \cdots < t_n = b$, is going to zero.

   (b) Show that, for almost all sample paths, $\{B_t\}$ has unbounded variation.

All the best.

Statistical Methods in Epidemiology & Ecology

Mid-Semestral Examination

M.Stat.- II Year, 2017-2018

Total Marks - 60 Time: 3 hr.

Attempt all questions:

1. Let us consider the Blumberg's growth equation

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

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 of a bounded growth law and its parameters, as an extension of Malthus's unbounded growth law, using the concept of basic and two opposite growth pulses.

(b) Find the approximate analytical solution of the growth equation (1), assuming size to be bounded by the carrying capacity. Comment on the stability aspect of this approximate growth curve, with/without using the Lypunov stability theorem.

(c) Find the condition for which the growth curve derived from equation (1), is always symmetric with respect to its point of inflexion.

(d) Evaluate the point of inflection of the absolute growth rate function of the growth curve as defined in (1) and compare it with the logistic case.

(e) Define Allee effect and its critical size. Why the strong Allee effect is more serious in the context of ecology. After proper modification, the equation (1) can be used to model the Allee effect phenomena with positive/null critical size - Discuss.

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

2. (a) Define Fisher's Relative Growth Rate(RGR) based on the size measurement at two specific time points.

(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 $E(R(t)) = \theta(t) = be^{-at}t$, an extended Gompertz rate profile. Suggest an estimate of the rate parameter "a" based on a specific time interval $(t, t + \Delta t)$.

(c) Using the approximate expression for the expectation and variance of the logarithm of the ratio of RGR for two consecutive time points, find the asymptotic distribution of the decay parameter a based on the data matrix $X(n \times q)$. Note that, in the data matrix $X(n \times q)$ any row corresponds to a q - variate size measurements available at q time points on one of the n individuals. Construct an asymptotic test for the null hypothesis of the equality of expected rate parameters for the two separate species. Also, suggest required modifications of the test statistic when the time points are not equispaced (for simplicity, you may assume only three time points - for example 1, 2, 5.)

[2+3+(5+3+4) = 17]

3. (a) Define equilibrium probabilities of a general birth death process.

   (b) The θ - logistic growth equation for a single species dynamics is defined as follows:

$$\frac{dX(t)}{dt} = aX(t) - bX(t)^{\theta+1}, \tag{2}$$

   where parameters have their usual interpretations.

   Define the birth and death rates clearly in the context of the above growth equation and derive the equilibrium probabilities.

   (c) Derive the limiting form of the following growth model when $\theta \to 0$ and parameters have their usual meanings

$$\frac{1}{X(t)}\frac{dX(t)}{dt} = rX(t)^{\gamma}\left(\frac{x(t)}{a} - 1\right)\left(1 - \left(\frac{x(t)}{k}\right)^{\theta}\right). \tag{3}$$

   (d) Discuss the application of the limiting model in the context of ecology.

[3 + (3 + 5) + 4 + 2 = 16]

INDIAN STATISTICAL INSTITUTE

M. Stat 2nd Year

Weak Convergence and Empirical processes

Date: 21.02.18 Maximum Marks: 30 Duration: 2 hours

Total exam is of 30 marks. Anybody caught using unfair means will immediately get 0. Please try to explain every step. Use of any kind of electronic gadgets is completely prohibited.

(1) Suppose that $\mathcal{U} \subset \mathcal{B}$ is a family of sets closed under finite unions and intersections and each open set is a countable union of sets in $\mathcal{U}$. Assume that $\mu_n(A) \to \mu(A)$ for every $A \in \mathcal{U}$. Show that $\mu_n \Rightarrow \mu$. [5 points]

(2) For any probability measures P and Q on the real line, denote by F_P and F_Q their distribution functions. Define

$$d(P,Q) = \frac{1}{2}\int_{-\infty}^{\infty} |F(x) - G(x)| e^{-|x|} dx.$$

(a) Show that d is a metric on the space of probability measures on real line.

(b) Show that d metrizes weak convergence. [10 points]

(3) Let C be the collection of bounded real continuous functions f on $(-\infty, \infty)$ which have finite limits at $-\infty$ and $+\infty$. For any probability P on $\mathbb{R}$ define an operator T_P by

$$T_P f(x) = \int f(x-y) P(dy).$$

(a) Show that T_P maps C into C.

(b) Show that $P_n \Rightarrow P$ if and only if for every $f \in C$, $T_{P_n} f \to T_P f$ uniformly. [10 points]

(4) Let $(W_t; 0 \leq t \leq 1)$ be standard Brownian motion. Show that $\tilde{W}_t = W_{1-t} - W_1$ is again Brownian motion on $0 \leq t \leq 1$. [5 points]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2017 – 18

MStat (2nd Year)

Financial Econometrics

Date: 22nd February 2018 Maximum Marks: 30 Duration: 2 Hours

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

2. 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'. [2 + 4 + 4 = 10]

3. (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. [3 + 6 + 4 = 13]

Indian Statistical Institute
Semester 2, Academic Year: 2017-18
Mid-Semester Examination
Course: M. Stat 2nd Year
Subject: Brownian Motion and Diffusions

Total Points: 30

Date: 23.2.2018 **Time: 10:30-12:30 AM**

Answers must be justified with clear and precise arguments. If you refer to a theorem/result proved in class, state it explicitly. More than one answer to a question will not be entertained and only the first uncrossed answer will be graded.

1. Using a suitable martingale constructed from Brownian motion show that for $\lambda, \beta > 0$,

$$P(\max_{0 \le s \le t} \left(W_s - \frac{\lambda}{2} s\right) \ge \beta) \le e^{-\lambda\beta}.$$

You can use other inequalities for martingales/submartingales extending them in the continuous time case. 4 pts.

2. Consider the following process $V(t) = \int_0^t e^{-(t-u)} dW(u), V(0) = 0$. Note that the finite dimensional distributions will be multivariate normal with zero mean. $5 + 5 = 10$ pts.

(a) Find the covariance function $EV(s)V(t)$ and explain if the distribution of $(V(t+t_1), V(t+t_2), \ldots, V(t+t_n))$ depends on t.

(b) Now suppose $Z \sim N(0, 1/2)$ and is independent of $V(t), t \ge 0$. Find the covariance function of the process $U(t) = V(t) + e^{-t} Z$ and explain if the distribution of $(U(t+t_1), U(t+t_2), \ldots, U(t+t_n))$ depends on t.

3. Consider a progresively measurable step function $g(t) = a_0 1_{(0,t_1]}(t) + a_1 1_{(t_1,t_2]}(t)$ where a_i is $\mathcal{F}_{t_i}$ measurable. $4 + 4 = 8$ pts.

(a) Show that with $M_t = \int_0^t g(s) dW(s)$, one has M_t, $M_t^2 - \int_0^t g(s)^2 ds$, are $\mathcal{F}_t$ martingales. It will be enough to consider just one case: $t_1 < s < t < t_2$. **P.T.O.**

P.T.O

(b) Suppose f is a progressively measurable function satisfying $\int_0^T Ef(s)^2 ds < \infty$ and consider progressive measurable step functions f_n such that $f_n \to f$ in the norm $||g||^2 = \int_0^T Eg^2(s)ds$. Show that the limiting Ito integral $M(t) = \int_0^t f(s)dW(s)$ is a martingale, by considering $C \in \mathcal{F}_s$ and the equality $EM_n(t)1_C = EM_n(s)1_C$ for $s < t$, where $M_n(t) = \int_0^t f_n(s)dW(s)$.

4. Suppose f is twice continuously differentiable on $[0, T]$. In addition, assuming $\int_0^T E\{f'(B(s))\}^2 ds < \infty$, the Ito formula was derived as $f(B(t)) - f(B(0)) = \int_0^t f'(B(s))dB(s) + \frac{1}{2}\int_0^t f''(B(s))ds, 0 \le t \le T$.

$4 + 4 = 8$ pts.

(a) Write down the Ito formula for $f(x) = x^4$ and compute $E\int_0^t B(s)^2 ds$ using this formula.

(b) Compute $E\{B(t)^2 \times \int_0^t B(s)dB(s)\}$ (using a different Ito formula if necessary). One cannot just take $B(t)^2$ inside the integral (and interchange expectation and integration) because Ito integrals have been defined for progressively measurable integrands only.

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: (2017–2018)

M. Stat Second Year

Inference for High Dimensional Data

Date: 23/02/2018 Marks: 40 Duration: $2\frac{1}{2}$ hours

Attempt all questions

1. (a) Assume that you need to estimate a function $F(\theta_1, \ldots, \theta_p)$ from noisy observations $X_j = \theta_j + \epsilon_j$. Assume that the noise variables $\epsilon_1, \ldots, \epsilon_p$ are all centered with variance σ^2. Under appropriate assumptions show that the mean square error can be very large for large p even if σ^2 is small.

   (b) Give an example of a common statistical model where the above situation arises, and illustrate the phenomenon in details for your example.

[5+5=10]

2. Let $f : \mathbb{R} \mapsto \mathbb{R}$ be L-Lipschitz and let $X_1, \ldots, X_n$ be *iid* with finite variance σ^2.

   (a) Prove that

$$\lim_{n\to\infty} P\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i) - E\left[f(X_1)\right] \geq \frac{L\sigma}{\sqrt{n}}x\right) \leq e^{-x^2/2} \text{ for } x > 0.$$

   (b) Under appropriate distributional assumptions on $X_1, \ldots, X_n$, show that the above probability inequality can be achieved non-asymptotically.

   (c) Consider the normal regression problem with *iid* error variance σ^2 in the co-ordinate sparse setting. Assuming that it is not known *a priori* that the true regression function belongs to a known linear subspace of $\mathbb{R}^n$, propose a method to estimate the true regression function.

   (d) Also obtain an unbiased estimator of the risk of the estimators associated with the various linear subspaces.

[3+3+2+2=10]

3. Let $Z \sim N\left(0, \sigma^2 I_d\right)$, a d-dimensional zero-mean normal random variable with variance $\sigma^2 I_d$, where $\sigma > 0$ and I_d is the identity matrix of order d. Let $\|Z\|$ denote the Euclidean norm of Z.

   (a) Prove that the map $Z \mapsto \|Z\|$ is 1-Lipschitz.

   (b) Using the Gaussian concentration inequality, prove that the variance of $\|Z\|$ can be bounded above independently of the dimension d.

[3+7=10]

4. 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, and that the components of ϵ are *iid* $N(0, \sigma^2)$. 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 $\lambda = K\left(1 + \sqrt{2\log(p)}\right)^2$, where K is a constant, this penalty is approximately equal to the penalty

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

   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 $\mathbf{X}$.

   (c) Now assume that $\mathbf{f}^* = \mathbf{0}$.

      (i) Obtain the oracle model.

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

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

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

      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.

[2+2+1+2+3=10]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination 2017-18

M.Stat. - 2nd Year
Survival Analysis

26th 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. Suppose that the lifetime variable T has a mixed distribution with a continuous hazard $\lambda_c(t)$ and discrete hazards $\alpha_1, \alpha_2, \ldots$ at the points $x_1 < x_2 < \ldots$, respectively. Derive the survival function of T. Hence or otherwise describe a method for simulating an observation from T. How will your simulation method change under the hybrid censoring scheme? [2+2+2]=6

2. In a parallel system with two independent components each having constant hazard λ, once a component fails the hazard of the other changes to α.

(a) Derive the Survival function and the mean residual life function of the overall system lifetime.

(b) Based on the random right censored data from n such systems, derive the MLE score equations for λ and α and an asymptotic test for testing $H_0 : \lambda = \alpha$ against the omnibus alternative. [(2+2)+(3+4)]=11

3. Proof that a continuous lifetime random variable has the memoryless property if and only if it is exponentially distributed. [4]

4. Consider the random right censored data from a lifetime distribution F with $t_1 < \ldots < t_k$ being the observed distinct failure times.

(a) Derive the Kaplan Meier estimator $\widehat{S}_{KM}$ of $S = 1 - F$ as a non-parametric MLE of the survival function.

(b) What is the asymptotic variance of $\widehat{S}_{KM}$? How will you estimate it consistently?

(c) How will this MLE change if it is known a priori that the median of F is $m > t_1$?

(d) Based on $\widehat{S}_{KM}$ in (a), derive an estimate of mean of F along with the asymptotic variance estimate. [4+3+3+5]=15

P. T. O

5. Suppose we have additional covariate observation $z_i \in \mathbb{R}^p$ associated with the random right censored observation (x_i, δ_i) for $i = 1, \ldots, n$, respectively. We want to study the effect of covariates on the underlying lifetime variable.

(a) Define the PH and the AFT regression models for this purpose.

(b) Find out all the common members of the PH and the AFT regression model families.

(c) Consider a parametric AFT model with logistic error distribution (having a variance σ). Write down the likelihood function for the underlying parameters and describe a method to find out their MLEs.

(f) How can you construct a test for testing the significance of the regression model under the assumption in Part (c)? [3+4+(3+3)+5]=18

6. Consider random right censored data from two lifetime distributions F_1 and F_2 respectively. Suppose we want to test for $H_0 : F_1 \equiv F_2$ against the omnibus alternative.

(a) Describe the Grehan test procedure for this problem. Show that the Grehan test statistic coincides with the Wilcoxon test statistic in case of no censoring.

(b) Describe the Tarone-Ware Class of tests for this problem and show that it contains the Grehan test for properly chosen weights. [(3+2)+(1+4)]=10

7. Consider random right censored data from two discrete lifetime variables having constant hazards λ_1 and λ_2, respectively. Derive the Score test, Wald test and the Likelihood ratio test for testing $H_0 : \lambda_1 = \lambda_2$ against the omnibus alternative. [3×4]=12

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination : 2017-18

Percolation Theory

M.Stat 2nd year

26^{th} February, 2018

Maximum marks: 35 | Duration: 3hr

1. Let A and C be increasing events and B and D be decreasing events. The four aforementioned events A, B, C, and D depend on the edges from the finite sets $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$, and $\mathfrak{D}$ respectively. If $\mathfrak{A} \cap \mathfrak{B} = \mathfrak{A} \cap \mathfrak{D} = \mathfrak{B} \cap \mathfrak{C} = \phi$, show that

$$\mathbb{P}[A \cap B \cap C \cap D] \geq \mathbb{P}[A \cap B]\mathbb{P}[C \cap D]$$ **[8]**

2. Prove that, for Bernoulli bond percolation with any parameter $p \in (\frac{1}{d-1}, 1)$ on d-regular tree ($d \geq 3$), $\mathbb{P}_p[N = \infty] = 1$, where N is the number of infinite open connected components. ($d \geq 3$) **[8]**

3. Fix any $d \in \mathbb{N}$. For Bernoulli bond percolation on $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$, we define the following critical threshold:

$$p_c^m = sup\{p : \mathbb{E}[|C(O)|^m] < \infty\},$$

where $C(O)$ is the open connected cluster containing origin and $m \in \mathbb{N}$. Show that $p_c^m = p_c(\mathbb{Z}^d)$ for any m. **[9]**

4. Let $B(n) = [-n, n]^d$, and $F_1(n), F_2(n), \ldots, F_{2d}(n)$ be $2d$ faces of it. Show that, for Bernoulli bond percolation on $\mathbb{L}^d$ with a parameter p such that $\theta(p) > 0$ and any $n > m$,

$$\lim_{m\to\infty} \mathbb{P}_p[\bigcap_{i=1}^{2d}\{B(m) \text{ is connected by open path to } F_i(n)\}] = 1$$

Using this, and uniqueness of the infinite cluster, deduce that under the same assumptions, $\lim_{n\to\infty} \mathbb{P}_p[LR(n)] = 1$, where $LR(n)$ means two opposite faces of $B(n)$ (in some prefixed direction) are connected by an open path. **[5+5=10]**

5. Let T be a rooted tree that is defined as follows. t_0 is the root and it has two children, each of which has 3 children and so on (any k-th generation child has $k+2$ children, the root being 0-th generation). Find out $p_c(T)$. **[10]**

INDIAN STATISTICAL INSTITUTE
M.STAT Second Year
2017-18 Semester II

Computational Finance
Midterm Examination

Date: 27/02/2018

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

1. (5+15+10) 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 $\Delta t \to 0$ and to find the Binomial option price numerically.
   (a) Write down the relationship induced by no arbitrage, between prices of an option 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 the relationship in (a) leads to the Black Scholes PDE.
   (c) Write a program to implement the Binomial model above to price a CALL option. The inputs are the asset price S_0, the strike price K, the riskless interest rate r, the current time t and the maturity time T, the volatility σ, and the number of steps n.

2. (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.

3. (10) Describe the Mersenne twister random number generator algorithm.

4. (15) Obtain the price of a geometric average PUT option analytically when the underlying STOCK follows geometric Brownian motion.

5. (10+10) Let f be a twice continuously differentiable function on $[a, b]$. Suppose $f(x) = P_1(x) + E(x)$, where P_1 is the polynomial of degree one in the trapezoid rule. Show that $| E(x) |\leq f''(\xi) * (x - a)(x - b)/2$ with $\xi \in (a, b)$. Assume that $|f''| \leq M$ is bounded. Then show that $| \int_a^b E(x)dx |\leq= M/12(b - a)^3$.

6. (5+10) Consider the random vector (Y_1, Y_2, Y_3) with expectation $(\theta_1, \theta_1 + \theta_2, \theta_2)$, var($Y_i$)=1, $i = 1, 2, 3$, cov(Y_i, Y_j)=-1/2, for $i \neq j$. The objective is to estimate θ_1. One unbiased estimator is Y_1. Use the idea of control variates to propose another estimator and compare the number of simulations required for the same degree of accuracy. [It is assumed that one simulation is generation of one observation from the joint distribution of (Y_1, Y_2, Y_3) and takes the same computational effort at generating one observation from the marginal distribution of Y_1.].

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination : Semester II (2017-2018)

Course Name : BSDA (M. Stat. 2nd year)

Subject Name : Clinical Trials

Date : 27.02.2018 , Maximum Marks : 40. Duration : 2 hrs.

1. Define type I error spending function $\alpha^*(t)$. Discuss how type I error spending function can be derived for a simple Brownian Motion process.

   Suppose for all s, t such that $0 \leq s < t \leq 1$, the type I error in the interval $(s, t]$ will be proportional to $(t-s)^\delta$ for some real δ. Find all possible values of δ and interpret. [3+8+7]

2. Suppose three treatments A, B and C are being compared in a clinical trial. The first patient is treated randomly by choosing any treatment with same probability. For any subsequent patient i, if the response of the $(i-1)$th patient is a success, we treat the ith patient by the same treatment as the $(i-1)$th patient. If the response of the $(i-1)$th patient is a failure, we treat the ith patient by any of the remaining two treatments by tossing a fair coin. If the success probabilities of the three treatments are 0.7, 0.4 and 0.3 respectively, find the probabilities that the 4th patient is treated by A and the 5th patient results in a success.

   [12]

3. Discuss Ehrenfest Urn Design for two treatments A and B. Find the conditional probability of allocating any patient to treatment A. Suggest a possible extension of the design for three treatments. [5+2+3]

Indian Statistical Institute
Second Semestral Examination 2016-17
M. Stat. II yr

Statistical Inference III

Date: February 28, 2018 (14:30 hrs) Maximum marks: 60 Duration: 2.5 hrs.

Answer all Questions. ***Question paper carries 70 marks.***

1. Define the deficiency $\delta(\mathcal{E}, \mathcal{F})$ between two experiments according to Theorem 2 of Le Cam and Yang (as discussed in class). Show that there exists a transition kernel K^* which minimizes the deficiency criterion in case of experiments with finite sample spaces.

[8+12 =20]

2 (a) Let the experiment $\mathcal{E}_0$ be defined as an infinite sequence of Bernoulli trials $\{X_1, X_2, X_3, \ldots\}$ with unknown success probability $0 < \theta < 1$. Consider two sub-experiments $\mathcal{E}_n$ and $\mathcal{E}^k$, $n \geq 1, k \geq 1$, where $\mathcal{E}_n$ is the first n outcomes from $\mathcal{E}_0$, while $\mathcal{E}^k$ is the sequence of outcomes $\{X_1, \ldots, X_\tau\}$ stopped at the random index τ where the k successes occur for the first time. Show that $\delta(\mathcal{E}_n, \mathcal{E}^k) = 0$ if $n < k$.

(b) Calculate the deficiency (using the formulation in Theorem 2 of Le Cam and Yang or otherwise) $\delta(\mathcal{E}_2, \mathcal{E}^1)$.

[15+ 10=25]

3 (a) Describe the three basic principles of statistical inference namely, Sufficiency, Conditionality and Likelihood, and the Evidence with appropriate examples. (*Answer should be brief and to the point.*)

(b) Consider a statistical model $(\Omega, \mathcal{A}, \mathcal{P})$ in standard notations. A statistic T is pairwise sufficient for $\mathcal{P}$ if it is sufficient for all two point submodels $\{P_0, P_1\} \subset \mathcal{P}$. Show that if $\mathcal{P}$ is countable and T is pairwise sufficient then it is sufficient for $\mathcal{P}$. Under what condition the previous conclusion remains valid for pairwise sufficiency if $\mathcal{P}$ is uncountable?

[10+ 15=25]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination : 2017-18

MStatII Directional Data Analysis

5 March 2018 Maximum Marks: 15 Duration 2.5 hrs

Note: Attempt all questions

1. (4 +4) (a) Prove that the "circular" distribution derived from the linear distribution through its characteristic function is the same as the corresponding wrapped distribution. Prove the result(s) you may need for the above.

   (b) Let the pdf of a circular r.v. Θ be given as,

   $$f(\Theta) = K.\,[\,1 + \delta \cos \Theta + \eta \sin^2 \Theta].$$

   Obtain K. Using the Fourier series representation of $f(\Theta)$ or otherwise, identify δ and η in terms of the trigonometric moments of Θ. Hence or otherwise derive consistent estimators of δ and η, when a random sample of n observations is obtained from $f(\Theta)$.

2. (4+3) (a) Derive the general Fourier series representation of the pdf of any circular r.v. Θ.

   (b) Using Möbius transformation, obtain the pdf of the linear r.v. corresponding to that of the circular r.v. having the pdf $f(\Theta)$ given in Problem 1(b) above.

****

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Supplementary Examination: (2017–2018)

M. Stat Second Year

Inference for High Dimensional Data

Date: 16/04/2018 Marks: 40 Duration: 2 hours.

Attempt all questions

1. (a) For $W_1, \ldots, W_p \overset{iid}{\sim} N(0, 1)$, for $\alpha > 0$ and as $p \to \infty$,

$$P\left(\max_{j=1,\ldots,p} |W_j| \geq \sqrt{\alpha \log(p)}\right)$$
$$= 1 - \exp\left(-\sqrt{\frac{2}{\alpha\pi}} \frac{p^{1-\frac{\alpha}{2}}}{(\log p)^{1/2}} + O\left(\frac{p^{1-\frac{\alpha}{2}}}{(\log p)^{3/2}}\right)\right).$$

(b) Consider a simple normal linear model given by $Z_j^{(i)} = \theta_j + \epsilon_j^{(i)}$; $j = 1, \ldots, p$; $i = 1, \ldots, n$, with $\epsilon_j^{(i)} \overset{iid}{\sim} N(0, 1)$, corresponding to $i = 1, \ldots, n$ individuals and $j = 1, \ldots, p$ genes. Assume that the goal is to detect those θ_j such that $\theta_j \neq 0$.

(i) Devise a simple statistical method in this regard.

(ii) In the light of 1. (a), discuss the consequences of your method when $p \to \infty$.

[5+2+5=12]

2. (a) What is variation sparsity?

(b) Show that variation sparsity can be represented in terms of a linear regression model.

(c) Define the collection of models in this case, with explicit mathematical details.

[4+4+4=12]

3. Consider the linear model setting $Y_i = f_i^* + \epsilon_i$; $i = 1, \ldots, n$, with $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. Let $\mathbf{Y}$, $\mathbf{f}^*$ denote the vectors gathering the components Y_i and f_i^* respectively. Let $\hat{\mathbf{f}}_m$ denote the projection of $\mathbf{Y}$ on the linear space S_m spanned by a specific set of m co-ordinates. Let

$$\hat{m} \in \arg\min_m \left\{ \|Y - \hat{f}^m\|^2 + \sigma^2 pen(m) \right\},$$

where $pen(m) = K\left(\sqrt{d_m} + \sqrt{2\log\left(\frac{1}{\pi_m}\right)}\right)^2$. Here $K > 1$, $d_m = dim(S_m)$ and $\pi_m \geq 0$ for all m such that $\sum_m \pi_m = 1$. Now consider the risk bound in the co-ordinate sparse setting:

$$E\left[\|\hat{\mathbf{f}}_{\hat{m}} - \mathbf{f}^*\|\right] \leq C_K \min_{m} \left\{ E\left[\|\hat{\mathbf{f}}_m - \mathbf{f}^*\|^2\right] + \sigma^2 \log\left(\frac{1}{\pi_m}\right) + \sigma^2 \right\}.$$

where $C_K > 1$ depending only on $K > 1$.

(a) Prove that when $\log\left(\pi_m^{-1}\right) \leq d_m$, the risk of $\mathbf{f}_m$ is bounded by a constant times the risk of the oracle $\hat{\mathbf{f}}_{m_o}$.

(b) With a suitable example (you may consider $\mathbf{f}^* = \mathbf{0}$ and $pen(m) = K\left(\sqrt{d_m} + \sqrt{2\log\left(\frac{1}{\pi_m}\right)}\right)^2$, with $\pi_m = \left(1 + \frac{1}{p}\right)^{-p} p^{-|m|}$) demonstrate that the Akaike Information Criterion (AIC) performs better when p is fixed, compared to when $p \to \infty$.

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

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

In the context of 3. (b), demonstrate that BIC can perform much better than AIC when p is fixed and $n \to \infty$.

[5+5+6=16]

Indian Statistical Institute

Second Semestral Midterm Supplementary Examination 2017-18

M. Stat. II yr

Statistical Inference III

Date: April 16, 2018 | Maximum marks: 60 | Duration: 2 hrs.

Answer all Questions.

1. (a) Define the deficiency $\delta(\mathcal{E}, \mathcal{F})$ between two experiments according to Theorem 2 of Le Cam and Yang (as discussed in class). Let $(\mathcal{X}, \mathcal{A}, P_\theta)$ be an experiment (denoted by $\mathcal{E}$) and $(\mathcal{S}, \mathcal{B}, P_\theta^T)$ be a subexperiment induced by a statistic T on $\mathcal{X}$ and a uniform random variable U which is independent of $\mathcal{E}$ (denoted by $\mathcal{F}$). If the defficiencies $\delta(\mathcal{E}, \mathcal{F}) = \delta(\mathcal{F}, \mathcal{E}) = 0$, is the statistic T sufficient? If yes, justify your answer, else provide a counterexample.

   (b) Consider the experiment $\mathcal{E}$ in (a). Suppose there is a countable set of values $\{\theta_i\}$ and a sequnce of positive constants $\{c_i\}$ satisfying $\lambda(A) = \sum_1^\infty c_i P_{\theta_i}(A) = 0$ if and only if $P_\theta(A) = 0$ for every θ and $A \in \mathcal{A}$. Then show that a statistic T is sufficient in $\mathcal{E}$ if
$$E_\theta\, u(X) = E_\lambda\, u(X) g(\theta, T(X))$$
for some measurable and integrable, nonnegative kernel g on $\Theta \times \mathcal{X}$ for every bounded measurable statistic $u : \mathcal{X} \longrightarrow \mathbb{R}$.

[15+20 =35]

2 (a) Describe the three basic principles of statistical inference namely, Sufficiency, Conditionality and Likelihood, and the Evidence with appropriate examples. (*Answer should be brief and to the point.*)

   (b) Show that the likelihood principle implies and is implied jointly by the sufficiency and conditionality principles.

[10+ 15=25]

Indian Statistical Institute

M.Stat. Second year, Second Semester Examination: 2017-18

Topic: Statistical Computing II

Maximum Marks: 50, Duration: 3 hours 23/04/2018

1. Consider a multiple linear regression model with n subjects and p predictors: $Y = X\beta + \epsilon$, where $\epsilon \sim N(0, \sigma^2 I)$.

   (a) For $n > p$, following Park and Casella (2008), develop a Bayesian hierarchical model for variable selection through LASSO by considering a Laplace prior on the regression coefficients. Discuss the Bayesian estimation method for model parameters by Markov Chain Monte Carlo (MCMC).

   (b) For $n < p$, it is known that LASSO consistently estimates maximum upto n predictors. Thus for ultrahigh dimensional data LASSO provides poor estimates. Propose a suitable two-stage method for variable selection based on LASSO for ultrahigh dimensional case. [(6+4)+5]

2. (a) Define a semi-parametric model. Show that Cox Proportional Hazards model is a semi-parametric model.

   (b) Suppose we observe a dataset with n data points $Y_1, Y_2, \ldots, Y_n$. We assume that these data points come from a mixture of three multivariate normal densities, where μ_k, Σ_k, and π_k respectively denote the mean, covariance matrix and mixing proportion of the k-th component, k=1,2,3.

   Develop an EM algorithm for estimating the model parameters. Explicitly write down E-step and M-step, and give closed form expression of the estimates.

   (c) Write down the stick-breaking representation of Dirichlet Process (DP). Discuss the role of concentration parameter α in this representation. [3+8+4]

3. Financial regulations require banks to report their daily risk measures called Value at Risk (VAR). Let Y be the financial return of the bank and then for a given θ $(0 < \theta < 1)$, the VAR is the value y^* satisfying $P(Y \leq y^*) = \theta$. The financial return depends on exchange rate (x).

   (a) Based on a sample of n data points (Y_i, x_i), we need a suitable statistical approach for estimating VAR at θ=0.85. Suggest an appropriate statistical model for this problem. Derive a simple Gibbs sampler algorithm for the parameter estimation.

   (b) Suppose now the problem is to estimate VAR at θ=0.25, 0.50, 0.65, 0.80, and 0.95. Suggest a model based approach which can simultaneously estimate all these five quantiles and can also avoid quantile crossing. [5+5]

4. We need to analyze data from a clinical trial to compare three alternative dose regimens of haloperidol for schizophrenia patients. Sixty-five patients with diagnosis of

schizophrenia were assigned to receive 5, 10, or 20 mg/day of haloperidol for 4 weeks. The outcome variable Y was the Brief Psychiatric Rating Scale Schizophrenia (BPRSS) factor, measured at $j = 1$ (baseline), $j = 2$ (week 1), and $j = 3$ (week 4). The main parameters of interest were the average change in BPRSS between baseline and week 4 for each dose group. Twenty-nine patients dropped out of the study at $j = 3$, with dropout rates varying across dose groups. Accordingly, the missingness indicator $R_i = 1$, if Y_{i3} is observed; and $R_i = 0$, if Y_{i3} is missing. A poor BPRSS outcome may cause patients to leave the study, particularly if combined with unpleasant side effects associated with the drug (particularly at high doses).

Develop a suitable Pattern-Mixture Model for the above dataset. Write down the likelihood function and discuss the parameter estimation based on the likelihood. [10]

Indian Statistical Institute
Second Semestral Examination 2017-18
M. Stat. II yr

Statistical Inference III

Date: April 25, 2018 (14:30 hrs) Maximum marks: 100 Duration: 3 hrs.

Answer all Questions

1 (a) State the three principles of foundation of inference, namely: sufficiency, conditionality and likelihood principles (stating the context clearly)

(b) Show that in discrete cases sufficiency and conditionality principles imply and are implied by the likelihood principle.

[8+12 =20]

2 (a) Define notions of specific sufficiency, θ-oriented statistic and partial sufficiency in a model (Ω, A, M) where $M = \{P_{\theta,\phi}\}$ with ϕ being a nuisance parameter.

(b) Consider the statistical model where n observations are iid from $\mathrm{U}(\theta-\sigma, \theta+\sigma), \theta \in \mathbb{R}, \sigma > 0$. Obtain a minimal sufficient statistic of the form $S = (U, V)$ where V is σ-oriented. Also find conditionally exact confidence intervals for θ using conditional distribution of U given V.

[6+ 14=20]

3 (a) Define an invariantly sufficient statistic in a parametric statistical model which is closed under the action of some group G. Consider a model with discrete sample space, $X_\Theta = (\mathcal{X}, \mathcal{A}, P_\Theta)$, closed under the action of a finite group G which induces a group G_s on the range (sample space) of some sufficient statistic S. Let u be a maximal invariant for the induced group G_s on the range of S. Prove that $V = u \circ S$ is invariantly sufficient for X_Θ under G. (Hint: Arguments provided in Hall, Wijsman and Ghosh)

(b) Let $X_1, X_2, \ldots, X_n$ be iid on $(0,1)$ and consider testing the hypothesis H_0 that the common distribution is $\mathrm{U}(0,1)$ against a two point alternative $H_1 = \{p_1, p_2\}$ where $p_2(x) = p_1(1-x)$ for some known density p_1 on $(0,1)$. Show that there exists a UMP invariant test in the problem and find the exact general form of the optimal test statistic.

[20+15=35]

4. In a multiple testing problem discuss the notion of false discovery rate (FDR) along with arguments in support of FDR over FWER. (answer must be brief, to the point and must contain a motivation for considering FDR). Describe the Benjamini-Hochberg procedure for multiple testing in the independent case. State the main theorem on the validity of BH procedure in controlling FDR from the flagship article mentioned in the class. [15]

P.T.O

5. Let d be a measure of distance of an alternative θ from a given hypothesis H. A level α test is ϕ_0 is said to be locally most powerful (LMP) if, given any other level α test ϕ, there exists a $\Delta > 0$ (may depend on ϕ) such that $\beta_{\phi_0}(\theta) \geq \beta_{\phi}(\theta)$ for all θ with $0 < d(\theta) < \Delta$. An LMP test is locally uniformly most powerful (LUMP) if there exists a $\Delta > 0$ such that $\beta_{\phi_0}(\theta) \geq \beta_{\phi}(\theta)$ for all θ with $0 < d(\theta) < \Delta$ for any level α test ϕ.

Let the data (X, Y) be independent Poisson random variables with means λ and $\lambda + 1$ respectively ($\lambda > 0$). It is desired to test $H_0 : \lambda = \lambda_0$ against $H_1 : \lambda > \lambda_0$. Determine the LMP test for this problem and show that the LMP test is not LUMP.

[10]

MSTAT II - Stochastic Calculus for Finance

Final Exam. / Semester I 2017-18

Date - April 25, 2018

Time - 3 hours/ Maximum Score - 50

<u>NOTE</u> : **THE PAPER HAS QUESTIONS WORTH 56 MARKS. SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.**

1. (a) (4 marks) Let $\{B_t\}_{t\geq 0}$ be a standard Brownian motion. Define a process $\{X_t\}_{t\geq 0}$ as follows:

$$X_t = \begin{cases} tB_{\frac{1}{t}} & \text{for } t \neq 0 \\ 0 & \text{for } t = 0. \end{cases}$$

Show that $\{X_t\}$ is a standard Brownian motion.

(b) (4 marks) Let $a > 0$ be a real number. Define, $\Lambda = \sup\{t > 0 : B_t = at\}$. Here Λ is the last time, the (standard) Brownian motion $\{B_t\}$ touches/crosses the line $y = at$. Find the distribution of Λ.

(c) (5 marks) Let $\sigma > 0$. For $t \geq 0$, define,

$$M_t = e^{\sigma B_t - \frac{\sigma^2 t}{2}}.$$

Let M_t be the price of a US dollar in terms of Euro at time t (assuming both of their domestic interest rates same). Let $\sigma = 0.35$ be the volatility per annum for the process. Find the probability that the price of a US dollar would never be more than one Euro after 3 years from its birth (i.e., $t = 0$).

2. An oil refinery in India plans to buy crude oil for their refinery, which is denominated in US dollar. Assume that the price of one barrel of oil at time t is S_t and it follows the model in the risk-neutral world,

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where $\{W_t\}$ is a standard Brownian motion, $\mu = r_d$ is the risk-free interest rate in the country that oil is being produced and σ is the constant volatility of the oil price during the period considered. Since the refinery is India-based they have to purchase dollar with Indian rupees only and then settle the bill for oil. Let P_t be the price of a dollar in rupees and it follows the model in the risk-neutral world,

$$dP_t = \mu_0 P_t dt + \sigma_0 P_t dB_t,$$

P. T. O

where $\{B_t\}$ is a standard Brownian motion and for any $s < t$, $COV((B_s, W_s)', (B_t, W_t)) = \begin{pmatrix} s & \rho s \\ \rho s & s \end{pmatrix}$ where $\rho \in (-1,1)$, with $\mu_0 = r_i - r_u$ = difference between the risk-free interest rate in India and the risk-free interest rate in the US, whereas σ_0 is the constant volatility of the dollar (with respect to the rupees) during the period considered.

(a) (8 marks) Let $Z_t = P_t S_t$ be the price of the stock in rupees. Use Itô's formula to show that Z_t satisfies,

$$dZ_t = \mu_1 Z_t dt + \sigma_1 Z_t d\hat{W}_t,$$

for some μ_1, σ_1 and a standard Brownian motion $\{\hat{W}_t\}$. Express μ_1 and σ_1 in terms of the above parameters and $\{\hat{W}_t\}$ in terms of the above parameters and standard Brownian motions $\{B_t\}$ and $\{W_t\}$.

(b) (8 marks) To hedge against the price of the oil, the refinery buys a derivative on oil that pays off $Rs.(Z_t - 5000)$ at maturity time T if $Z_T \geq 5000$, and pays nothing otherwise. Draw a diagram for the profit (against Z_T). Use the risk-neutral valuation to calculate the price of the derivative at time 0, when $r_i = 7\%$, $r_u = 1\%$, $r_d = 5\%$, $\sigma = 0.3$, $\sigma_0 = 0.4$, $\rho = 0.1$, $S_0 = 60$ dollar, $P_0 = Rs.60$ and $T = 3$ months.

3. For $i = 1, 2, 3$, let $\{W_{i,t}\}$ be three independent standard Brownian motion defined on a probability space. Let us denote its natural filtration by $\{\mathcal{F}_t\}$. Let $R_t = \sqrt{W_{1,t}^2 + W_{2,t}^2 + W_{3,t}^2}$. Define $Z_t = 1/R_t$, for all $t \geq 1$.

   (a) (4 marks) Find dZ_t.

   (b) (4 marks)) Show that $\sup_{t \geq 1} E(Z_t^p) < \infty$, for each $p \in (0, 3)$.

   (c) (4 marks) Is $\{Z_t\}$ a martingale with respect to $\{\mathcal{F}_t\}$? Justify your answer.

4. (5+5+5=15 marks) Write TRUE or FALSE and justify your answer.

   (a) Let $a < b$ be two reals, $\{B_t\}$ be a standard Brownian motion and $\{\Pi_n\}$ be a sequence of nested partitions of $[a, b]$ where Π_n is given by, $a = t_0^n < t_1^n < \cdots < t_{k(n)}^n = b$ and $k(n) \to \infty$ as $n \to \infty$. Then $\sum_{i=1}^{k(n)} (B_{t_i^n} - B_{t_{i-1}^n})^2 \to (b - a)$ almost surely, as mesh of the partition, Π_n, is going to zero whenever $\bigcup_{n \geq 0} D_n$ is dense in $[a, b]$, where D_n is set of all end points of the nth partition Π_n.

   (b) Let (Y, d) be a complete separable metric space and $\mathcal{B}(Y)$ be the Borel σ-field generated by the open sets of (Y, d). Then any finite Borel measure μ defined on $\mathcal{B}(Y)$ is tight.

   (c) Let $\{B_t\}$ a standard Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$ and let $\{\mathcal{F}_t\}$ be its natural filtration. Let $\{X_t\}$ be a continuous stochastic process defined on the same probability space and adapted to the same filtration, such that $E(\int_0^t X_s^2 ds) < \infty$ for each fixed $t > 0$. Define, $Y_t = (\int_0^t X_s dB_s)$. Then quadratic variation of $\{Y_t\}$ over $[0, T]$ is $\int_0^T X_s^2 ds$.

All the best.

INDIAN STATISTICAL INSTITUTE
Second Semester Examination 2017-18

MSTAT II
Directional Data Analysis

Date: 27 April 2018 | Maximum Marks: 50 | Duration: 3 hrs

Note: Attempt all problems. Show all of your work.

1. (a) Using Mobius transformation, show how random variates from a t-distribution can be simulated from an appropriate circular distribution. [8]

   (b) Construct a family of toroidal distributions whose conditionals are von Mises distributions. Show that this family admits of a sub-family for which the sin-sin correlation is zero but the circular variables are not independent. [7]

2. (a) Derive the Locally Most Powerful Invariant test for Isotropy against the family of symmetric wrapped stable distribution. [9]

   (b) Obtain the interval of alternatives under which the test in (a) has a monotone power function. [6]

3. (a) Describe one model each for Cylindrical, Toroidal and Spherical regressions and their related estimation methods. [12]

   (b) Explain why the Cramer - von Mises functional test for Goodness-of-Fit is not applicable to a circular probability density function and prove that Watson's U^2 test overcomes this problem. [8]

****

26.4.2018.

INDIAN STATISTICAL INSTITUTE

M. Stat 2nd Year

Weak Convergence and Empirical processes

Date: 02.05.2018 Maximum Marks: 70 Duration: 3 hours

Important Note: You can use **handwritten class notes** for your reference during the examination.

(1) Let $B_t = (1+t)W^{\circ}_{t/(t+1)}$ where W° is the Brownian Bridge on $[0,1]$. Show that process $(B_t)_{t\geq 0}$ has continuous paths, the finite dimensional distributions are Gaussian and $\mathsf{E}[B_t] = 0$ and $\mathsf{E}[B_t B_s] = s \wedge t$ and hence process represents Brownian motion on $[0, \infty)$. [10 pts]

(2) Let

$$\omega'_x(\delta) = \inf_{\{t_i\}} \max_{1\leq i\leq v} \sup_{s,t\in[t_i,t_i-1)} |x(s) - x(t)|,$$

where infimum is taken over all δ spare sets $\{t_i\}$. Show that for $x : [0,1] \to \mathbb{R}$, $x \in D[0,1]$ if and only if $\omega'_x(\delta) \to 0$ as $\delta \to 0$. Let $a < 1$, find out what $\omega'_x(\delta)$ is for $x = \mathbb{1}_{[0,a)}$. [6+4=10 pts]

(3) Let $X_1, X_2, \cdots$, be iid (on some probability space) having mean 0 and finite variance σ^2. Let $S_n = X_1 + \cdots + X_n$ for $n \geq 1$ and $S_0 = 0$. Let

$$Z_n(t) = \frac{S_{\lfloor nt \rfloor}}{\sigma\sqrt{n}} + (nt - \lfloor nt \rfloor)\frac{X_{\lfloor nt \rfloor + 1}}{\sigma\sqrt{n}}.$$

Assuming the Donsker's theorem done in class, that is, $Z_n \Rightarrow W$, and the derivation of $\sup_{t\in[0,1]} W_t$ show that $\sup_{t\in[0,1]} |W_t|$ has finite fourth moment. Hence using this show that

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \lambda^2 \mathsf{P}(\max_{k\leq n} |S_k| \geq \lambda\sigma\sqrt{n}) = 0.$$

[7+8=15 pts]

(4) If P_n, P are all probabilities on $D[0,1]$ and suppose you only know $P_n(\pi_{t_1}, \cdots \pi_{t_k})^{-1} \Rightarrow P(\pi_{t_1}, \cdots \pi_{t_k})^{-1}$ for all $(t_1, \cdots, t_k)$, where $(\pi_{t_1}, \cdots \pi_{t_k})$ denotes the coordinate projections from $D[0,1$

to $\mathbf{R}^k$. Then you cannot conclude $P_n \Rightarrow P$ on $D[0,1]$. Give an example to justify this.

[10 pts]

(5) State whether the following are true or false (with appropriate reasons):

(a) Brownian Bridge has independent increments.

(b) For $x \in D[0,1]$, the maximum jump j_x is continuous in the Skorohod topology.

(c) $\mathcal{G} = \{\mathbf{1}_{(-\infty,t]} : t \in \mathbf{R}^d\}$ is a VC (Vapnik-Chervonenkis) class of functions.

(d) Consider $C[0,1]$. For each $n \geq 1$ let f_n be the polygonal line joining

$$(0,0),(1/n,n),(2/n,0),(1,0).$$

Let μ_n be the point mass at f_n so that μ_n is a probability on $C[0,1]$. Then $(\mu_n)_{n\geq 1}$ is not a tight family.

[5+ 5+ 8+7=25pts]

Indian Statistical Institute
Semester 2, Academic Year: 2017-18
Final Examination
Course: M. Stat 2nd Year
Subject: Brownian Motion & Diffusions

Total Points: 70
Date: 4.5.2018 **Time: 2 hours 30 minutes**

Answers must be justified with clear and precise arguments. If you refer to a theorem/result proved in class, state it explicitly. More than one answer to a question will not be entertained and only the first uncrossed answer will be graded.

1. From the martingale $e^{B(t)-(t/2)}$ where $B(t)$ denotes standard Brownian motion, consider the stopped martingale $e^{B(T_a \wedge t)-(T_a \wedge t)/2}$ where for fixed $a > 0$, T_a is the hitting time of a. Using this stopped martingale at times zero and t, one gets

$$Ee^{B(T_a \wedge t)-(T_a \wedge t)/2} = 1.$$

(a) Justify why one can make $t \to \infty$ inside the expectation of the above equality and find the value of the Laplace transform $Ee^{-(T_a/2)}$.

(b) How do you get $Ee^{-\beta T_a}$ for any $\beta > 0$ and what is its value?

$10 + 10 = 20$ pts.

2. (a) Suppose f, g are nonrandom functions in $L^2[0,1]$. Consider the Wiener integrals $M_t = \int_0^t f(s)dW(s)$ and $N_t = \int_0^t g(s)dW(s)$. Evaluate $E\{[M_t.N_t - \int_0^t f(u)g(u)du] | \mathcal{F}_s^W\}$ for $s < t$.

(b)Suppose f is a bounded continuous progressively measurable process on the standard Wiener space. If $0 = t_0 < t_1 < t_2 < \ldots < t_n = t$, is a partition of $[0,t]$, show that the Riemann sums $\sum_{i=0}^{n-1} f(t_i)(W(t_{i+1}) - W(t_i))$ converge to $\int_0^t f(u)dW(u)$ in mean square as the norm of the partition goes to zero. You can assume the Ito isometry for stochastic integrals. $10 + 10 = 20$ pts.

3. (a) Suppose ϕ is a nonnegative, measurable function on $[0,1]$ that satisfies $\phi(t) \le \beta \int_0^t \phi(s_2)ds_2$, for any $0 \le t \le 1$, where β is a positive constant. If $\phi(t) \le c, 0 \le t \le 1$, then using inequalities like $\phi(t) \le \beta^2 \int_0^t \int_0^{s_2} \phi(s_1)ds_1\, ds_2$, show that $\phi(t) = 0, 0 \le t \le 1$. (Application of Gronwall's inequality will not be accepted as an answer.) **P. T. O**

(b) Consider a solution to the SDE $X_t = \xi + \int_0^t \sigma(X_s)dW_s$, over $[0, T]$ where $|\sigma(x) - \sigma(y)|^2 \leq K.|x - y|^2$ for some positive constant K. You may assume that the solution satisfies $\int_0^T EX_t^2 dt < \infty$. If Y_t is another such solution with the same properties, then using $\sup_{0\leq t\leq t'} |X_t - Y_t|^2 = \sup_{0\leq t\leq t'} |\int_0^t \{\sigma(X_s) - \sigma(Y_s)\}dW_s|^2$ prove the following inequality 5 + 15 = 20 pts.

$$E(\sup_{0\leq t\leq t'} |X_t - Y_t|^2) \leq 4K \int_0^{t'} E(\sup_{0\leq s\leq t} |X_s - Y_s|^2)dt.$$

4. For each of the following two processes satisfying the SDE mentioned

(i) the Ornstein-Uhlenbeck process satisfying $dX_t = -X_t dt + dW_t$,

(ii) the geometric Brownian motion satisfying $dX_t = \mu X_t dt + \sigma X_t dW_t$, $\mu \in \mathbb{R}, \sigma > 0$,

(a) Write down the generator of the process taking x as the space variable.

(b) Assuming $X_0 = \xi$, write down the Ito formula for $f(X_t) - f(X_0)$ where f is C_b^2. 5 + 5 = 10 pts.

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2017–2018)

M. Stat Second Year

Inference for High Dimensional Data

Date: 04.05.18 Marks: ..60.. Duration: .3 hours.

Attempt all questions

1. (a) For $W_1, \ldots, W_p \stackrel{iid}{\sim} N(0,1)$, for $\alpha > 0$ and as $p \to \infty$, prove that

$$P\left(\max_{j=1,\ldots,p} |W_j| \geq \sqrt{\alpha \log(p)}\right)$$
$$= 1 - \exp\left(-\sqrt{\frac{2}{\alpha\pi}} \frac{p^{1-\frac{\alpha}{2}}}{(\log p)^{1/2}} + O\left(\frac{p^{1-\frac{\alpha}{2}}}{(\log p)^{3/2}}\right)\right).$$

(b) Consider a simple normal linear model given by $Z_j^{(i)} = \theta_j + \epsilon_j^{(i)}$; $j = 1, \ldots, p$; $i = 1, \ldots, n$, with $\epsilon_j^{(i)} \stackrel{iid}{\sim} N(0,1)$, corresponding to $i = 1, \ldots, n$ individuals and $j = 1, \ldots, p$ genes. Assume that the goal is to detect those θ_j such that $\theta_j \neq 0$.

(i) Devise a simple statistical method in this regard.

(ii) In the light of 1. (a), discuss the consequences of your method when $p \to \infty$.

[6+4+6=16]

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* $N(0, \sigma^2)$. 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 $\mathbf{f}^* \in S_{m^*}$ with $|m^*| = D^*$.

(a) Writing $\mathbf{P}_j$ for the projection on the line spanned by $\mathbf{X}_j$, prove that

$$\|\hat{\mathbf{f}}_{\hat{m}_\lambda} - \mathbf{f}^*\|^2 = \|\mathbf{f}^* - \sum_{j \in \hat{m}_\lambda} \mathbf{P}_j \mathbf{f}^*\|^2 + \sum_{j \in \hat{m}_\lambda} Z_j^2 \sigma^2$$
$$\geq \sum_{j \in \hat{m}_\lambda \setminus m^*} Z_j^2 \sigma^2 \geq (|\hat{m}_\lambda| - D^*) \lambda \sigma^2,$$

where, for any m, $\hat{\mathbf{f}}_m$ is the projection of $\mathbf{Y}$ on S_m and $\hat{m}_\lambda = \left\{j : \left(\mathbf{X}_j^T \mathbf{Y}\right)^2 > \lambda \|\mathbf{X}_j\|^2 \sigma^2\right\}$.

(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 $E[|\hat{m}_\lambda|] \geq pP\left(\tilde{Z}^2 \geq \lambda\right)$, 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 $D \in \{1, \dots, p\}$, 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 $\mathbf{f}^* \in V_{D^*}(\mathbf{X})$,

$$E\left[\|\hat{\mathbf{f}}_{\hat{m}_{2K\log p}} - \mathbf{f}^*\|^2\right] \geq R[\mathbf{X}, D^*], \quad \text{as } p \to \infty.$$

[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 $\hat{\mathbf{f}}_m$ be the projection of the response data $\mathbf{Y}_n = (y_1, \dots, y_n)^T$ 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 $\beta > 0$.

(a) Prove that the weights $\mathbf{w} = \{w_m : m \in \mathcal{M}\}$ minimizes the functional

$$G(\mathbf{q}) = \sum_{m\in\mathcal{M}} q_m\hat{r}_m + \frac{\sigma^2}{\beta}\mathcal{K}(\mathbf{q}, \boldsymbol{\pi}),$$

where $\mathbf{q} = \{q_m : m \in \mathcal{M}\}$, $\boldsymbol{\pi} = \{\pi_m : m \in \mathcal{M}\}$, and $\mathcal{K}(\mathbf{q}, \boldsymbol{\pi}) = \sum_{m\in\mathcal{M}} q_m \log\left(\frac{q_m}{\pi_m}\right)$ is the Kullback-Leibler divergence between the probabilities $\mathbf{q}$ and $\boldsymbol{\pi}$.

(b) Propose, with proper mathematical justification, an algorithm for computing $\hat{\mathbf{f}}$.

[6+6=12]

4. (a) Show analytically how the Lasso estimator selects variables in the orthonormal set-up.

   (b) Show analytically how the Elastic-Net estimator selects variables.

[6+6=12]

Statistical Methods in Epidemiology & Ecology

Date: 07.05.18

Semestral Examination

M.Stat.- II Year, 2017-2018

Total Marks - 100 Time: 3 hrs. 30 mins.

Note :

(i) Question no. 1, 4, 5, 6 are compulsory.

(ii) Additionally you need to answer either question no 2 or 3.

(iii) Total you need to answer five questions.

1. (a) Define Fisher's Relative Growth Rate (henceforth, RGR) based on the size measurement at two specific time points. Comment on its extension and growth law non-invariant form. Derive the expression of the extended RGR metric for the Gompertz growth law based on the size measurement at three consecutive time points. How is this extended metric affected by the reading/measurement error? Construct a suitable test for showing the efficacy of the extended metric over Fisher's relative growth rate based on non-overlapping time points.

   (b) Let us define, $X(t)$ be the size of a species measured at time point t. We assume that $(X(1), \dots X(q))' \sim N_q(\theta, \Sigma)$, where $E(X(t)) = \theta(t)$, an exponential quadratic growth curve profile. Suppose we interested in testing the hypothesis of the exponential quadratic growth curve model (EQGCM), i.e., to test

   $$H_0 : \theta(t) = \exp\{b_0 + b_1 t + b_2 t^2\} \text{ against } H_1 : \text{not } H_0$$

   Using the approximate expression for the 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 matrix $X(n \times q)$. Note that. in the data matrix $X(n \times q)$ any row corresponds to a q - variate size measurements available at q time points on one of the n individuals. Also, suggest required modifications of the test statistic when the size variables are non normal.

   (c) Suggest an estimate of RGR for any time interval (t, t + 1] based on the data available as in question no 1(b). Assuming suitable multiplicative error structure between observed and true size variables, find the bias and MSE of this estimate under second order of approximation (here order is defined by the "power of relative errors" for this specific time interval).

[2+(3+4+4+8)+(10+3)+(2+5) = 41]

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

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

   where a, b and c are positive constants.

   (b) Let us rewrite the equation (1) as

   $$R(t) = \frac{1}{x(t)} \frac{dx(t)}{dt} = be^{-at^c} t^{c-1} + \epsilon_t, \qquad (2)$$

   where, $R(t)$ is the empirical estimate of RGR and ϵ_t is the error of the nonlinear regression at time point t. Show that for the model (2) the nonlinear least square estimates exist, and are consistent and asymptotically normal.

[3+10=13]

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

   (b) The θ-logistic growth equation for a single species population dynamics is defined as follows:

$$\frac{dx(t)}{dt} = rx(t)\left(1 - \left(\frac{x(t)}{k}\right)^{\theta}\right), \qquad (3)$$

   where, parameters have their usual interpretations.

   Consider a stochastic model by incorporating a suitable random uctuation in the above model equation (3). Determine the expression for the approximate mean and variance of this stochastic model.

[3+10=13]

4. (a) Define Lyapunov stability.

   (b) State and prove the theorem of asymptotic stability for a single species dynamics.

   (c) Define Allee effect and critical density. The θ-logistic model as defined in equation (3) can be converted to the extended Allee effect model - Justify. Comment on the equilibrium points and stability criteria of this model without using the stability theorem.

[2+6+(3+2+4)=17]

5. (a) Write down the Kermack - Mckendrick model of general epidemics.

   (b) What are the basic hypotheses of the model ?

   (c) Deduce the Kermack - Mckendrick threshold phenomenon and state clearly the justification of such threshold phenomenon in order to build up epidemic.

   (d) Prove that the spread of the disease will not stop for the lack of susceptible populations.

[3+3+5+4+4=19]

6. (a) Stating proper assumptions write down the deterministic model of the HIV infection based on the dynamics of CD4+T' cells and the free HIV virus.

   (b) Construct the transition probabilities for the stochastic analogue of this deterministic model as proposed in 6(a).

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

[3+3+4=10]

INDIAN STATISTICAL INSTITUTE

M.STAT Second Year

2017-18 Semester II

Computational Finance

Final Examination

Date: 08/05/2018

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. (10+5) 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 $\int_a^b f(x)w(x)dx = \sum_{i=1}^n w_i f(x_i)$, where given a non-negative weight function $w(x)$, the nodes x_i and weights w_i have to be chosen properly. What are the proper weights when $w(x)$ is the Gaussian weight function?

2. (8+5+7) We are interested in generating the path of a Brownian motion at time points $0 < t_1 < t_2 \cdots < t_n$

   (a) Describe the Cholesky construction by explicitly deriving the lower triangular matrix A such that $AA^T = C$, where C is the covariance matrix.

   (b) How does the random walk construction use the form of A above to reduce the number of computations from $O(n^2)$ to $O(n)$?

   (c) Write one advantage and one disadvantage of using the principal component construction over the random walk construction.

3. (10+10) Let C denote the copula of the two random variables X and Y. Assume that the marginal cdfs are continuous and strictly increasing.

   (a) Show that $P[\max(X,Y) \le t] = C(F_X(t), F_Y(t))$

   (b) Prove that the Spearman correlation coefficient $\rho(X,Y)$ is given by the formula

$$\rho(X,Y) = 12\int_0^1\int_0^1 uvC(u,v)dudv - 3$$

4. (8+7+10) Consider the generalized Vasicek family below to model the term structure of interest rates:

$$Y(x) = \theta_1 - \theta_2\theta_4\frac{1-e^{-x/\theta_4}}{x} + \theta_3\theta_4\frac{(1-e^{-x/\theta_4})^2}{4x}$$

   (a) Derive an analytical formula for the instantaneous forward rate.

   (b) Derive an analytical formula for the price of zero coupon bonds.

   (c) Given data on the price of bonds of different maturities for a series of days, how would you forecast the prices for the next day using this model and an AR(1) structure on the evolution of the parameters.

P.T.O

5. (7+8+5) Suppose ϵ_t is stationary strong ARCH(1) process; that is $\mathrm{E}(\epsilon_t|\mathcal{F}_{t-1}) = 0$, $\mathrm{Var}(\epsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2$.

   (a) Show that ϵ_t is white noise.

   (b) In addition, assume that $\mathrm{E}(\epsilon_t^4) = c < \infty$ and $Z \sim \mathcal{N}(0,1)$. Show that ϵ_t^2 follows an AR(1) process.

   (c) What is the main difference between GARCH-type models and Stochastic Volatility models?

INDIAN STATISTICAL INSTITUTE

End-Semester Examination : 2017-18

Percolation Theory

M.Stat 2nd year

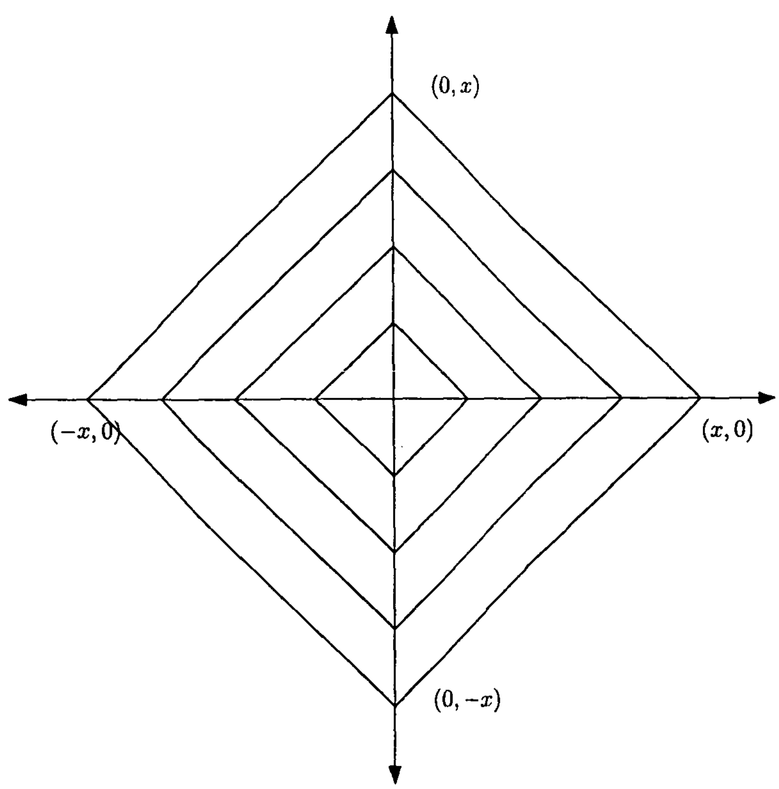
9^{th} May, 2018

Maximum marks: 50 Duration: 3 hr 30 min

1. Let A and B be increasing and decreasing events respectively, both depending on finitely many edges. Show that $A \circ B = A \cap B$, where $\circ$ indicates disjoint occurrence. [8]

2. Let $\mathbb{S}_k = \mathbb{Z}^2 \times \{0, 1, \dots, k\}$. Show that $p_c(\mathbb{S}_k) > p_c(\mathbb{S}_{k+1})$. [8]

3. Let us construct the following graph H with vertex set $V = \{(x, y) : xy = 0, x, y \in \mathbb{Z}\}$. (a, b) is an edge of H if either $|a| = |b|$ or $|a - b| = 1$ for $a, b \in V$. Find out $p_c(H)$. [8]



4. Assuming that $\theta(p_c) > 0$ for Bernoulli bond percolation on $\mathbb{Z}^3$, show that with probability 1, any infinite path from origin has to intersect every octant infinitely often. [8]

P.T.O

5. G be any bounded degrees infinite connected graph with $0 < p_c(G) < 1$. For some parameter $p > 0$, you decide to observe N outcomes of the Bernoulli bond percolation on G and superimpose them, i.e. construct a resultant graph where an edge of G is open if it was also open in any of the N outcomes. Show that, given any p, you can find with probability 1, an open infinite cluster in the resultant graph for N large enough. Can you write down minimum such N, as a function of p and $p_c(G)$? **6+2=8**

6. Consider critical Bernoulli bond percolation on $\mathbb{Z}^2$. Let us call $B(n) = [-n, n]^2$, $L(n) = \{-n\} \times [-n, n]$, $R(n) = \{n\} \times [-n, n]$, and $\alpha(n) = \mathbb{P}_{p_c}[O \rightsquigarrow \partial B(n)]$. We define the horizontal crossing collection as the following:

$$HC(n) = \{v \in B(n) : L(n) \rightsquigarrow v \rightsquigarrow R(n) \text{ by open paths inside } B(n)\}.$$

   (a) Using Russo-Seymour-Welsh Theorem, or otherwise, show that $\exists\ C > 0$ such that $\alpha(n/2) \leq C\alpha(n)$.

   (b) Prove that $\exists\ C' > 0$ such that $\mathbb{E}_{p_c}[|HC(n)|] \leq C'n^2\alpha(n)$. **6+4=10**

7. Consider super-critical Bernoulli bond percolation on $\mathbb{Z}^2$. Let $L_{m,n}$ be the event that the rectangle $[0, m] \times [-n, n]$ contains an open vertical crossing.

   (a) Using duality or otherwise, show that $\exists$ constants $A, \sigma(p) > 0$ such that $\mathbb{P}_p[L_{m,n}] \geq 1 - Ane^{-\sigma(p)m}$.

   (b) Given any $\epsilon > 0$, show that there exists ν large enough, such that any vertex of $B(n - \nu \log n)$ which is in an infinite cluster has to be in the crossing cluster of $B(n)$ with probability exceeding $1 - \epsilon$. **8+7=15**

You are allowed to use your own handwritten notes.

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2017-2018

PROGRAMME NAME: BSDA (M. STAT. 2nd YEAR)

Course Name: Clinical Trials

Date: 09/05/2018 , Maximum Marks: 60. Duration : 3 hrs

1. What is Selection Bias in clinical trials? Illustrate Selection Bias with an example. [3+12]

2. What is a play-the-winner rule? Obtain the unconditional probability of success by the n-th patient in a play-the-winner design. [4+11]

3. Using the idea of a delayed response indicator and stating the assumptions, obtain the transition probability matrix when both the response time and interarrival times are exponentially distributed. [12]

4. The success probabilities of two treatments are p_A and p_B. Assuming that the prior distribution of p_A is such that p_A can take the values a and b $(a < b)$ with probabilities $1/3$ and $2/3$. The prior for p_B is same, and p_A and p_B are independent. If n_k patients are treated by treatment k and s_k successes and f_k failures are observed, $k = A, B$, define and obtain a suitable Bayesian adaptive allocation design for the $(n+1)$st patient, where $n = n_A + n_B$. [10]

5. Discuss how Optimal Safe Dose can be obtained in phase II clinical trials. [8]

Indian Statistical Institute
Vectors and Matrices II
B.Stat (hons) 1st year
Second Semestral Examination

Date: May 11, 2018 Duration: 3hrs.

This paper carries 55 marks. Attempt all questions. The maximum you can score is 50. Justify all your steps. This is a closed book, closed notes examination. You are not allowed to use the exercises from the class webpage without proof.

If copying is detected in the solution of any problem, all the students involved in the copying will get 0 for that problem.

1. For any real square matrix A, show that the matrices $A'A$ and AA' are similar.

   [10 marks]

2. Let A be a square matrix with nonegative real entries, with all the row sums equal to 5. Show that 5 must be an eigenvalue of A. Also show that if λ is any eigenvalue of A then $|\lambda| \le 5$. [3+7 marks]

3. Let A and B be positive definite matrices. Show that

$$\max\left\{\frac{\vec{x}'A\vec{x}}{\vec{x}'B\vec{x}} \;:\; \vec{x} \neq \vec{0}\right\}$$

   is the largest eigenvalue of AB^{-1}. [10 marks]

4. If P is an orthogonal projector then show that $P^{+} = P$. Here P^{+} denotes the Moore-Penrose pseudoinverse of P. [5 marks]

5. If the Jordan Canonical Form of a complex square matrix A contains a Jordan block of size 7 with diagonal entries 3, then show that the minimal polynomial $p(\lambda)$ of A must be divisible by $(\lambda - 3)^7$. [10 marks]

6. Submit the "QR decomposition by Householder method" project.

   [10 marks]

INDIAN STATISTICAL INSTITUTE
Semester Examination 2017-18

M.Stat. - 2nd Year
Survival Analysis

11th May, 2018

Maximum Marks: 100 Time: 4 hours

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

1. Consider randomly right-censored data from a lifetime distribution F with no ties and the problem of non-parametric estimation of its survival function $S = 1 - F$.

(a) Define self-consistency of an estimate of S in this context.

(b) Prove that the Kaplan Meier estimator $\widehat{S}_{KM}(t)$ is the unique self-consistent estimator of $S(t)$ for t less than the largest observation. [2+6]=8

2. Consider a series system with two independent components having survival functions $S_1(\cdot)$ and $S_2(\cdot)$, respectively. We have randomly right-censored data on the life time of each of the components tested separately before being assembled for the system.

(a) Obtain the nonparametric maximum likelihood estimate of the survival function $S(\cdot)$ of the system and, hence, find an estimate of its asymptotic variance.

(b) Obtain a 95% confidence band for S containing only the permissible values. [(2+5)+3]=10

3. Consider the K-sample problem with $\lambda_i(t)$ denoting the hazard for the i-th population, $i = 1, \ldots, K$. In the i-th group, the randomly right-censored data are of the form (x_{ij}, δ_{ij}), for $j = 1, \ldots, n_i$, where x_{ij} denotes the observation time and δ_{ij} the censoring indicator for the jth individual in the ith group.

(a) Describe, in detail, the log-rank test for testing homogeneity of K samples under consideration.

(b) Suppose we assume the model $\mathcal{M} : \lambda_i(t) = a_i\lambda_1(t)$, for $i = 2, \ldots, K$, with $\lambda_1(t)$ being totally unknown and arbitrary. Suggest a graphical test for the model $\mathcal{M}$.

(c) Develop a test for homogeneity under the model in (b), giving full details. [5+4+6]=15

4. Suppose that data have been generated from two groups of individuals by the Cox proportional hazards model and that we observe the following values:

$$(U_i, \delta_i, z_i) = (16,1,1), (13,1,0), (21,1,1), (11,1,0), (12,1,1), (14,0,1), (24,0,0),$$

where U_i denotes the observation time, δ_i is the censoring indicator (1 if uncensored and 0 if censored) and z_i is the group indicator (0 for Group I and 1 for Group II).

(a) Write down the expression for the Cox partial likelihood given the above data.

(b) Derive the maximum likelihood estimator (MLE) of the regression coefficient β.

(c) Derive the Breslow's estimate of the base-line hazard under appropriate assumptions (to be stated by you).

(d) How will the partial likelihood obtained in (a) change if the observation times U_is (measured in months) are recorded as $30 \times U_i$s (measured in days)?

[4+3+6+2]=15

5. Let $(t_1, z_1), \ldots, (t_n, z_n)$ be independent uncensored observations from Weibull regression model with hazard $\lambda(t, z) = \lambda t^p e^{z\beta}$, $t > 0$, where p is a known positive number and $\lambda > 0$ is the nuisance parameter.

(a) Under what group of transformations on t will the inference on β be invariant? Hence, or otherwise, give a marginal sufficient statistic for β.

(b) Obtain the marginal likelihood for β and compare it with the full likelihood.

[5+5]=10

6. Consider randomly right-censored data from the Accelerated Failure Time (AFT) model $Y = \log T = \alpha + \beta z + \epsilon$ and the associated linear rank test statistics W_n depending on two sets of constants c_i and c_i^*.

(a) Derive the general formula of the constants c_i and c_i^* for which W_n is the locally most powerful rank test (clearly state the notations and assumptions, if any).

(b) Let $u_{(1)}, \ldots, u_{(n)}$ denote the order statistics from n IID $U(0,1)$ random variables. Then, prove that

$$\int \cdots \int_{u(1)<\ldots<u(k)} \log(1-u_{(i)}) \prod_{j=1}^{k} n_j (1-u_{(i)})^{m_j} du_{(j)} = -\sum_{j=1}^{i} \frac{1}{n_j}.$$

(c) Using part (b), or otherwise, find out the locally most powerful rank test for covariate effects under the AFT model with extreme-value error distribution $f(\epsilon) = \exp(\epsilon - e^{\epsilon})$.

[13+7+5]=25

7. Consider the competing risk model given by the cause-specific hazards $\lambda_j(t,z) = \lambda_0(t)e^{\gamma_j+z\beta_j}$ for $j = 1,\ldots,m$, with $\gamma_1 = 0$ and $\lambda_0(t)$ being unknown and arbitrary.

(a) Describe the relationship between the different cause specific hazard rates for a fixed covariate z.

(b) Obtain an appropriate partial likelihood to estimate the parameters γ_j's and β_j's based on randomly right-censored data. Derive the score test for no covariate effects (i.e., $\beta_j = 0$ for all j) based on this partial likelihood.

(c) Let J denote the random variable for cause of death. Find $P[J = j; z]$ for a fixed covariate z, and hence prove that the life time T and cause of death J are independent.

(d) Let us assume that $\lambda_0(t) = \lambda$ for all t. Based on randomly right-censored data, derive the fully parametric score test for no covariate effects and compare it with the test obtained in (b). [Hint: Use the re-parametrization $\lambda_j = \lambda e^{\gamma_j}$]

[3+(4+5)+3+5]=20

8. Suppose $X_1,\ldots,X_n$ be independent non-negative life-time random variables and the distribution of each X_i is absolutely continuous with hazard rate $\alpha_i(t,\theta) = \mu_i(t)\alpha_0(t)$ for $i = 1,\ldots,n$, where μ_i are known hazard rates and α_0 is an unknown relative mortality rate common to all i.

(a) Describe this problem in terms of a suitable (multivariate) counting process model and derive its intensity process (if exists).

(b) Using the aggregated univariate process, or otherwise, derive the Nelson-Aalen estimator of $A_0(t) = \int_0^t \alpha_0(s)ds$ and its (asymptotic) variance estimate.

(c) Let us assume the parametric form $\alpha_0(t) = \alpha_0(t,\theta)$. Write down the likelihood function for θ and, hence find out a sufficient statistic (process) for θ.

[(3+6)+(5+3)+5]=22