Statistical Genomics

M-Stat (2nd Year)

End Semester Examination 2018-19

Group B

This group carries 25 marks and is an open notes examination. Answer all questions.

1. Consider two autosomal biallelic loci with alleles (*D*,*d*) and (*M*,*m*) respectively and recombination fraction θ between the loci . In a family, the father is *DdMm*, the mother is *Ddmm* while the offspring are *DDMm*, *DdMm* and *DDmm*. Assuming that both haplotype phases have equal prior probabilities, obtain the posterior probability of the father being in the coupling phase. [6]

2. Consider the following genotype data on 500 randomly selected individuals at two autosomal biallelic loci with alleles (*A*,*a*) and (*B*,*b*) respectively. Do the data provide evidence of linkage disequilibrium between the two loci? [8]

	BB	*Bb*	*bb*
AA	133	89	19
Aa	33	147	88
aa	5	16	20

---

3. Consider a dominant disorder controlled by an autosomal biallelic locus. Suppose that a marker locus is in linkage disequilibrium with the disease locus. For what value of the disease allele frequency will the difference between the marker allele frequencies among cases and controls be the maximum? [6]

4. Suppose genotype data are available on *M* randomly selected trios suitable for performing the classical TDT and on *N* randomly selected sibships (without parental information) suitable for performing the Sib TDT. How can you combine the two sets of data to perform a suitable test of linkage in the presence of association? [5]

INDIAN STATISTICAL INSTITUTE

Statistical Genomics
M-Stat (2nd Year) 2018-2019
Mid-Semester Examination

Date: 7.9.2018 **Time: 2 hours**

This is an open notes examination. The paper carries 30 marks.

1. (a) Consider a disorder controlled by L unlinked autosomal biallelic loci such that an individual is affected if and only if he/she is homozygous recessive in at least one of these loci. If both parents are heterozygous at all L loci, what is the probability that an offspring will be affected?

(b) Consider two biallelic loci with alleles (A,a) and (B,b) such that the recombination fraction between the two loci is θ. Suppose A is dominant over a and B is dominant over b. Hence, there are four possible phenotype combinations with respect to the two loci: dominant-dominant, dominant-recessive, recessive-dominant and recessive-recessive. Consider families where both parents are double heterozygous with one parent in the coupling phase and the other in the repulsion phase. Obtain the probability distribution of the different phenotype combinations of an offspring in such a family. Given the frequency distribution of the phenotype combinations of offspring in a random set of such families, describe an algorithm to obtain the maximum likelihood estimate of θ. Show all computational steps clearly.

[3+12]

2. (a) Consider two DNA sequences,

Seq1: ACACAACGG
Seq2: AAAACG.

Compute the best global alignment of the two sequences assuming match score of 1, mismatch penalty of -3, gap opening penalty of -3 and gap extension penalty of -1.

(b) Describe the alignment you would expect if the gap opening penalty was infinity. What kind of alignment would you expect if the gap start penalty and gap extension penalty were both zero?

[6+2+2]

3. (a) Let $\boldsymbol{P}$ be the PAM-1 transition matrix and p_j be the background frequency of amino acid j. Give an expression for the PAM-120 log odds scoring matrix in terms of $\boldsymbol{P}$ and p_j.

(b) Can the BLOSUM-40 matrix be derived directly from the BLOSUM-62 matrix? If not, what additional information would you need?

[2+3]

M.Stat Second Year, First Semester, 2018-19

Statistical Computing

Mid-Semestral Examination

Time: 3 hours | Full Marks =60 | Date: 03.09.2018.

[Answer as many as you can. The maximum you can score is 60.]

1. (a) Describe how will you generate an observation from a four-dimensional distribution with p.d.f. $f(\mathbf{x}) \propto \|\mathbf{x}\| \exp\{-\|\mathbf{x}\|\}$, where $\|\cdot\|$ denotes the usual Euclidean norm. [6]

   (b) Describe how you will generate an observation from the uniform distribution on the region $S = \{(x,y,z):\ |x| + 2|y| + 3|z| \leq 6\}$. [6]

   (c) In order to get observations from the uniform distribution on the d-dimensional unit sphere $S_d = \{\mathbf{x} \in R^d :\ \|\mathbf{x}\| \leq 1\}$, one generates observations from the uniform distribution on $[-1,1]^d$ and then uses the rejection method. Let X be the number of trials needed to get an observation from the uniform distribution on S_d. Show that $E(X)$ diverges to infinity as the dimension increases. [6]

2. (a) Consider a function h, which is symmetric in its arguments. Let $X_1, \ldots, X_n$ be n independent observation from a distribution F with $\int h(x,x)dFx < \infty$ and $\int h(x_1,x_2)dFx_1dFx_2 < \infty$.

   (i) Show that $T_n = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} h(X_i,X_j)$ is a biased estimator of $\theta = \int h(x_1,x_2)dFx_1dFx_2$ and its bias is of the order $O(n^{-1})$. [3]

   (ii) Show that the corresponding bias corrected jackknife estimator of θ is given by $T_{JACK} = \binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} h(X_i,X_j)$. [4]

   (b) Consider an auto-regressive model $X_t = \theta X_{t-1} + \epsilon_t;\ \ t = 1,2,\ldots,n$, where $X_0 = 1$ and the ϵ_t's are i.i.d. with mean 0 and variance 1.

   (i) Find the least square estimate of θ. [2]

   (ii) Describe how you will use the residual bootstrap method to construct a 95% confidence interval for θ. [6]

   (c) Consider a moving average model $X_t = \mu + \epsilon_t + \frac{1}{2}\,\epsilon_{t-1};\ \ t = 1,2,\ldots,n$, where $\epsilon_0, \epsilon_1, \ldots, \epsilon_n$ are i.i.d. with mean 0 and variance $\sigma^2 > 0$.

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

   (ii) Let $X_1^*, X_2^*, \ldots, X_n^*$ be a bootstrap sample generated from the empirical distribution with mass points $X_1, X_2, \ldots, X_n$. Does the bootstrap distribution of $\sqrt{n}(\bar{X}^* - \bar{X})$ give a good approximation for the distribution of $\sqrt{n}(\bar{X} - \mu)$? Justify your answer. [6]

[P. T. O.]

3. (a) Show that the influence function for the mean of a distribution is unbounded but that for the median is bounded. [3+5]

(b) Prove or disprove the following statements.

(i) If the half-space depth of a point with respect to a continuous probability distribution is less than 0.5, it cannot be the half-space median of the distribution. [3]

(ii) Half-space median of a continuous probability distribution cannot lie outside the support of the distribution. [3]

(iii) Simplicial depth is affine invariant. [3]

(iv) If the underlying distribution is elliptically symmetric with the mean $\boldsymbol{\mu}$ and the dispersion matrix $\boldsymbol{\Sigma}$, then the half-space depth of a point $\mathbf{x}$ is a decreasing function of $(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$. [4]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : 2018-2019

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 4 September, 2018 Max. Marks: 60 Duration: 2 Hours

1. Describe how the posterior distribution can be used for estimation of a real parameter. How do you measure the accuracy of an estimate?

[7]

2. What is a conjugate prior? Give an example to show that a conjugate prior can be interpreted as additional data.

[5]

3. Consider the linear regression model $\boldsymbol{y} = \boldsymbol{X\beta} + \epsilon$ 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 ϵ 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 $\boldsymbol{\beta}$.

[5+4+10=19]

4. Let $X_1, \ldots, X_n$ be i.i.d. with a common density $f(x|\theta)$ where $\theta \in R$ and $\pi(\theta)$ be a prior density of θ.

(a) 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.

(b) Show that under suitable regularity conditions, for any $\delta_0 > 0$, with probability one,

$$\int_{\{|t|>\delta_0\sqrt{n}\}} \pi(\widehat{\theta}_n + tn^{-1/2}) \exp[L_n(\widehat{\theta}_n + tn^{-1/2}) - L_n(\widehat{\theta}_n)]dt \to 0$$

as $n \to \infty$ where $L_n(\theta) = \sum_{i=1}^{n} \log f(X_i|\theta)$ and $\widehat{\theta}_n$ is the MLE.

(c) Under suitable regularity conditions find a large sample approximation to a $100(1-\alpha)$ % HPD credible interval for θ.

(d) Use a stronger version of the result on asymptotic normality of posterior (to be stated by you) to prove that $\sqrt{n}(\widetilde{\theta}_n - \widehat{\theta}_n) \to 0$ with probability one, where $\widetilde{\theta}_n$ and $\widehat{\theta}_n$ denote respectively the posterior mean and MLE.

[5+13+5+6=29]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: 2018-19 (First Semester)

M. STAT. II YEAR
Commutative Algebra

Date : 05.09.2018 Maximum Marks : 30 Duration : $2\frac{1}{2}$ Hours

ANSWER ANY FOUR QUESTIONS.

R denotes a commutative ring, $\mathbb{Q}$ the field of rational numbers and $\mathbb{R}$ the field of real numbers.

1. (i) Let I, P_1, P_2 be ideals of R and P_3 a prime ideal of R such that $I \subseteq P_1 \cup P_2 \cup P_3$. Prove that $I \subseteq P_j$ for some j, $1 \leq j \leq 3$.

   (ii) Give an example to show that the result does not hold without the hypothesis that P_3 is prime. [6+3=9]

2. (i) Let $\phi : M \to F$ be a surjective R-linear map from a finitely generated R-module M onto a free R-module F. Show that the kernel of ϕ is a finitely generated R-module.

   (ii) Let k be a field and $R = k \oplus R_1 \oplus \cdots \oplus R_n \oplus \cdots$ be a graded ring. Prove that if R is Noetherian, then R is a finitely generated algebra over k. [4+5=9]

3. Let $f(X) = X^4 - 1 \in \mathbb{Q}[X]$ and $A = \mathbb{Q}[X]/(f(X))$.

   (i) Prove that there exists a ring-isomorphism from A to the direct product of fields $B = \mathbb{Q} \times \mathbb{Q} \times \mathbb{Q}(i)$.

   (ii) Identify a 3-tuple (a_1, a_2, a_3) in the product ring B which corresponds to the element $\bar{X}$ of A under the above isomorphism.

   (iii) Which element in A corresponds to the element $(0, 0, 1)$ under the above isomorphism? [5+2+2=9]

4. Let $C = \mathbb{R}[X, Y]/(X^2 + Y^2 - 1)$ and M the ideal $(x, y - 1)$ where x and y denote respectively the images of X and Y in C. Prove that:
   (i) C is a Noetherian domain.
   (ii) M is a maximal ideal of C.
   (iii) For every prime ideal P of C, M_P is a free C_P-module of rank one. [3+2+4=9]

5. Find an element f in the ring $D = \mathbb{C}[X, Y, Z]/(X^2 + Y^2 + Z^2)$ such that $D/fD \cong \mathbb{C}[T, 1/T]$. (Prove all the isomorphism steps.) [9]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: Semester I (2018-19)

M. Stat. II Year

PATTERN RECOGNITION

Date: September 6, 2018 | Maximum Marks: 50 | Duration: 2 hr

Note: *Answer as many questions as you can. The maximum you can score is 50.*

1. Consider the problem of discrimination between two populations ω_1, ω_2 based on observations on a random variable X. It is assumed that in ω_i, i=1,2, X is distributed as

$$
\begin{aligned}
p(x|\omega_1) &= \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{x-a}{2}\right)^2}, \\
p(x|\omega_2) &= \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{x-b}{2}\right)^2},
\end{aligned}
$$

where $0 \leq a < b$. Assume that π_1=$\pi_2 = \frac{1}{2}$, where π_i denotes the prior probability for the population $\omega_i,\ i = 1,2$.

a) Show that the Bayes rule for discriminating between ω_1 and ω_2 is linear.
b) Under zero-one loss, show that the overall error probability for this rule is

$$\frac{1}{2} - \frac{1}{\pi}\tan^{-1}\frac{b-a}{4}.$$

c) Write down the modified Bayes discriminant rule if, instead of zero-one loss, the following loss function is used:

$$
l(i,j) = \begin{cases} 0 & \text{if } i = j, \\ 1 & \text{if } i = 1, j = 2, \\ 2 & \text{if } i = 2, j = 1. \end{cases}
$$

Here $l(i,j)$ denotes the loss incurred when an observation from ω_i is allocated to $\omega_j, i,j = 1,2$. How does the overall error probability change under this loss? Explain.

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

(PLEASE TURN OVER)

2. Consider a two-category discrimination problem based on two discrete variables X and Y where X is a binary variable which takes values 0 or 1, while Y takes one of four ordinal values, $a < b < c < d$. Suppose the following observations (in the form of pairs xy) are available from the two categories, where the 1st and 2nd elements in each pair represents the observation on X and Y respectively:

Category 1	0a	0d	1a	1b	1d
Category 2	0b	1c	0c	0d	

a) Use the misclassification impurity measure to create an unpruned binary classification tree from this data choosing, at each node, variables and split values optimally so as to maximize the decrease in impurity as a result of the split. Explain each step clearly, preferably through tables associated with each node which contain the relevant quantities.

b) Explain in brief the rationale behind cost-complexity pruning of classification trees and write down the main steps of the algorithm that finds the optimal pruned tree in this sense.

[12+(4+4)=20]

3.

a) Formulate, as an appropriate optimization problem, the problem of training a Support Vector Machine (SVM) for discriminating between two classes, using n observations on a d-dimensional feature vector $\boldsymbol{X}$.

b) In the context of SVMs, explain, with the help of an example, how the choice of a suitable kernel function can take care of the issue of transformation of the feature vector to a higher-dimensional space without having to define the transformation explicitly.

[8+7=15]

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination (2018–2019)

M STAT II

Functional Analysis

Date : 07.09.2018 Maximum Marks : 60 Time : 2 hrs.

This paper carries 65 marks. Maximum you can score is 60. Precisely justify all your steps. Carefully state all the results you are using.

1. Let X, Y be normed linear spaces and $T : X \to Y$ be a linear map (not necessarily continuous). If T is an open map, then T is onto. [10]

2. Let X and Y be normed linear spaces and $T : X \to Y$ be a linear map. Define

$$\|x\|_1 = \|x\| + \|T(x)\|, \quad x \in X.$$

   (a) Show that $\|\cdot\|_1$ is a norm on X. [5]

   (b) Show that $\|\cdot\|_1$ is equivalent to $\|\cdot\|$ if and only if T is continuous. [5]

3. Let X be a normed linear space. Consider the following statements :

   (a) X is finite dimensional

   (b) Any two norms on X are equivalent

   (c) For any normed linear space Y and any linear map $T : X \to Y$, T is continuous

   (d) Any linear map $f : X \to \mathbb{K}$ is continuous

   Assuming $(a) \Rightarrow (b)$ (done in class), show that all the statements are equivalent. [20]

4. A subset $A \subseteq X^*$ is said to separate points of X if whenever $x_1 \neq x_2 \in X$, then there exists $f \in A$ such that $f(x_1) \neq f(x_2)$.

   Let X, Y be Banach spaces. Let $A \subseteq Y^*$ separate points of Y. Let $T : X \to Y$ be a linear map. Show that T is continuous if and only if $g \circ T$ is continuous for all $g \in A$. [10]

5. Let $\{e_n\}$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. Suppose $\{f_n\} \subseteq \mathcal{H}$ is an orthonormal set such that

$$\sum_{n=1}^{\infty} \|e_n - f_n\|^2 < 1.$$

   Show that $\{f_n\}$ is also an orthonormal basis for $\mathcal{H}$. [15]

INDIAN STATISTICAL INSTITUTE

MID-SEMESTRAL EXAMINATION 2018-2019

M.STAT 2nd year. Advanced Design of Experiments

September 7, 2018, Total marks 30 Duration: Two hours

Answer Question 1 or 2 and all questions 3-5.

Keep your answers brief and to the point.

1. Suppose you are to weigh 4 objects using a weighing balance with two pans; the balance needing zero bias correction. You are allowed to make 4 observations and suppose all observations are independent with constant variance σ^2.

   Derive a lower bound to the variance of the best linear unbiased estimators of these weights. [5]

2. When is a block design said to be balanced? Given a Balanced Block Incomplete block (BIB) design with parameters $v = b, r = k, \lambda$, one block from the BIB design is deleted and all treatments which appear in this block are deleted from the other blocks. Show that the remaining blocks will give a BIB design. Find it pàrameters. [5]

3. a) Define mutually orthogonal Latin squares.

   b) Let p be a prime number. Consider the squares A_j, $j = 1, \ldots, p-1$, constructed as follows:

   $$\begin{array}{ccccc} 0 & 1 & 2 & \ldots & p-1 \\ j & 1+j & 2+j & \ldots & p-1+j \\ 2j & 1+2j & 2+2j & \ldots & p-1+2j \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (p-1)j & 1+(p-1)j & 2+(p-1)j & \ldots & (p-1)+(p-1)j \end{array}$$

   where all entries in A_j are reduced mod p. Prove that the above squares form a set of mutually orthogonal Latin squares.

   (c) Construct two mutually orthogonal Latin squares of order 8.

   (You need to show the first square and only three rows of the 2nd square.) [2+4+4=10]

4. a) Define a Hadamard matrix.

   b) Can there exist a Hadamard matrix of order 15? Justify your answer with a proof. [2+4=6]

5. a) Describe an experimental situation where you have to compare 4 treatments and you would use a row-column design. Justify your answer.

   b) Give the model for analysing data from an experiment conducted using your design in (a) above.

   c) Write down the information matrix for (a) above under the model in (b). [3+3+3=9]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2018 – 19

M. Stat (2nd Year)

Time Series Analysis

Date: 11 September 2018 Maximum Marks: 30 Duration: 2 Hours

1. Consider the random walk with drift model

$$x_t = \delta + x_{t-1} + w_t$$

for t = 1,2,..., with $x_0 = 0$, where w_t is white noise with variance σ_w^2.

(a) Show that the model can be written as $x_t = \delta t + \sum_{k=1}^{t} w_k$.

(b) Find the mean function and the auto-covariance function of x_t.

(c) Argue that x_t is not stationary.

(d) Show that the ACF $\rho_x(t-1, t) = \sqrt{\frac{t-1}{t}} \to 1 \text{ as } t \to \infty$. What is the implication of this result?

(e) Suggest a transformation to make the series stationary, and prove that the transformed series is stationary. [1 + (2 + 3) + 2 + 3 + (2 + 2) = 15]

2. Find the conditions on the parameters of the AR(2) model,

$$(1 - \varphi_1 B - \varphi_2 B^2)x_t = w_t,$$

for causality. Show this causality region in the (φ_1, φ_2) – space diagrammatically.

[4 + 3 = 7]

3. Find the ACF of an ARMA (1, 1) process

$$x_t = \varphi x_{t-1} + \theta w_{t-1} + w_t,$$

where $|\varphi| < 1$. [8]

INDIAN STATISTICAL INSTITUTE

(203 B. T. Road, Kolkata 700 108)

Master of Statistics (M.Stat.) IInd Year
Academic Year 2018 - 2019: Semester I

Martingale Theory
Midterm Examination
Instructor: Antar Bandyopadhyay

Date: September 10, 2018 **Total Points: 25**
Time: 02:30 PM - 04:30 PM **Duration: 120 minutes**

Note:

- **Please write your roll number on top of your answer paper.**
- **There are five problems carrying 6 points each. Solve as many as you can. Maximum you can score is 25.**
- **Show all your works and write explanations when needed.**
- **This is an open note examination. You are allowed to use your own hand written notes (such as class notes, your homework solutions, list of theorems, formulas etc). No printed materials or photo copies are allowed, in particular you are not allowed to use books, photocopied class notes etc.**

1. Suppose μ is a *signed measure* on a measurable space $(\Omega, \mathcal{F})$. Then show that there exist a probability ν on $(\Omega, \mathcal{F})$ and a random variable X defined on it, such that, $\mu(A) = \mathbf{E}[X \, 1_A]$ for all $A \in \mathcal{F}$,where the expected value is computed with respect to the probability ν.

2. Show that *Lebesgue σ-algebra* on $(0,1)$ is not *countably generated.*

3. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $(X_i)_{i \geq 1}$ be *i.i.d.* random variables defined on it. Assume $\mathbf{E}[|X_1|] < \infty$. Let $S_n := X_1 + X_2 + \cdots + X_n$, $n \geq 1$. Find $\mathbf{E}\left[X_1 \,\middle|\, S_n\right]$.

4. Suppose X and Y are two independent random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a bounded Borel measurable function. Let $h : \mathbb{R} \to \mathbb{R}$ be a function defined as $h(x) := \mathbf{E}[\phi(x, Y)]$, $x \in \mathbb{R}$. Then show that $\mathbf{E}\left[\phi(X,Y) \,\middle|\, X\right] = h(X)$ a.s..

5. Let $(X_n)_{n \geq 1}$ be *i.i.d.* ± 1-variables with $p := \mathbf{P}(X_1 = +1) < \mathbf{P}(X_1 = -1)$. Put $S_0 = 0$ and $S_n := X_1 + X_2 + \cdots + X_n$, $n \geq 1$. Show that

   (a) $Y_n := \left(\frac{1-p}{p}\right)^{S_n}$, $n \geq 0$ is a *martingale* with respect to the filtration $(\sigma(X_1, X_2, \ldots, X_n))_{n \geq 0}$; and

   (b) $\mathbf{E}\left[\sup_{n \geq 0} S_n\right] \leq \frac{p}{1-2p}$.

Good Luck

INDIAN STATISTICAL INSTITUTE

(203 B. T. Road, Kolkata 700 108)

Master of Statistics (M.Stat.) IInd Year
Academic Year 2018 - 2019: Semester I

Martingale Theory
Final Examination
Instructor: Antar Bandyopadhyay

Date: November 09, 2018 Total Points: 50
Time: 10:30 AM - 02:30 PM Duration: 4 hours

Note:

- **Please write your roll number on top of your answer paper.**
- **There are four problems carrying 15 points each. Solve as many as you can. Maximum you can score is 50.**
- **Show all your works and write explanations when needed.**
- **This is an open note examination. You are allowed to use your own hand written notes (such as class notes, your homework solutions, list of theorems, formulas etc). No printed materials or photo copies are allowed, in particular you are not allowed to use books, photocopied class notes etc.**

1. State whether the following statements are ***true*** or ***false***. Write brief reasons supporting your answers. For each **correct guess you will get +1 point** but for each **wrong guess you get −2 point**. If your guess is **correct and your reasoning is also correct then you will get an additional +4 points**. However, if you give a **wrong reasoning then you will receive additional −3 points**. [$(1+4)\times 3 = 15$]

   (a) Suppose $(M_n, \mathcal{F}_n)_{n\geq 0}$ is a *bounded increment* martingale with $M_0 = 0$. Then there exists a sequence of positive real numbers, say $(\sigma_n)_{n\geq 1}$, such that,
   $$\frac{M_n}{\sigma_n} \xrightarrow{d} \text{Normal}\,(0,1)\,.$$

   (b) Suppose $(X_n)_{n\geq 1}$ is a sequence of random variables such that $(X_i, X_j) \stackrel{d}{=} (X_j, X_i)$, for all $1 \leq i \neq j < \infty$. Then $(X_n)_{n\geq 1}$ is *exchangeable.*

   (c) Let X be a random variable such that $|X| \sim$ Exponential (1). Then the distribution of X and the *Lebesgue measure* on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ are *equivalent.*

[please turn over]

2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X_n)_{n\geq 1}$ be i.i.d. Bernoulli $\left(\frac{1}{2}\right)$ random variables defined on it. Let $\mathbf{Q}$ be the law of the sequence $(X_n)_{n\geq 1}$ on $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$. Let $Y_{k,n} := \mathbf{1}(X_n = 1, X_{n+1} = 1, \cdots, X_{n+k} = 1)$ for $k \geq 1$ and $n \geq 1$. For each $k \geq 1$, denote by $\mathbf{P}_k$, the law of the process $(Y_{k,n})_{n\geq 1}$ defined on $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$. Define $\mathbf{P} := \sum_{k=1}^{\infty} \frac{1}{2^k} \mathbf{P}_k$.

   (a) Show that for any $k \geq 1$

   $$\frac{1}{n} \sum_{j=1}^{n} Y_{k,j} \longrightarrow \frac{1}{2^{k+1}} \quad \text{a.s.}$$

   (b) Show that $\mathbf{P}_k$ and $\mathbf{P}_{k'}$ are *mutually singular* for any $1 \leq k \neq k' < \infty$.

   (c) Find the *Lebesgue decomposition* of $\mathbf{P}$ with respect to $\mathbf{Q}$.

   $[8 + 5 + 2 = 15]$

3. Suppose $(X_n)_{n\geq 1}$ is a sequence of random variables, such that for any $k \geq 1$ and $1 \leq i_1 < i_2 < \cdots < i_k < \infty$, the joint distribution of $(X_{i_1}, X_{i_2}, \cdots, X_{i_k})$ is *multivariate normal*. Then show that $(X_n)_{n\geq 1}$ is *exchangeable*, if and only if,

   $$X_n = \theta + Z_n, \ \forall n \geq 1,$$

   where $\theta \sim \text{Normal}(\mu, \sigma^2)$, for some $\mu \in \mathbb{R}$ and $\sigma^2 \geq 0$; $(Z_n)_{n\geq 1}$ are i.i.d. $\text{Normal}(0, \eta^2)$ for some $\eta^2 \geq 0$; and θ is independent of $(Z_n)_{n\geq 1}$. [15]

   Note: Here $Z \sim \text{Normal}(a, 0)$ means Z is the degenerate random variable $Z \equiv a$.

4. Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and all random variables are defined on it.

   (a) Let $(X_n, \mathcal{F}_n)_{n\geq 0}$ be a non-negative super-martingale. Then show that there exists $N \subseteq \Omega$, such that, $\mathbf{P}(N) = 0$ and for all $\omega \notin N$, if $X_k(\omega) = 0$ then $X_{n+k}(\omega) = 0$, for all $n \geq 0$. [7]

   (b) Suppose $(X_n, \mathcal{F}_n)_{n\geq 0}$ is a super-martingale and and $(Y_n, \mathcal{F}_n)_{n\geq 0}$ is a sub-martingale, such that, $Y_n \leq X_n$ a.s. for all $n \geq 0$. Then show that there exists a martingale $(M_n, \mathcal{F}_n)_{n\geq 0}$, such that, $Y_n \leq M_n \leq X_n$ a.s. for all $n \geq 0$. [8]

   Hint: Consider $Z_{n,k} := \mathbf{E}\left[X_{n+k} \,\middle|\, \mathcal{F}_n\right]$, for $k \geq 1$.

Good Luck

INDIAN STATISTICAL INSTITUTE

First Semester Examination : 2018-19

Course Name: M.Stat. 2nd Year

Subject Name: Advanced Design of Experiments

Date: 9th November, 2018 Total Marks: 70.

Duration: 3 hours Note: Answer **question 1 and any four** Questions.

Answers should be to the point. Marks may be deducted for unduly lengthy answers.

1. (a) Define E-optimality and explain its statistical significance.

   (b) Prove that a regular generalized Youden Square design is universally optimal in a suitable class of row-column designs. Where do you need the 'regular' property in the proof?

   [2+4=6]

2. a) Construct a Hadamard matrix of order 12, clearly stating the result you use to construct this (proof of result not required).

   b) Prove that the existence of a Hadamard matrix of order $4t$ implies the existence of a symmetric BIB design (the parameters of the BIB design are to be determined by you).

   c) Prove that a Youden square can always be constructed from a symmetric BIB design.

   [(4+6+6=16]

3. (a) A factorial experiment is to be conducted with 3 factors F_1, F_2 and F_3 at levels 2, 3 and 4, respectively. Using suitable notation, **write down**(without proof) the expression for a full set of orthonormal treatment contrasts belonging to the interaction F_1F_3.

   (b)Prove that the set of contrasts written by you in (a) is indeed a set of 'contrasts'.

   (c)Prove that the set of contrasts written by you in (a) is indeed a 'full set' of contrasts.

   (d) Prove that contrasts belonging to any two distinct interactions are always mutually orthogonal. [4+4+4+4=16]

4. (a) Show that if a factorial design is balanced for a factorial effect, then the BLUE's of every two mutually orthogonal contrasts belonging to this effect are uncorrelated.

   (b) Define a fractional factorial plan. Obtain a necessary condition for the estimability of contrasts belonging to the factorial effects if the experiment is to be conducted in N runs (clearly state the results you are using).

   (c) What is meant by a resolution (2,3) plan for a factorial experiment?

   (d) Obtain a 9-run main effect plan for a 3^4 experiment.

   [4+4+4+4=16]

5. (a) Define an orthogonal array $OA(N, k, s, t)$ with s symbols and strength t,

   (b) For each of the following statements **write** whether it is True or False.

   (i) An $OA(N, k, s, t)$ is also an $OA(N, k, s, t')$, where $0 < t' < t$.

P.T.O

(ii) An $OA(s^2, k, s, 2)$ always exists if $k \leq s - 1$.

(iii) Not all $N \times k'$ subarrays of an $OA(N, k, s, t)$ are orthogonal arrays of strength t.

(iv) Suppose $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ is an $OA(N, k, s, t)$. If A_1 is itself an $OA(N_1, k, s, t_1)$, then A_2 is an $OA(N - N_1, k, s, t_2)$ with $t_2 \geq t_1$.

(c) Prove that for $N \geq 4$, the existence of a Hadamard matrix of order N implies the existence of an $OA(N, N-1, 2, 2)$.

(d) From the OA of (c) above, how can you get an orthogonal array of strength 3?

[2+(2× 4) + 3+3=16]

6. (a) Describe an experimental situation where you would use cross-over designs for experimentation. State the model and associated assumptions for analysing data from these designs.

(b) Construct a balanced uniform crossover design with 6 treatments and the minimum possible number of periods and subjects needed.

(c) Give the conditions on the parameters under which a strongly balanced uniform crossover design may exist.

(d) Prove that for a strongly balanced uniform crossover design, the information matrix for direct effects is completely symmetric.

[4+4+4+4=16]

M.Stat II / Quantitative Finance

Midsem. Exam. / Semester I 2018-19

Time - 2 and 1/2 hours/ Maximum Score - 30

Date: September 10, 2018

1. Let $X_n = X_0 + R_1 + \cdots + R_n$ and $Y_n = Y_0 + S_1 + \cdots + S_n$, for all $n \geq 1$, for all $n \geq 1$, where (R_k, S_k) are i.i.d. random variables with Bivariate Normal distribution with mean $(0, 0)$ and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with $-1 < \rho < 1$.

   (a) (4 marks) Assuming $X_0 \equiv 0 \equiv Y_0$, give brief reason that X_n and Y_n are martingale with respect to appropriate filtering $\{\mathcal{F}_n\}$ after describing this $\{\mathcal{F}_n\}$.

   (b) (6 marks) Find the compensator for $\{X_n^4\}$ and for $\{X_n Y_n\}$.

2. Consider the portfolio optimization problem with N risky assets, as given by the data.

   (a) (4 marks) Based on 3/4 of (time series) data, find the (i) minimum variance portfolio vector; (ii) minimum variance portfolio vector with shorsale constraint; (iii) minimum variance portfolio vector which gives 8% return; (iv) minimum variance portfolio vector with shorsale constraint which gives 8% return, .

   (b) (5 marks) Assume there is also a risk-free asset in the market of N risky asset that pays $r = 6\%$ **per annum**.

   Describe the risk-free asset and the corresponding tangent portfolio convex combination of which generates an expected return of 8% and with minimum possible variance. Find the variance.

   (c) (6 marks) Use remaining 1/4 of data to find the out-of-sample variance of all the returns corresponding to portfolio vectors described above and compare it with the corresponding in-sample variances.

3. Let the market has 5 risky asset and one risk-free asset. where the intial price vector of the asset is given by $(1, 0.8982, 1.1914, 1.1634, 1.2955, 1.1275)'$.

   At the next time period possible change of prices are given as follows.

P.T.O.

$$\begin{pmatrix} 1.05 & 1.05 & 1.05 & 1.05 & 1.05 & 1.05 \\ 0.8359 & 0.8649 & 0.8367 & 0.8051 & 0.8328 & 0.8795 \\ 1.308 & 1.296 & 1.329 & 1.229 & 1.276 & 1.203 \\ 1.0957 & 1.1395 & 1.1147 & 1.1489 & 1.1172 & 1.1590 \\ 1.2950 & 1.3401 & 1.3161 & 1.3722 & 1.2588 & 1.3032 \\ 1.1405 & 1.1795 & 1.1522 & 1.1970 & 1.1613 & 1.1298 \end{pmatrix}$$

(a) (5 marks) Is the market arbitrage free? Justify your answer by showing the existence of probability measure w.r.t. which the assets are (discounted) markingale.

(b) (3 marks) Is the market complete? Justify your answer.

INDIAN STATISTICAL INSTITUTE

End-Semester Examination: Semester I (2018-19)

M. STAT. II Year

Course: PATTERN RECOGNITION

Date: November 12, 2018 Maximum Marks: 100 Duration: 3 hr

Note: *Answer as many questions as you can. The maximum you can score is 100. Use of calculators is permitted.*

1.

(a) For solving a c-class pattern recognition problem based on a training set of size n, consisting of $n_i(> 0)$ observations from class ω_i, $i = 1,2,\cdots,c\ (\geq 2)$, with $\sum_{i=1}^{c} n_i = n$, describe the k-Nearest Neighbour (k-NN) rule, where k is a pre-specified positive integer such that $k \ll n$.

(b) Show that, as $n \to \infty$, the limiting value of the overall error probability of the 1-NN rule is bounded above by

$$P^*\left(2 - \frac{c}{c-1}P^*\right),$$

where P^* is the Bayes error for the problem.

(c) In the context of a 2-class problem, describe the *editing* procedure for reducing the training set size, starting with a k-NN rule.

(d) Show that the k-NN rule with editing has an error rate approximately equal to that of the Bayes rule under certain conditions.

(e) Consider a 2-class problem based on a d-dimensional feature vector x whose components $x_j, j = 1,2,\cdots,d$, are independently and identically distributed in class $\omega_i, i = 1,2$, as

$$p_{ij}(x) = \begin{cases} \theta_i^x(1-\theta_i)^{1-x}, & \text{for } x = 0,1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta_i \in (0,1)$, $i = 1,2$.

Suppose d is an odd integer greater than 1. Further, let $\theta_1 = p\ (> 0.5)$, $\theta_2 = 1 - p$, and $\pi_1 = \pi_2 = 0.5$. Derive the Bayes classification rule for the problem and determine the limiting value of the Bayes error as $p \to 0.5$.

(**Hint**: You may use the fact that the cumulative distribution function $P(x) = P[X \leq x]$ of a binomial random variable with parameters m and p can be written as the Incomplete Beta function $I_{1-p}(m - x, x + 1)$.)

[4+6+4+6+(4+6)=30]

(PLEASE TURN OVER)

2.

(a) Formulate the problem of clustering n d-variate observations in an unlabeled training set $\mathcal{D}$, where $d \geq 1$, as a problem of estimating the (unknown) parameters of a finite mixture of probability densities, making appropriate assumptions.

(b) Presuming the underlying mixture model to be identifiable, deduce an iterative algorithm for computing the maximum likelihood estimates of the unknown parameters, including the mixing proportions π_i, for the special case where the mixture density has d-variate Gaussian components with (unknown) mean vectors $\boldsymbol{\mu}_{i,} i = 1,2, \cdots, k$, and a common dispersion matrix $\sigma^2 \boldsymbol{I}$, σ being known.

(c) Write down the main steps of the k-means algorithm for grouping the n d-variate observations in the unlabeled training set $\mathcal{D}$ into k clusters.

(d) Show that k-means algorithm in (c) may be looked upon as a limiting form of the iterative algorithm in part (b) of this problem. Clearly state any assumption(s) you may need to make for this purpose.

[2+5+5+8=20]

3.

(a)

i. Briefly discuss the principle of *Boosting* in the context of classification problems.

ii. For a two-class problem, write down the main steps of the popular *AdaBoost* algorithm used for boosting the performance of a weak classifier.

(b) Describe the *Random Forests* algorithm for classification and explain how it improves upon over the Bootstrap Aggregating (*Bagging*) of tree classifiers.

[(3+7)+(8+2)=20]

4.

(a) Describe Rousseeuw's *silhouette index* and explain how it can be used for cluster validation as well as for determination of the optimal number of clusters.

(b) What is a Receiver Operating Characteristic (ROC) curve? Explain its utility in the context of assessment of the performance of supervised classification methods.

[(4+3+3)+(4+6)=20]

5.

(a) What are *Radial Basis Function* (RBF) neural networks? Discuss how they can be used for supervised classification, providing details of any method known to you for training such networks.

(b) Show that an RBF network having two hidden nodes with respective activation functions $\varphi_j(\boldsymbol{x}) = e^{-\|x-c_j\|^2}, j = 1, 2$, where $\boldsymbol{c}_1 = (0,0)', \boldsymbol{c}_2 = (1,1)'$, can solve the 2-class exclusive-OR (XOR) problem, for which the training data is given in the following table:

i	x_{1i}	x_{2i}	Class label
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0

Assuming all input-to-hidden weights to be equal to 1, estimate the hidden-to-output weights.

[(4+3+6)+7=20]

INDIAN STATISTICAL INSTITUTE

Semestral Examination (2018–2019)

M. STAT. II (MSP)

Functional Analysis

Date : 13.11.2018 Maximum Marks : 100 Time : $3\frac{1}{2}$ hrs.

This paper carries 125 marks. Maximum you can score is 100. Precisely justify all your steps. Carefully state all the results you are using.

1. (a) Let X be a normed linear space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and $f : X \to \mathbb{K}$ be a linear map. Show that $\ker f = \{x \in X : f(x) = 0\}$ is either closed or dense in X. [5]

   (b) Let X be a Banach space and $f : X \to \mathbb{K}$ be a linear map. Show that f is continuous if and only if $\ker f$ is a G_δ-set in X. [15]

2. Let (X, d) be a metric space and $f : X \to \mathbb{R}$. Show that the following are equivalent :

   (a) f is lower semi-continuous

   (b) For $x_n, x_0 \in X$, $x_n \to x_0$ implies $\liminf f(x_n) \geq f(x_0)$

   (c) for any $x_0 \in X$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in B(x_0, \delta)$, $f(x) > f(x_0) - \varepsilon$. [12]

3. Let S be a linear subspace of $C[0,1]$. Show that S is closed in $L^2[0,1]$ if and only if S is finite dimensional, by proving the following :

   (a) Any finite dimensional subspace of any normed linear space is closed. [8]

   (b) Show that S is also closed in $C[0,1]$. [5]

   (c) Show that there is a constant $M > 0$ such that for all $f \in S$, [7]

   $$\|f\|_2 \leq \|f\|_\infty \leq M\|f\|_2$$

   (d) If $\{f_1, f_2, \ldots, f_n\}$ is a finite orthonormal set in $S \subseteq L^2[0,1]$, then [10]

   $$\sum_{k=1}^{n} |f_k(x)|^2 \leq M^2 \text{ for all } x \in [0,1].$$

   (e) Show that $\dim(S) \leq M^2$. [5]

   [Note : The constant M in (d) and (e) is same as that in (c).] **[PTO]**

4. Let X, Y be Banach spaces. Let $T \in \mathcal{L}(X, Y)$. Show that there is a constant $c > 0$ such that $\|Tx\| \geq c\|x\|$ for all $x \in X$ if and only if T is 1-1 and $T(X)$ is closed in Y. [20]

5. Let $(\alpha_n) \in \ell^\infty$. Define a linear map $T : \ell_2 \to \ell_2$ by

$$T((x_n)) = (\alpha_n x_n), \qquad (x_n) \in \ell_2.$$

Show that

(a) T is continuous, normal and $\|T\| = \|(\alpha_n)\|_\infty$. [10]

(b) T is invertible if and only if $\{\alpha_n\}$ is bounded away from 0. [8]

(c) λ is an eigenvalue of T if and only if $\lambda = \alpha_n$ for some $n \geq 1$. [4]

(d) $\sigma(T) = \overline{\{\alpha_n : n \geq 1\}}$. [4]

(e) T is compact if and only if $(\alpha_n) \in c_0$. [12]

Date: 16.11.2018

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, First Semester, 2018-19

Semestral Examination

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

[Answers should be brief and to the point. Answer as many as you can. The maximum score is 100.]

1. (a) Describe how you will generate an observation from the bivariate distribution with probability density function [4]

$$f(x,y) = \frac{e^{-(x+y)}}{\pi\sqrt{xy}}, \quad 0 < x, y < \infty.$$

(b) Consider the discrete univariate distribution with probability mass function

$$f(x) = \begin{cases} \frac{1}{4x(x-1)} & \text{if } x = -1, -2, \ldots \\ \frac{1}{2} & \text{if } x = 0 \\ \frac{1}{4x(x+1)} & \text{if } x = 1, 2, \ldots. \end{cases}$$

Using only one random number from the $U(0,1)$ distribution, how will you generate an observation from this distribution? [8]

2. Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with distribution function F satisfying $\int x^2 \, dF(x) < \infty$. Define F_n to be the empirical analog of F.

(a) Let $T_n = T(F_n)$ be an estimator of $\theta = T(F)$ with $Bias(T_n) = O(n^{-\delta})$ for some $\delta \in (0,1)$. Find the order of bias for the corresponding jackknife estimator. [6]

(b) If $T(F) = \int x \, dF(x)$, find the jackknife estimator and the bootstrap estimator of $Var(T_n)$. Hence comment on their unbiasedness. [6]

3. (a) Find the half-space median, the simplicial median and the spatial median of the univariate distribution with probability density function [2+2+2]

$$f(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

(b) Consider the following observations generated from a distribution with location parameter θ.

$$x_1 = 12.9 \quad x_2 = 10.1 \quad x_3 = 11.2 \quad x_4 = 9.8 \quad x_5 = 13.5$$
$$x_6 = 10.6 \quad x_7 = 11.5 \quad x_8 = 8.9 \quad x_9 = 12.2 \quad x_{10} = 11.8$$

(i) Find the least median of squares estimate for θ. [3]

(ii) Find the set of all values of α that minimize $f(\alpha) = 3\sum_{i=1}^{10} |x_i - \alpha| + \sum_{i=1}^{10}(x_i - \alpha)$. [3]

4. (a) Show that the iteratively re-weighted least squares algorithm used for LAD regression can be viewed as a majorization-minimization (MM) algorithm. Does this algorithm always converge? Justify your answer. [5+3]

(b) Check whether the following bivariate functions belong to the class of (i) additive models, (ii) projection pursuit models. [4]

$$(i)\ f_1(x_1, x_2) = log(x_1 x_2) \quad (ii)\ f_2(x_1, x_2) = (1 + x_1)(1 + x_2)$$
$$(iii)\ f_3(x_1, x_2) = e^{x_1 + x_2} \quad (iv)\ f_4(x_1, x_2) = \sin(x_1 + x_2)\sin(x_1 - x_2)$$

5. Consider the following data set.

x	-5	-4	-3	-2	-1	0	1	2	3	4	5
y	25	16	9	4	1	0	1	4	9	16	25

(a) Find the equation of the line $y = \alpha + \beta x$ that passes through the point $(0, 4)$ and minimizes $\sum_{i=1}^{10} |y_i - \alpha - \beta x_i|$. (Here $(x_i, y_i)'$ denotes the i-th observation $(i = 1, 2, \ldots, 11)$). [6]

(b) Find the regression depth of the linear fit obtained in (a). Is this a linear fit with the maximum regression depth? Justify your answer. [3+3]

6. (a) Consider the following data set with two missing values

x	-5	-4	-3	-2	-1	1	2	3	4	5
y	10.7	9.5	10.3	8.9	*	*	9.7	8.5	8.1	7.3

Let $\hat{f}_h$ be the Nadaraya-Watson estimate of the regression function obtained from this data set when a Gaussian kernel with bandwidth h is used. If it is known that $\lim_{h \downarrow 0} \hat{f}_h(0.5) = \lim_{h \uparrow \infty} \hat{f}_h(0.5) = 9.0$, find the limiting value of the local linear estimate of the regression function at 0.5 as the associated bandwidth tends to infinity. [6]

(b) Consider a two-class problem. In order to construct a classification tree, one needs to find the best split based on a categorical measurement variable C, which has three categories C_1, C_2 and C_3. The number of observations in these three categories are given below

Category	C_1	C_2	C_3	Total
No. of. observations from class-1	30	40	30	100
No. of. observations from class-2	70	0	30	100
Total	100	40	60	200

Assuming the impurity function to be concave and symmetric in its arguments, find the best split based on C. [6]

7. (a) Describe the Metropolis-Hastings algorithm for generating observations from a target distribution π. Show that if $X_t \sim \pi$, all subsequent observations generated by this algorithm will follow the same distribution. [3+3]

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

$$f(y, z) = C\left(\frac{y}{1-y}\right)\frac{\{10(1-y)\}^z}{(z-1)!}, \quad 0 \le y \le 1; \quad z = 1, 2, \ldots,$$

where C is a normalizing constant. [6]

8. Computer assignments [20]

MSTAT II - Quantitative Finance

Final Exam. / Semester I 2018-19

Date - November 19, 2018

Time - 3 hours/ Maximum Score - 50

<u>NOTE</u> : **THE PAPER HAS QUESTIONS WORTH 57 MARKS. SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED. STANDARD CALCULATORS AND TABLES, AS PER RULE OF ISI, MAY BE USED.**

1. Let $X_n = X_0 + R_1 + \cdots + R_n$ and $Y_n = Y_0 + S_1 + \cdots + S_n$, for all $n \geq 1$, where (R_k, S_k) are i.i.d. random variables with $P(R_k = 1, S_k = 1) = p = P(R_k = -1, S_k = -1)$ and $P(R_k = -1, S_k = 1) = q = P(R_k = 1, S_k = -1)$, where $p > q > 0$, and $p + q = 1/2$.

   (a) (2 marks) Assuming $X_0 \equiv 7 \equiv Y_0$, note that X_n and Y_n are martingale with respect to appropriate filtration $\{\mathcal{F}_n\}$. Describe this $\{\mathcal{F}_n\}$.

   (b) (8 marks) Find the compensators for $\{X_n Y_n\}$. Is $M_n = (q/p)^{Z_n}$ a martingale with respect to the above filtration, where $Z_n = |X_0 + Y_0| + \sum_{k=1}^{n} |R_k + S_k| - n$? Justify.

   (c) (5 marks) If the price of an asset on the nth day is M_n Rupees, then find the expected time that $\log(M_n)/\log(q/p)$, the price, will hit either 0 or 30 Rupees.

2. Let $\{S_t\}$ be a price of a share at time t, given by

$$dS_t = \mu dt + \sigma dB_t,$$

   where $\{B_t : t \geq 0\}$ is a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, and μ is the real world growth rate and the $\sigma > 0$ is the instantaneous volatility.

   (a) (5 marks) Show that $S_t = S_0 \exp\{(\mu - \sigma^2/2)t + \sigma B_t\}$.

   (b) (5 marks) Let $r > 0$ be the risk-free interest rate. Find the measure P^* with respect to which $\{S_t\}$ grows at the rate of risk-free rate r (instead of real world growth rate μ).

   (c) (5 marks) Find the distribution of $M_t = Max_{0 \leq u \leq t} S_u$.

3. (a) (5 marks) Write the Put-Call parity equation for currency options (European type). Justify the equation using no-arbitrage argument.

P.T.O

(b) (5 marks) Use Black-Scholes-Merton's framework for risk-neutral valuation to show that the European type Call (on currency) option price at time $t \geq 0$ and to mature at time $T \geq t$, is

$$C_t = e^{-r_f(T-t)} S_t \Phi(d_{1t}) - K e^{-r_d(T-t)} \Phi(d_{2t})$$

where S_t is the price at time t and K is the strike price, $d_{1t} = [\ln(S_t/K) + (r_d - r_f + \sigma^2/2)(T-t)]/[\sigma\sqrt{T-t}]$, $d_{2t} = d_{1t} - \sigma\sqrt{T-t}$, and r_d and r_f are domestic and foreign risk-free interest rates, respectively.

(c) (5 marks) Find the Black-Scholes-Merton's p.d.e. for above call options.

4. (4+4+4=12 marks) For any of the following, write TRUE or FALSE and justify your answer. Do any three.

(a) Suppose two portfolios have equal expected return. Then the portfolio with more volatility is preferable to the portfolio with less volatility, whether the expected return is positive or negative.

(b) With one risky asset and one risk-less asset, if a no-arbitrage market is incomplete then there is a unique martingale measure to determine the prices of any derivative (or claim).

(c) The explicit difference method is always more accurate than the implicit difference method and always similar to the trinomial method.

(d) VaR is a better measurement of risk than Expected Shortfall.

All the best.

INDIAN STATISTICAL INSTITUTE
Semestral Examination: 2018-19 (First Semester)

M. STAT. II YEAR
Commutative Algebra

Date : 19.11.2018 Maximum Marks : 70 Duration : 3 Hours

k denotes a field and R a commutative ring with 1.

GROUP A
ATTEMPT ANY FIVE QUESTIONS.

1. Let I be a finitely generated ideal of R satisfying $I^2 = I$. Prove that there exists $f \in I$ for which $f^2 = f$ and $I = (f)$. [9]
2. Let B be a subring of an integral domain A. Suppose that there exists a nonzero element $f \in B$ such that $B[1/f] = A[1/f]$ and $fA \cap B = fB$. Prove that $B = A$. [9]
3. Prove that R is a reduced ring (i.e., does not have nonzero nilpotents) if and only if R_P is a reduced ring for every prime ideal P of R. [9]
4. Prove that if M and N are Noetherian R-modules, then so is $M \otimes_R N$. [9]
5. Prove that any k-subalgebra of $k[X]$ is a finitely generated k-algebra. [9]
6. Prove that $A = \mathbb{C}[X, Y]/(Y^2 - X^2 - X^3)$ is a non-normal integral domain. [9]
7. Let V be an affine algebraic set in $\mathbb{C}^n$ and g an element of the coordinate ring $\mathbb{C}[V]$ such that $g(x) \neq 0$ for each $x \in V$, considering g as a polynomial function on V. Prove that g is a unit in $\mathbb{C}[V]$. [9]

GROUP B
ATTEMPT ANY TWO QUESTIONS.

1. Let $A = \mathbb{R}[X, Y, Z]/(X^2 + Y^2 + Z^2 - 1)$.
(i) Prove that A is a Noetherian integral domain.
(ii) Examine whether A is a UFD. (Clearly state the result(s) that you use.) [5+10=15]
2. (i) Compute $k[X]/(X^4) \otimes_{k[X]} k[X]/(X^6)$.
(ii) Let M and N be finitely generated R-modules such that $M \otimes_R N = 0$. Prove that $Ann_R M + Ann_R N = R$. [Hint: Use Local-Global Technique.] [5+10=15]
3. (i) Let A be an integral domain which is finitely generated as a k-algebra and L the field of fractions of A. Prove that if $A \neq L$, then $A[1/f] \neq L$ for any $f \in A$.
(ii) Give an example of a proper subring B of $\mathbb{Q}$ and an element $f \in B$ such that $B[1/f] = \mathbb{Q}$. [10+5=15]

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2018-2019

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 22 November, 2018 Max. Marks: 100 Duration: 3 Hours

Answer all questions

1. (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) Suppose we have observations $X_1, \ldots, X_n$. Under M_0, X_i are i.i.d. $N(0, 1)$ and under M_1, X_i are i.i.d. $N(\theta, 1)$, $\theta \in R$. Consider the noninformative prior $g_1(\theta) \equiv 1$ for θ under M_1. Find the intrinsic prior for θ for the AIBF and show that the ratio of the AIBF and the BF with this intrinsic prior tends to one as n tends to infinity.

(c) Let $X_1, \ldots, X_n$ be i.i.d. observations with a common density $f(x|\theta) = \exp(-(x-\theta))$, $x > \theta$. Our problem is to compare the models $M_0 : \theta = 0$ with $M_1 : \theta > 0$. Using the standard noninformative prior under M_1, calculate the fractional Bayes factor (FBF) and comment on its suitability as a model selection criterion.

[(3+7)+11+12=33]

2. Consider p independent random samples, each of size n, from p normal populations $N(\theta_j, \sigma^2)$, $j = 1, \ldots, p$. Assume σ^2 to be known. Also assume that $\theta_1, \ldots, \theta_p$ are i.i.d. $N(\eta_1, \eta_2)$. 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 is a suitable shrinkage estimate expected to perform better than this estimate?

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

[5+(15+7)=27]

P.T.O

3. Suppose we are studying the distribution of the number of defectives X in the daily production of a product. Consider the model $X|Y,\theta \sim \text{Bin}(Y,\theta)$ where Y, a day's production, follows Poisson (λ). The difficulty is that Y is not observable and inference has to be made on the basis of X only.

Consider a Beta(α,β) prior for θ, and show how Gibbs sampling can be used to sample from the posterior distribution of θ. Find all the required conditional distributions.

[17]

4. Write down the Laplace approximation for an integral in the multidimensional case. Describe how this approximation can be used to obtain the Bayesian information criterion (BIC) for model selection.

[8]

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 uniform prior over $[-K,K]$ for a very large K under M_1.

(a) Explain why this is not a reasonable specification of prior for this problem.

(b) Will there be a conflict between P-value and the posterior probability of M_0? Justify your answer.

[8+7=15]

INDIAN STATISTICAL INSTITUTE

Semester Examination: 2018-19 (First Semester)

M. Stat. II Year

Fourier Analysis

Date: 22/11/2018 Maximum Marks: 60 Duration: $2\frac{1}{2}$ Hours

---

Note: Give proper justification to your answers. State clearly all the results you are using.

---

(1) For $f \in L^1(\mathbb{T})$, let $S_N f$ be the N-th partial sum of the Fourier series of f. Define

$$S^*f(x) = \sup_{N\in\mathbb{N}} |S_N f(x)|,$$

the maximal operator associated with $\{S_N\}$. Let $1 < p < \infty$ and assume that the operator $S^* : L^p(\mathbb{T}) \to L^p(\mathbb{T})$ is bounded. Prove that

(a) $|\{x \in \mathbb{T} : |f(x)| \geq \lambda\}| \leq C\frac{\|f\|_p^p}{\lambda^p}$, for some $C > 0$.

(b) $S_N f$ converges to f almost everywhere for all $f \in L^p(\mathbb{T})$. [15]

(2) Suppose $f, \hat{f} \in L^1(\mathbb{R}^n)$ be such that

$$|f(x)| + |\hat{f}(x)| \leq C\frac{1}{(1+|x|)^{n+\delta}}$$

for some $C, \delta > 0$. Define $F(x) = \sum_{k\in\mathbb{Z}^n} f(x + 2k\pi)$. Show that $F \in L^1(\mathbb{T}^n)$ and compute the Fourier coefficients of F in terms of $\hat{f}$. [10]

(3) Let $1 \leq p, q, r \leq \infty$ satisfy $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$. Prove that

$$\|f * g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$$

for all $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$. [10]

(4) Let f be a real-valued Schwartz class function on $\mathbb{R}$. Show that

$$(Hf)^2 = f^2 + 2H(fHf),$$

where H is the Hilbert transform. [10]

(5) Prove that for almost every $\xi \in \mathbb{R}$,

$$\sum_{k\in\mathbb{Z}} \frac{1}{(\xi + 2k\pi)^2} = \frac{1}{4\sin^2(\xi/2)}.$$

[Hint. Consider the scaling function of the Haar MRA and use the fact that integral translates of a scaling function are orthonormal.] [5]

(6) Let φ be the scaling function of a multiresolution analysis so that there exists a sequence $\{a_k : k \in \mathbb{Z}\} \in \ell^2(\mathbb{Z})$ with

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k\varphi_{1,k}(x) = \sum_{k\in\mathbb{Z}} a_k 2^{1/2}\varphi(2x - k).$$

Show that

(a) $\sum_{k\in\mathbb{Z}} |a_k|^2 = 1$,

(b) $\sum_{k\in\mathbb{Z}} a_k\overline{a_{k+2l}} = \delta_{l,0}$ for all $l \in \mathbb{Z}$. [10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2018 – 19

M. Stat (2nd Year)

Time Series Analysis

Date: 26 November ~~December~~ 2018 | Maximum Marks: 100 | Duration: 3 Hours

Attempt ALL questions. Be brief and to the point.

You may use standard textbook notation automatically.

1. Consider the two series $x_t = w_t$ and $y_t = w_t - \theta w_{t-1} + u_t$, where w_t and u_t are independent white noise series with variances σ_w^2 and σ_u^2, respectively, and θ is an unspecified constant.

(a) Express the Autocorrelation function (ACF), $\rho_y(h)$ for h = 0, $\pm1, \pm2, \ldots$ of the series y_t as a function of σ_w^2, σ_u^2 and θ

(b) Determine the Cross-correlation function (CCF), $\rho_{xy}(h)$ relating x_t and y_t.

(c) Show that x_t and y_t are jointly stationary. [5 + 5 + 5 = 15]

2. Consider a process consisting of a linear trend with an additive noise term consisting of independent random variables w_t with zero means and variance σ_w^2,

$$x_t = \beta_0 + \beta_1 t + w_t,$$

where β_0 and β_1 are fixed constants.

(a) Prove x_t is non-stationary.

(b) Prove that the first difference series $\nabla x_t = x_t - x_{t-1}$ is stationary by finding its mean and auto-covariance function.

(c) Repeat part (b) if w_t is replaced by a general stationary process, say y_t, with mean function μ_y and auto-covariance function $\gamma_y(h)$. [4 + 6 + 5 = 15]

3. Consider the MA(1) series $x_t = w_t + \theta w_{t-1}$ where w_t is white noise with variance σ_w^2.

(a) Derive the minimum mean-square error one-step forecast based on the infinite past, and determine the mean-square error of this forecast.

(b) Let $\tilde{x}_{n+1}^{n} = \phi_1 \tilde{x}_n^n + \cdots + \phi_p \tilde{x}_{n+1-p}^n + \theta_1 \widetilde{w}_n^n + \cdots + \theta_q \widetilde{w}_{n+1-q}^n$ be the truncated one-step-ahead forecast. Show that

$$E[(x_{n+1} - \tilde{x}_{n+1}^n)^2] = \sigma^2(1 + \theta^{2+2n}).$$

Compare the result with (a), and indicate how well the finite approximation works in this case. [7 + (7 + 6) = 20]

4. first-order autoregressive model is generated from the white noise series w_t using the generating equation $x_t = \phi x_{t-1} + w_t$,

where ϕ, with $|\phi| < 1$, is a parameter and w_t are independent random variables with zero mean and variance σ_w^2.

(a) Show that the power spectrum of x_t is given by

$$f_x(\omega) = \frac{\sigma_w^2}{1 + \phi^2 - 2\phi\cos(2\pi\omega)}.$$

(b) Show that the inverse transform of the auto-covariance function of this process:

$$\gamma_x(h) = \frac{\sigma_w^2 \phi^{|h|}}{1 - \phi^2},$$

for h = 0, ±1, ±2, ... is indeed the spectrum derived in (a). [7 + 7 = 14]

5. Show that the 2 X 1 gradient vector for an ARCH(1) model

$$r_t = \sigma_t \varepsilon_t \text{ and } \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$$

is given by

$$\begin{pmatrix} \partial l / \partial \alpha_0 \\ \partial l / \partial \alpha_1 \end{pmatrix} = \sum_{t=2}^{n} \begin{pmatrix} 1 \\ r_{t-1}^2 \end{pmatrix} X \frac{\alpha_0 + \alpha_1 r_{t-1}^2 - r_t^2}{2(\alpha_0 + \alpha_1 r_{t-1}^2)^2}.$$

Then use the result to calculate the 2 x 2 Hessian matrix

$$\begin{pmatrix} \partial^2 l / \partial \alpha_0{}^2 & \partial^2 l / \partial \alpha_0 \partial \alpha_1 \\ \partial^2 l / \partial \alpha_0 \partial \alpha_1 & \partial^2 l / \partial \alpha_1{}^2 \end{pmatrix}.$$

[9 + 9 = 18]

6. Consider the model $y_t = x_t + v_t$, where v_t is Gaussian white noise with variance σ_v^2, x_t are independent Gaussian random variables with mean zero and $\text{var}(x_t) = r_t \sigma_x^2$ with x_t independent of v_t, and $r_1, \dots, r_n$ are known constants. Show that applying the EM algorithm to the problem of estimating σ_x^2 and σ_v^2 leads to updates (represented by hats)

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{t=1}^{n} \frac{\sigma_t^2 + \mu_t^2}{r_t} \text{ and } \hat{\sigma}_v^2 = \frac{1}{n} \sum_{t=1}^{n} [(y_t - \mu_t)^2 + \sigma_t^2],$$

where, based on the current estimates (represented by tildes),

$$\mu_t = \frac{r_t \tilde{\sigma}_x^2}{r_t \tilde{\sigma}_x^2 + \tilde{\sigma}_v^2} \text{ and } \sigma_t^2 = \frac{r_t \tilde{\sigma}_x^2 \tilde{\sigma}_v^2}{r_t \tilde{\sigma}_x^2 + \tilde{\sigma}_v^2}.$$

[18]

INDIAN STATISTICAL INSTITUTE

(203 B. T. Road, Kolkata 700 108)

Master of Statistics (M.Stat.) IInd Year
Academic Year 2018 - 2019: Semester I

Martingale Theory
Back Paper Examination
Instructor: Antar Bandyopadhyay

Date: 03.12.2018
Time: ______

Total Points: 100
Duration: 4 hours

Note:

- Please write your roll number on top of your answer paper.
- There are four problems carrying 25 points each. Solve as many as you can. The maximum you can score is 45.
- Show all your works and write explanations when needed.
- This is an open note examination. You are allowed to use your own hand written notes (such as class notes, your homework solutions, list of theorems, formulas etc). No printed materials or photo copies are allowed, in particular you are not allowed to use books, photocopied class notes etc.

1. State whether the following statements are ***true*** or ***false***. Write brief reasons supporting your answers. $[(1+4)\times 5 = 25]$

   (a) There is no i.i.d. sequence of random variables $(X_n)_{n\geq 0}$ which is a martingale.

   (b) Let $(X_n)_{n\geq 0}$ be an *exchangeable* sequence of random variables taking values in $\{0,1\}$. Then the tail σ-algebra of $(X_n)_{n\geq 0}$ is always trivial.

   (c) Suppose $\mu \ll \nu$ are two finite measures on a measure space $(\Omega, \mathcal{F})$. Let λ be another measure which is singular with respect to ν. Then $\mu \perp \lambda$.

   (d) It is not possible to have a pair of random variables (X, Y) such that the marginal distributions of X and Y are continuous, but the conditional distribution of X given Y is always discrete.

   (e) A predictable super-martingale converges on $[-\infty, \infty]$.

[please turn over]

2. (a) Give an example of a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and two sub-σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$, and a random variable X defined on $(\Omega, \mathcal{F}, \mathbf{P})$ such that

$$\mathbf{E}\left[\mathbf{E}\left[X \,\middle|\, \mathcal{F}_1\right] \,\middle|\, \mathcal{F}_2\right] \neq \mathbf{E}\left[\mathbf{E}\left[X \,\middle|\, \mathcal{F}_2\right] \,\middle|\, \mathcal{F}_1\right].$$

[5]

(b) Suppose $\mathbf{E}[X^2] < \infty$ and let $\mathcal{G}$ be a sub-σ-algebra. Then show that

$$\mathbf{Var}(X) = \mathbf{E}\left[\mathbf{Var}\left(X \,\middle|\, \mathcal{G}\right)\right] + \mathbf{Var}\left(\mathbf{E}\left[X \,\middle|\, \mathcal{G}\right]\right).$$

[10]

(c) Show that if X and Y are two square integrable random variables such that $\mathbf{E}\left[X \,\middle|\, Y\right] = Y$ and $\mathbf{E}\left[Y \,\middle|\, X\right] = X$. Then $X = Y$ a.s. [10]

3. (a) Let $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ be positive integrable and adapted to a filtration $(\mathcal{F}_n)_{n\geq 0}$. Assume

$$\mathbf{E}\left[X_{n+1} \,\middle|\, \mathcal{F}_n\right] \leq (1 + Y_n)\, X_n \;\; \forall\, n \geq 0.$$

Show that the two events $[\sum_{n=1}^{\infty} Y_n < \infty]$ and $\left[(X_n)_{n\geq 0} \text{ converges}\right]$ are almost surely equal. [10]

(b) If $(X_n)_{n\geq 0}$ is a non-negative super-martingale and $N \leq \infty$ is a stopping time then

$$\mathbf{E}[X_N] \leq \mathbf{E}[X_0].$$

[**Note:** Here we write X_∞ as the almost sure limit of X_n's which exists.] [15]

4. Suppose $(X_n)_{n\geq 1}$ be an exchangeable sequence of square integrable random variables.

(a) Show that $\mathbf{E}[X_i X_j] \geq 0$ for all $i, j \geq 1$. [10]

(b) Show that

$$\frac{1}{n(n-1)} \sum_{1\leq i<j\leq n} (X_i - X_j)^2 \xrightarrow{\mathcal{L}_1} \mathbf{Var}\left(X_1 \,\middle|\, \mathcal{E}\right) \text{ and also a.s.,}$$

where $\mathcal{E}$ is the exchangeable σ-algebra. [15]

Good Luck

INDIAN STATISTICAL INSTITUTE

Mid-Semester (Supplementary) Examination: Semester I (2018-19)

M. Stat. II Year

PATTERN RECOGNITION

Date: December 10, 2018 Maximum Marks: 50 Duration: 2 hr

1. Consider a two-class discrimination problem based on a univariate feature X. It is assumed that in the class ω_i, i=1,2, X is distributed as $\mathcal{N}(\mu_i, \sigma^2)$. Let π_1=π_2 $= \frac{1}{2}$, where π_i denotes the prior probability for the population ω_i, $i = 1,2$.

   a) Show that the minimum probability of error for a linear classification rule based on X for discriminating between ω_1 and ω_2 is

   $$P_e = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-z^2/2}\, dz,$$

   where

   $$a = \frac{|\mu_1 - \mu_2|}{2\sigma}.$$

   b) Show that $P_e \to 0$ as $\frac{|\mu_1-\mu_2|}{2\sigma} \to \infty$.

[10+5=15]

2.

   (a) What is meant by a kernel density estimator? Explain

   (b) Define the *Mean Integrated Square Error* (MISE) criterion for assessing the goodness of such estimators.

   (c) For kernel density estimation of a probability density function f which has continuous derivatives of all orders, if a symmetric probability density function with expectation equal to 0 and variance equal to $k_2(\neq 0)$ is used as the kernel, determine the optimal window width h which minimizes the Mean Integrated Square Error (MISE) in an approximate sense, as a function of the kernel K, f'' and the number of observations, n.

(PLEASE TURN OVER)

(d) If the density f is $\mathcal{N}(\mu, \sigma^2)$ and the kernel K used for estimating it on the basis of n observations is $\mathcal{N}(0, \sigma^2)$, show that the optimal value of h is approximately equal to

$$\frac{1.06\sigma}{\sqrt[5]{n}},$$

given that $\int[\phi''(x)]^2 dx = \frac{3}{8\pi}$, where ϕ denotes the probability density function of the standard normal distribution.

[3+3+10+4=20]

3. Create a binary tree classifier for a univariate two-class pattern recognition problem, presuming the availability of a very large number of observations on the feature X from the two classes ω_1 and ω_2, with

(i). $\pi_1 = \pi_2 = \frac{1}{2}$, where π_i denotes the prior probability for the population ω_i, $i = 1,2$;

(ii). $p(x|\omega_1) \equiv \mathcal{N}(0,1)$ and $p(x|\omega_2) \equiv \mathcal{N}(1,2)$.

The tree should be small with one root node, two other nonterminal nodes and four leaf nodes. All nonterminal nodes should have decisions of the form "Is $x \leq x_s$?" with respect to the misclassification impurity measure, deduce, in terms of the c.d.f. Φ of the standard normal distribution,

(a) the split value x_s at each of the three terminal nodes,

(b) the final test error of the tree classifier.

[12+3=15]

M.Stat Second Year, First Semester, 2018-19

Statistical Computing

Mid-Semestral (Supplementary) Examination

Time: $2\frac{1}{2}$ hours | Full Marks =60 | Date: 11.12.2018.

1. (a) Using only two random numbers from $U(0,1)$ distribution, how will you generate an observation from the two-dimensional uniform distribution on $S = \{(x_1, x_2) : 1 \leq x_1^2 + x_2^2 \leq 4\}$? Find the mean vector and the dispersion matrix of this distribution. [6+2+6]

   (b) Consider a non-negative continuous function $\psi : [a, b] \to [0, c]$, In order to estimate $\theta = \int_a^b \psi(x)\, dx$, a student generated n independent observations $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ from $U(a, b) \times U(0, c)$ and proposed the estimate $\hat{\theta}_1 = \frac{(b-a)c}{n} \sum_{i=1}^{n} 1_{\{y_i \leq \psi(x_i)\}}$. Another student generated n independent observations $z_1, z_2, \ldots, z_n$ from $U(a, b)$ and proposed the estimate $\hat{\theta}_2 = \frac{b-a}{n} \sum_{i=1}^{n} \psi(z_i)$. Which of these two estimates will you prefer? Justify your answer. [6]

2. (a) Show that simplicial depth is invariant under affine transformations. [3]

   (b) If F is a spherically symmetric distribution, which is symmetric about the origin, show that $SD(\mathbf{x}, F)$, the simplicial depth of $\mathbf{x}$ with respect to F, is a function of $\|\mathbf{x}\|$. [4]

   (c) Show that the simplicial depth of an observation $\mathbf{x}$ with respect to $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution is a function of the Mahalanobis distance $\{(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}^{1/2}$. [3]

3. (a) A a random sample of size 25 is to be drawn from $\{1, 2, \ldots, 25\}$ using SRSWR technique.

   (i) Calculate the total number of distinct samples that can be drawn from $\{1, 2, \ldots, 25\}$. [4]

   (ii) If $\tilde{X}$ denotes the median of a random sample, find $P(\tilde{X} = k)$ for $k = 1, 2, \ldots, 25$. [5]

   (iii) Check whether $\tilde{X}$ is unbiased for the population median. [5]

   (b) Show that the bootstrap estimate of $Var(\bar{X}_n)$ is biased, but the jackknife estimate of $Var(\bar{X}_n)$ is unbiased. [6]

4. (a) Give an algorithm for generating two random variables X_1 and X_2 such that $X_1 \sim U(0,1)$, $X_2 \sim U(0,1)$, they are uncorrelated but not independent. [5]

   (b) Is it possible to have a bivariate distribution, where the half-space depth of the spatial median is higher than 0.5? Justify your answer. [5]

INDIAN STATISTICAL INSTITUTE

Supplementary Mid-Semester Examination : 2018-2019

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 12 December, 2018 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]

Date: 31 · 12 · 2018

INDIAN STATISTICAL INSTITUTE

M.Stat Second Year, First Semester, 2018-19

Semestral (Backpaper) Examination

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

1. (a) Describe how you will generate observations on two random variables X and Y such that X follows $Bin(5, 0.5)$ distribution, Y follows $Bin(10, 0.5)$ distribution, and the correlation co-efficient between X and Y is -0.25. [6]

   (b) Let 0.3862, 0.7908, 0.0246 and 0.1325 be four independent observations generated from $U(0,1)$ distributions. Using these observations, generate an observation (X_1, X_2) from the following bivariate distribution with cumulative distribution function given by [8]

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

2. Consider a moving average process $X_t = \mu + \varepsilon_t + \varepsilon_{t-1}$ for $t = 1, 2, \ldots$, where the ε_t's are i.i.d. with the mean 0 and the variance σ_0^2.

   (a) Define $\bar{X}_n = \sum_{i=1}^n X_i/n$. Find the asymptotic distribution of $Y_n = \sqrt{n}(\bar{X}_n - \mu)$. [4]

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

3. Prove or disprove the following statements:-

   (a) Simplicial depth of an observation $\mathbf{x}$ with respect to the multivariate normal distribution $N(\boldsymbol{\mu}, \Sigma)$ is a function of the Mahalanobis distance $[(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})]^{1/2}$. [8]

   (b) Both least squares (LS) and least absolute deviation (LAD) estimators of regression coefficient in simple linear regression have zero asymptotic breakdown point. [6]

   (c) For univariate data, the Minimum Covariance Determinant (MCD) estimate of location turns out to be the Least Trimmed Squares (LTS) estimate of location. [8]

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

   (e) Consider a data set $\{(x_i, y_i);\ i = 1, \ldots, n\}$, where $y_i = f(x_i)$ for $i = 1, \ldots, n$ and f is monotonically increasing. For any fixed choice of the bandwidth h, the Nadaraya-Watson estimate $\hat{f}_h$ based on the Gaussian kernel will also be monotonically increasing. [8]

[P.T.O.]

4. (a) Consider a regression problem with observations $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Suppose that a linear spline with power spline basis and knots $t_0 < t_1 < \ldots < t_{10}$ is used to estimate the regression function f. If there are no x_is in the interval (t_4, t_6), show that the projection matrix is singular. [4]

(b) Give an example of a regression problem, where projection pursuit regression is expected to peform better than regression based on additive model. Describe how you will fit a projection pursuit regression model to a data set on a response variable Y and p covariates $X_1, X_2, \ldots, X_p$. [2+4]

(c) Let $\hat{f}_h$ be the Nadaraya Watson estimate of a regression function f based on a kernel K and bandwidth h. Assume that f is Lipschitz continuous, and K has a bounded support. Show that for any fixed x, bias of $\hat{f}_h(x)$ converges to 0 as the bandwidth h shirnks to 0. [4]

5. (a) Let $x_1, x_2, \ldots, x_n$ be n observations generated from a mixture of three univariate normal distributions $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ and $N(\mu_3, \sigma_3^2)$ (where $\mu_1 < \mu_2 < \mu_3$) with mixing proportions p^2, $2pq$ and q^2 $(p + q = 1)$, respectively. Describe how you will use the expectation-maximization (EM) algorithm (clearly state your choice for the initial values of the parameters and the convergence criterion) to estimate p. [8]

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

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

(i) $Bin(2m, 1/2)$ is the stationary distribution of the Markov chain $\{X_n;\ n \geq 0\}$. [5]

(ii) $\sup_k |P(X_n \leq k) - \sum_{i=0}^{k} \binom{2m}{i} (\frac{1}{2})^{2m}| \to 0$ as n diverges to infinity. [4]

(b) Describe how you will use the Gibbs sampling algorithm to generate observations from the following distribution

$$f(x, y, z) = C \left(\frac{y}{1 - y} \right)^x \frac{\{10(1 - y)\}^z}{x!\ (z - x)!}, \quad z = 1, 2, \ldots;\ x = 0, 1, \ldots, z;\ 0 \leq y \leq 1,$$

where C is a normalizing constant. [5]

INDIAN STATISTICAL INSTITUTE

Semester Examination: 2018-19 (First Semester)

M. Stat. II Year

Fourier Analysis (Backpaper)

Date: 01/02/2019 Maximum Marks: 60 Duration: 3 Hours

Note: Give proper justification to your answers. State clearly all the results you are using.

---

(1) Let f be integrable on $\mathbb{T}$ and is differentiable at $x_0 \in \mathbb{T}$. Show that $S_N f(x_0) \to f(x_0)$ as $N \to \infty$, where $S_N f$ is the N-th partial sum of the Fourier series of f. **[20]**

(2) Let $\{c_n : n \in \mathbb{Z}\} \in \ell^2(\mathbb{Z})$. Prove that there exists a unique $f \in L^2(\mathbb{T})$ such that $\hat{f}(n) = c_n$ for all $n \in \mathbb{Z}$. **[15]**

(3) Show that there exists a constant $C > 0$ such that

$$\sup_{0<r<1} |P_r f(x)| \leq CMf(x), \quad x \in \mathbb{T},$$

where P_r is the Poisson kernel and the M is the Hardy-Littlewood maximal operator on $\mathbb{T}$. **[15]**

(4) Compute the Fourier transform of the function f, where $f(x) = e^{-|x|^2/2}$, $x \in \mathbb{R}$. **[10]**

(5) Let δ_0 be the Dirac mass at 0. Compute the Fourier transform of the distribution $\partial^\alpha \delta_0$, where α is a multiindex. **[10]**

(6) Let $H = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ be the upper half-plane in $\mathbb{C}$. For $F \in L^2(0, \infty)$, define

$$f(z) = \int_0^\infty F(t)e^{izt}\, dt, \quad z \in H.$$

Show that f is analytic in H. **[10]**

(7) Let $f \in L^1(\mathbb{R}^n)$, $t \in \mathbb{R}^n$ and $\epsilon > 0$. Show that there exists $h \in L^1(\mathbb{R}^n)$ with $\|h\|_1 < \epsilon$ such that $\hat{h}(s) = \hat{f}(t) - \hat{f}(s)$ for all s in a neighbourhood of t. **[10]**

(8) Let $g \in L^2(\mathbb{R})$ such that $\{g(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal system in $L^2(\mathbb{R})$. Show that $|\operatorname{supp} \hat{g}| \geq 2\pi$. Also prove that equality holds if and only if $|\hat{g}| = \chi_K$ for some $K \subset \mathbb{R}$. **[10]**

INDIAN STATISTICAL INSTITUTE

Backpaper Examination

First Semester, 2018-2019

M.Stat. 2nd Year

STATISTICAL INFERENCE II

Date: 1 February, 2019 Max. Marks: 100 Duration: 3 Hours

Answer all questions

1. Let $X_1, \ldots, X_n$ be i.i.d. $N(\theta, \sigma^2)$ variables. Consider a standard noninformative prior for (θ, σ^2) and find the posterior distribution of θ. Also find the $100(1-\alpha)\%$ HPD credible set for θ.

[(6+4)=10]

2. Consider i.i.d. observations $X_1, \ldots, X_n$ with a common density $f(x|\theta) = \exp\{c(\theta)+\theta t(x)\}h(x)$ (one parameter exponential family). It is given that the usual regularity conditions hold and $c(\theta)$ is sufficiently smooth. A statistically natural parameter is $\mu = E_\theta t(X)$. Show that μ is a one-one function of θ. Find MLE of μ. Show that for a conjugate prior $\pi(\theta) = c\exp\{mc(\theta) + \theta s\}$, the posterior mean of μ is a weighted average of the prior estimate and the classical estimate (MLE). (assume that π is supported on $[a, b]$ where $\pi(a) = \pi(b) = 0$.)

[15]

3. What are the difficulties with improper noninformative priors in Bayes testing? Describe the intrinsic Bayes factor (IBF) and the fractional Bayes factor (FBF) as solutions to this problem with improper priors.

[(3+10)=13]

4. 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 $\theta_1, \ldots, \theta_p$.

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

[(15+7)=22]

5. Consider the set up of Question (4) 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 $\boldsymbol{\theta}, \sigma^2, \tau^2$.

[17]

6. 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.

[12]

7. Suppose X has density $e^{-(x-\theta)} I(x > \theta)$ and the prior density of θ is $p(\theta) = [\pi(1+\theta^2)]^{-1}$, $\theta \in R$. Find the Bayes estimate of θ for the loss function $L(\theta, a) = I(|\theta - a| > \delta)$ for some specified $\delta > 0$.

[11]

INDIAN STATISTICAL INSTITUTE

MID-SEMESTRAL EXAMINATION

M-Stat. (2nd year) 2018-19

Subject: Theory of Games and Decision

Date: 18.02.2019 | Duration: 1hr. 30mins | Full Marks: 40

Attempt ALL questions

1a) Consider the following game in Statistical Decision Theory,

Let set of states of Nature = $\{\theta_1, \theta_2, \theta_3\}$

Let $x_i = r(\theta_i, d)$, $i = 1, 2, 3$ and $d \in$ Decision space

Let risk set

$A = \{x = (x_1, x_2, x_3) : 2(x_1 - 3)^2 + (x_2 - 4)^2 + (x_3 - 6)^2 \leq 22\}$

Find value of the game, minimax point and Least favourable prior.

[2 + 2 + 4]

b) Let $V(A)$ = value of the mixed extension of matrix game A, for matrix $A : m \times n$.

Show that $V(A)$ is a continuous function from the set of mxn matrices to R.

[5]

c) Let A be an $n \times n$ positive definite matrix.

$V(A)$ = value of the mixed extension of the matrix game A

Show that $V(A) \geq (1/n)\, \lambda_{min}$,

Where λ_{min} = minimum eigenvalue of A

[5]

2. a) For $f(a, b)$ a real-valued function on $A \times B$, define $(a_0, b_0) \in A \times B$ to be a saddle point of f.

State and prove necessary and sufficient condition on the existence of saddle point in this context.

[2 + 12]

b) Define value of a two-person, zero-sum non-cooperative game.

Show that value of the game is unique in this context (if it exists)

Prove rectangular property of the set of equilibrium situations in this context.

[2 + 3 + 3]

******

Indian Statistical Institute

Second Midsemestral Examination 2018-19

M. Stat. II yr

Statistical Inference III

Date: February 19, 2019 (14:30 hrs) Maximum marks: 60 Duration: 2 hrs.

Answer all Questions.

1. Let $\{X_0, X_1, \ldots, X_k\}$ be a segment of a time-homogeneous discrete Markov chain with transition probability matrix P. Assume that the initial random variable X_0 has an initial pmf belonging to $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ (θ unknown) over the discrete state space of the Markov chain. Further let $T(X_0)$ be a sufficient statistic for the family $\mathcal{P}$. Next consider two experiments $\mathcal{E}^T = \{T\}$ and $\mathcal{E} = \{X_0, X_1, \ldots, X_k\}$ respectively. Show that deficiency $\delta(\mathcal{E}^T, \mathcal{E}) = 0$ using Theorem 2 of LeCam and Yang (2000) or otherwise.

[20]

2. Consider two random observations (Y, T) taking values jointly in $(\mathcal{Y} \times \mathcal{T}, \mathcal{A} \times \mathcal{B})$. Suppose that under two different hypotheses H_0 and H_1 their joint distributions P_0 and P_1 can be represented by joint densities $f(y)\,g(t)$ and $h(y,t)$ respectively with respect to a fixed product measure $\mu \times \nu$ on $(\mathcal{Y} \times \mathcal{T}, \mathcal{A} \times \mathcal{B})$. Assuming $h(y,t)/f(y)g(t) < \infty$ show that the marginal density of Y on $(\mathcal{Y}, \mathcal{A})$ is given by

$$p_Y(y) = f(y)\ \mathrm{E}_0\left[\frac{h(y,T)}{f(y)g(T)}\,\middle|\,Y = y\right] \quad \text{a.e } \mu,$$

where E_0 is expectation under H_0.

[20]

3. Consider two experiments on a sequence of Bernoulli trials $X_1, X_2, X_3, \ldots$ with success probability θ, $0 < \theta < 1$ (unknown) given by $\mathcal{E}_1 = \{X_1, X_2, \ldots, X_5\}$ and $\mathcal{E}_2 = \{\tau = \min\{m - 3 : X_1 + X_2 + \cdots + X_m \geq 3\}\}$ respectively. Suppose we perform a mixture experiment $\mathcal{E}^*$ by choosing one from $\mathcal{E}_1$ and $\mathcal{E}_2$ with probability 0.5. Write down the sample space of the mixed experiment explicitly and find a minimal sufficient statistic for $\mathcal{E}^*$.

[20]

Indian Statistical Institute

M.Stat. Second year, Mid Semester Exam: 2019

Topic: Statistical Computing II

Maximum Marks: 50, Duration: 2 hours Date: 20.2.2019

1. Consider a continuous longitudinal response variable Y and a set of covariates X, for n individuals. Under the usual assumptions, derive the generalized estimating equation (GEE) for this correlated data. Specify the underlying sandwich structure of the working covariace matrix, and mention why GEE results in a robust estimate for the model parameters. [10]

2. In a disease mapping project, suppose data are collected on the number of cases of disease from S different locations, with Y_s be the observed number of cases for the location s, where $s = 1, 2, \ldots, S$. Data are also collected on a set of related covariates, say, $x_1, x_2, \ldots, x_5$; from all S locations. In the data, $s_1 (< S)$ responses are found to be exactly equal to zero and non-zero responses are denoted by $Y_1, Y_2 \ldots, Y_{s_2}$, where $s_1 + s_2 = S$. Considering a suitable parametric covariance function, develop a Bayesian Two-Part model for the above dataset with the goal of predicting the number of cases of disease for a new location s^* given the relevant information on the covariates for that location. Show all the steps for your modelling and prediction. [15]

3. A study was conducted on metabolism in 15 year old females. Data are collected from 20 subjects, and let X_i be the energy intake over a 24 hour period for the i-th individual. We model $log(X)$ with Normal(θ, σ^2), and further consider a Normal(θ_0, τ^2) prior for θ, and an Inverse Gamma (a, b) prior for σ^2. Note that the density of an Inverse Gamma distribution with parameters (a, b) is given as the following:
$f(\sigma^2|a, b) \propto \frac{1}{(\sigma^2)^{a+1}} exp(\frac{1}{b\sigma^2})$.

   In a Bayesian framework, derive explicitly the joint posterior distribution, the full conditional distributions for θ and σ^2, and explain the Gibbs sampler algorithm (by identifying the respective full conditional distributions) for estimating the underlying model parameters. [15]

4. Define the following terms:
(i) Directed graph, (ii) Bayesian network, (iii) Importance sampling, (iv) Horseshoe prior for variable selection, (v) Quantile crossing. [10]

INDIAN STATISTICAL INSTITUTE

ROBUST STATISTICS

M-Stat II, Second Semester 2018-2019

Midterm Examination

Date: 21.02.2019 Maximum marks: 60 Duration: 2 hours.

1. (a) Given two densities g and f with respect to the same measure, define the Bregman divergence between them defined by a strictly convex function $B(\cdot)$. [Here, as well as in the other parts of this problem, you may assume that B is twice continuously differentiable.]

   (b) Suppose that a random sample $X_1, X_2, \ldots, X_n$ is available from an unknown distribution G, which is modeled by the parametric family $\{F_\theta; \theta \in \Theta\}$. Explain how one can set up an empirical version of the Bregman divergence between the unknown data density g and the model density f_θ.

   (c) Derive the influence function of the minimum Bregman divergence estimator of θ in terms of the defining convex function $B(\cdot)$.

[3+3+8=14]

2. Consider the location model $F_\mu(x) = F_0(x - \mu)$ where F_0 is symmetric about zero. Let ψ be a nondecreasing and bounded ψ function; let $k = \psi(\infty)$. Consider the mixture distribution $H = (1-\epsilon)F_\mu + \epsilon G$, where G is an arbitrary contaminating distribution. Define $m(b) = E_{F_0}\psi(X + b)$. Notice that $m(\cdot)$ is odd; it will be assumed that m is increasing. Let $T(F_\mu) = \mu$, and $T(H) = \mu_H$, where $T(\cdot)$ represents the functional corresponding to the M-estimator based on the function ψ. The quantity $\mu_H - \mu$ is the bias in estimation. Let $\epsilon < 0.5$. Show that the maximum bias in estimation is the solution b_ϵ of the equation
$$m(b) = \frac{k\epsilon}{1-\epsilon}.$$
[Hint: Without loss of generality, assume $\mu = 0$, so that the bias is simply μ_H. Look at the estimating equation for the mixture distribution.]

[10]

3. (a) Define a statistical functional $T(F)$. Give two examples of statistical functionals.

   (b) Under appropriate assumptions and notations, define the von-Mises derivative of a statistical functional. How is the influence function of a functional related to the von-Mises derivative?

(c) Explain the dual role of the influence function in describing the robustness and efficiency of an estimator.

[3+4+7=14]

4. (a) Let $X_1, \ldots, X_n$ be i.i.d. observations from a discrete distribution supported on $\chi = \{0, 1, 2, \ldots\}$. Let G be the true distribution function having probability mass function $g(x)$; let $f_\theta(x)$ represent the probability mass function of the model. Also let $d_n(x)$ be the relative frequency at $x \in \chi$. Define the disparity $\rho_C(d_n, f_\theta)$ between d_n and f_θ based on the function $C(\cdot)$. Describe the properties that the defining function $C(\cdot)$ must have.

(b) What is the residual adjustment function (RAF) $A(\cdot)$ of a disparity? How can one standardize the disparity so that $A(0) = 0$ and $A'(0) = 1$.

(c) Find the RAF of the Hellinger distance and the Neyman's chi-square.

[3+3+4=10]

5. (a) One definition of breakdown says that the functional $T(\cdot)$ breaks down at the distribution F at the level of contamination ϵ, if $|T(H_n) - T(F)| \to \infty$, as $n \to \infty$, where $H_n = (1 - \epsilon)F + \epsilon V_n$, and the sequence $\{V_n\}$ represents arbitrary contaminating distributions. When the parametric model is the family of Poisson(λ) distributions and the true distribution belongs to the model, show that the minimum Hellinger distance estimator has asymptotic breakdown point $\epsilon^* = 0.5$.

(b) Consider a location M-estimator corresponding to a ψ-function which is monotonic (but not necessarily odd). Suppose that $k_1 = -\psi(-\infty)$ and $k_2 = \psi(\infty)$ are finite. Show that the asymptotic breakdown point of the location M-estimator is given by

$$\epsilon^* = \frac{\min\{k_1, k_2\}}{k_1 + k_2}.$$

[6+6=12]

MSTAT II - Computational Finance

Midsem. Exam. / Semester II 2018-19

Date - February 22, 2019 / Time - 2 hours

Maximum Score - 30

<u>NOTE</u> : **SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.**

1. (5+7=12) Consider the followig Black-Schole-Merton PDE $\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} = rf$ with the boundary condition $f(S,T) = max(K-S,0)$ for some positive K, r and σ.

   (a) Use explicit difference method to formulate this equation and find the condition for stability, if any.

   (b) Transform the B-S-M equation using $Z = logS$ and then find the stability condition for explicit difference method for the transformed equation. Further find the stability condtion, if any, for implicit difference method for this transformed equation.

   [Hint: You may use that the eigenvalues of the tridiagonal matrix A of dimension $N \times N$ is of the form $\lambda_i = \alpha + 2\beta\sqrt{\frac{\gamma}{\beta}}Cos(\frac{i\pi}{N+1})$ where $A = ((a_{ij}))$ given by $a_{ij} = \alpha$ for $j = i$, $a_{ij} = \beta$ for $j = i+1$, $a_{ij} = \gamma$ for $j = i-1$ and $a_{ij} = 0$, otherwise.

   ]

2. (5+5=10) Evaluate the integration $\int_1^4 f(x)dx$ using the following two methods and compare.

   (i) Gaussian quadrature formula $\sum_{i=1}^{n} w_i f(x_i)$ for $n = 3$, where $f(x) = \exp(x^4)$, where the w_is are the appropriately chosen weights using Legendre polynomials over the interval $[-1,1]$ which are given by $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = (1/2)(3x^2-1)$ and $p_3(x) = (1/2)(5x^3-3x)$.

   (ii) Simpson's 1/3rd rule dividing them into 3 equal interval.

3. (5+5+3=13) Consider the SDE, $dr_t = k(\theta - r_t)dt + \sigma dW_t$ where k, σ and θ are positive constants and $\{W_t\}$ is a standard Brownian motion.

   (i) Solve for r_t, over the time period $[s,T]$, for $0 \le s \le T$.

   (ii) Let P_s^T be the price of a zero coupon bond at time s that matures at time $T(> s)$ and given by $E[\exp(-\int_s^T r_t dt) \mid \mathcal{F}_s]$ (under risk-neutral measure), where $\mathcal{F}_s = \sigma(\{W_u : 0 \le u \le s\} \cup \{r_0\})$, r_0 is independent of the Brownian motion $\{W\}$. Evaluate P_s^T.

   (iii) For $k = 2$, $\theta = 0.04$, $T = 3$, $\sigma = 0.25$, $r_s = 0.04$ and $s = 1$ months, find the value of P_s^T.

All the best.

Indian Statistical Institute
Mid-semestral Examination

September 25, 2018

Weak Convergence and Empirical Processes, M1

Total points: 20 Time: $1\frac{1}{2}$ hours

Note: This is a **closed notes/closed book** examination. Do any **four** of the following problems. Notations, if not explicitly explained, are to be interpreted as defined in class.

1. Closure of open balls. [2 + 3]

Let $(\mathcal{M}, \rho)$ be a metric space, $x \in \mathcal{M}, \epsilon > 0$. Show that $\overline{B(x;\epsilon)} \neq B[x;\epsilon]$ in general. If, however, $\mathcal{M}$ is a normed vector space, with ρ coming from the norm, then show that $\overline{B(x;\epsilon)} = B[x;\epsilon]$.

2. Tightness of a Gaussian family. [5]

Let $A \subset \mathbb{R} \times (0, \infty)$. Set $\Pi_A = \{\mathbb{P}_{\mu,\sigma^2} \mid (\mu, \sigma^2) \in A\}$, where $\mathbb{P}_{\mu,\sigma^2}$ is a $\mathcal{N}(\mu, \sigma^2)$ probability measure. Show that Π_A is tight if and only if A is bounded.

3. Weak convergence in countable and discrete spaces. [3 + 2]

Suppose that $\mathcal{M}$ is countable and discrete, and let $\mathbb{P}_n, \mathbb{P}$ be Borel probability measures on $\mathcal{M}$. Show that $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ if and only if $\mathbb{P}_n\{x\} \to \mathbb{P}\{x\}$ for all $x \in \mathcal{M}$. Show also that, in this case, $d_{\mathrm{TV}}(\mathbb{P}_n, \mathbb{P}) := \sup_{A \in \mathcal{B}(\mathcal{M})} |\mathbb{P}_n(A) - \mathbb{P}(A)| \to 0$.

4. Some Brownian computations. [2 + 3]

(a) Let X, Y, Z be random variables defined on a common probability space. If $(X, Y) \perp\!\!\!\perp Z$, then show that $\mathbb{E}[X \mid Y, Z] = \mathbb{E}[X \mid Y]$ a.s.

(b) Using the above result, or otherwise, compute (i) $\mathbb{E}[B(s) \mid B(r), B(t)]$, (ii) $\mathrm{Var}[B(s) \mid B(r), B(t)]$, where $0 \leq r < s < t \leq 1$.

5. The number of levels reached by a simple symmetric random walk. [2 + 3]

Consider a simple symmetric random walk $(S_k)_{k \geq 0}$, and let L_n denote the number of levels reached by the random walk by time n. Thus $L_0 = 1, L_1 = 2, L_2 \in \{2, 3\}$, and so on. Show that $\frac{L_n}{\sqrt{n}}$ converges weakly and give an expression for its limit CDF.

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2018-19

MStat II Year

Survival Analysis

Date: 25 February 2019 | Maximum marks: 60 | Duration: 2 hours

This examination is ~~open~~ closed books and ~~open~~ closed notes. The entire question paper is for 65 marks. The maximum score is 60.

1. Obtain a closed form expression for the mean residual life of a unit at age t, if its hazard rate function is $\lambda(t) = 2 + 4t$. [8]

2. Suppose $X_1, \dots, X_n$ is a randomly right censored sample from a life distribution having survival function $S(t) = e^{-(\lambda t)^2},\ t > 0$, where λ is an unknown and positive valued parameter.

   (a) Derive the maximum likelihood estimator of λ.

   (b) Derive the maximum likelihood estimator of $S(0.5)$.

   (c) Describe the asymptotic distribution of the estimator of part (b).

   (d) Describe a confidence interval of $S(0.5)$ with asymptotic coverage probability 0.95. [6+2+4+3=15]

3. Show that the Kaplan-Meier estimator of the survival function reduces to a simple function of the empirical distribution function when the data are not censored. [6]

4. For the data set 3, 5, 5+, 23, 35, 48+, 64, calculate the Kaplan-Meier estimator by using the redistribute-to-the right algorithm. Also obtain the nonparametric maximum likelihood estimate of the mean. [8+4=12]

5. Consider the counting process formulation of survival data with random right censoring.

   (a) Specify Aalen's class of two-sample test statistics for the equality of two deterministic intensity functions, after describing the set-up.

   (b) Describe the asymptotic null distribution of the test statistic, with justification. (There is no need to provide an expression for the asymptotic variance.)

   (c) How can the asymptotic variance of the test statistic be estimated consistently?

   (d) Show that the test of part (a) reduces to the log-rank test for a particular choice of the weight function.

   (e) How does the expression of part (c) simplify in the special case of (d)?

   [5+6+3+8+2=24]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2018 – 19

MStat (2nd Year)

Financial Econometrics

Date: 25th February 2019 Maximum Marks: 30 Duration: 2 Hours

1. (a) How do you calculate power of the 3 – period variance ratio statistic VR (3) under the AR(1) alternative?

(b) In particular, explicitly calculate the alternative distribution when the AR parameter $\varphi = 0.3$? [6 + 3 = 9]

2. Consider the model of non-synchronous trading. If we look at the return of q – period time aggregation for a portfolio, show that the maximal spurious autocorrelation $\rightarrow 1$ as the non-trading probability π_a, for portfolio a, $\rightarrow 1$, for a fixed q. [10]

3. How do you calculate power for tests using

(a) Cumulative abnormal return for security i, $\widehat{CAR_i}$?

(b) Cumulative average abnormal return $\overline{CAR}$? [6 + 5 = 11]

INDIAN STATISTICAL INSTITUTE

Mid-Sem Examination (2nd Semester): 2018-19

M. STAT. 2nd year

Subject Name: Clinical Trials

Date: 26 ·2, 2019, Maximum Marks: 40. Duration: 2 hrs.

Group A: Answer all questions.

1. (a) What are the advantages of Zelen's play-the-winner rule over an equal allocation rule? If n patients are treated by this rule in a two-treatment set up, find the expression of the proportion of allocation by a particular treatment. [4+8]

   (b) Suppose a Cyclic Play-the-Winner rule is carried out for a 4-treatment clinical trial with success probabilities 0.8, 0.6, 0.4 and 0.2, respectively. Find the probability of allocating the 4th patient to the first treatment. [8]

2. (a) Discuss how type I error spending function (and a discrete set of boundaries) can be constructed for a discrete stochastic process in group sequential set up. [8]

   (b) If there are only three groups in a group sequential study, give one form of type I error spending function which will spend the total type I error in a 1:7:56 fashion in the three groups (assuming the time intervals are in the ratio 1:1:2 for the three groups). [5]

3. What is Friedman and Wei's Urn Design? Write the expression for the proportion of balls of any kind in the urn up to the first n patients as a function of the difference of the number of patients treated by the two treatments up to that patient. [3+4]

Mathematical Biology
Mid Semestral Examination
M.Stat.-II Year. 2018-2019
Total Marks-40
Time-2 hrs.
Indian Statistical Institute
Kolkata-700108

1. Prey population grow in a logistic fashion and the interaction between predator and prey follow Holling type I functional response. The death of predator population is density dependent.

   (a) Following the above assumptions, write down the mathematical model in terms of ordinary differential equations.
   (b) Show that the solutions which initiate in R_2^+ are uniformly bounded.
   (c) Find out the condition(s) for which both the population will coexist?
   (d) By constructing a suitable Lyapunov function, show that the system is globally asymptotical stable around the positive interior equilibrium point.

(2+4+4+5)=15

2. Consider N population of Yeast cells are growing by fission. There is a certain probability that a particular cell will divide in a given time interval. Investigate the stochastic form of the process by calculating the mean and variance.

(12)

3. (a) Derive Holling disc equation with proper assumptions.
   (b) The Rosenzweig-MacArthur predator-prey model in dimensionless form is written as

$$\frac{dx}{dt} = x\left(1 - \frac{x}{K}\right) - \frac{axy}{1+x}$$
$$\frac{dy}{dt} = -cy + \frac{axy}{1+x}$$

   where the symbols have their usual meaning. By using Dulac criteria, show that there is no periodic orbit around the positive interior equilibrium point.
   (c) What is Allee effect. Write down a typical model of strong Allee effect with harvesting and point out the transcritical bifurcation point.

(3+5+5)=13

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination (2018–2019)

M STAT II

Advanced Functional Analysis

Date : 27.02.2012 Maximum Marks : 50 Time : $2\frac{1}{2}$ hrs.

This paper carries 55 marks. Maximum you can score is 50. Precisely justify all your steps. Carefully state all the results you are using.

1. Let Y be a w*-closed subspace of a dual Banach space X^*. Let $x^* \in X^*$. Show that there exists $y^* \in Y$ such that

$$\|x^* - y^*\| = \inf\{\|x^* - z^*\| : z^* \in Y\}.$$

[Hint : The norm on X^* is w*-lower semi-continuous.] [10]

2. (a) Show that the space c_0 is not isometric to the dual of any Banach space. [10]

   (b) Show that the space ℓ^1 is not reflexive. [10]

3. Let X be a Banach space. For a closed bounded convex set $K \subseteq X$, $x_0 \in K$ is called a *denting point of* K if, for each $\varepsilon > 0$,

$$x_0 \notin \overline{co}(K \setminus B_\varepsilon(x_0)).$$

   (a) Show that $x_0 \in K$ is a denting point of K if and only if for each $\varepsilon > 0$, there is a slice $S(K, f, \alpha)$ such that $x_0 \in S(K, f, \alpha)$ and the diameter of $S(K, f, \alpha)$ is less than ε. [10]

   (b) Show that a denting point is always an extreme point. [5]

   (c) Show that the converse holds if K is compact. [10]

INDIAN STATISTICAL INSTITUTE

M.STAT SECOND YEAR

Second Semester, 2018-19

Dt.. 22.04.19

Inference for High Dimensional Data

Time: $3\frac{1}{2}$ Hours | Semestral Examination | Full Marks: 100

[Answer all questions. The maximum you can score is 100.]

1. (a) Show that Benjamini-Hochberg's step up method controls the false discovery rate at the desired level even when the test statistics are positively regression dependent. [6]

   (b) What modification will you suggest for negatively regression dependent test statistics? Justify your answer. [2+4]

2. (a) Consider a data cloud $\Omega_n = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ consisting of d-dimensional observations, where $\mathbf{x}_i \neq \mathbf{0}$ for all $i = 1, 2, \ldots, n$. If $d > n$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent, show that the half-space depth of the origin with respect to this data cloud is 0. [6]

   (b) Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be independently distributed d-dimensional standard normal variates. Show that as the dimension diverges to infinity, $\mathbf{X}_1$, $\mathbf{X}_2$ and the origin tend to lie on the vertices of a right-angled isosceles triangle with probability tending to one. [6]

3. Consider a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is non-stochastic, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $Var(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$.

   (a) If $\mathbf{X}^\top\mathbf{X} = \mathbf{I}$, show that the LARS algorithm gives the entire LASSO path. [6]

   (b) What is LASSO modification of LARS algorithm? [2]

   (c) Give an example, where instead of LASSO, one should use the elastic net algorithm. Give justification for your answer. [4]

4. (a) Suppose that F is a multivariate distribution symmetric about $\boldsymbol{\theta}$, and $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are independent copies of $\mathbf{X} \sim F$. Also assume that $E(\|\mathbf{X}\|) < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. Define

$$T = \binom{n}{2}^{-1} \sum_{i<j} \Big\{ \|\mathbf{X}_i + \mathbf{X}_j\| - \|\mathbf{X}_i - \mathbf{X}_j\| \Big\}.$$

   Show that $E(T) \geq 0$, where the equality holds if and only if $\boldsymbol{\theta} = \mathbf{0}$. [6]

   (b) Let F and G be two multivariate distributions with finite first moments. If $\mathbf{X}_1, \mathbf{X}_2 \sim F$ and $\mathbf{Y}_1, \mathbf{Y}_2 \sim G$ are independent, show that

$$E\|\mathbf{X}_1 - \mathbf{X}_2\|_1 + E\|\mathbf{Y}_1 - \mathbf{Y}_2\|_1 \leq E\|\mathbf{X}_1 - \mathbf{Y}_2\|_1 + E\|\mathbf{Y}_1 - \mathbf{X}_2\|_1,$$

   where $\|\cdot\|_1$ denotes the ℓ_1 norm. When does the equality hold?. [4+2]

P.T.O.

5. (a) Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimensions p and q, respectively. Show that they are independent if and only if $\boldsymbol{\alpha}^\top \mathbf{X}$ and $\boldsymbol{\beta}^\top \mathbf{Y}$ are independent for all $\boldsymbol{\alpha} \in R^p$ and $\boldsymbol{\beta} \in R^q$. [3]

(b) Using spatial sign and rank functions, how will you generalize Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ for testing uncorrelatedness between two random vectors? [4]

(c) Suppose that one wants to construct a test of independence between $\mathbf{X}$ and $\mathbf{Y}$ based on a traversal of the edges of a minimum spanning tree constructed using observations on $\mathbf{X}$. Give an example to show that instead of random traversal, it is better to use traversal based on Prim's algorithm to construct the test statistic. [5]

6. Suppose that we have two sets of independent observations $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ from two d-variate normal distributions $N_d(\mathbf{0}, \mathbf{I})$ and $N_d(\boldsymbol{\mu}, 1.5\mathbf{I})$, respectively, where $\boldsymbol{\mu} = (0.7, \ldots, 0.7)^\top$.

(a) In order to test the equality of the two distributions, a person projects the observation along the direction $\bar{\mathbf{x}} - \bar{\mathbf{y}}$ and then computes the test statistic

$$T = \sup_t \left| \frac{1}{m} \sum_{i=1}^{m} I\{(\bar{\mathbf{x}} - \bar{\mathbf{y}})^\top \mathbf{x}_i \le t\} - \frac{1}{n} \sum_{i=1}^{n} I\{(\bar{\mathbf{x}} - \bar{\mathbf{y}})^\top \mathbf{y}_i \le t\} \right|.$$

The null hypothesis is rejected if the observed value of T is bigger than the threshold computed based on permutation principle. Compute the asymptotic power of this test as the dimension increases to infinity. [6]

(b) If $m/n > (1+\alpha)/(1-\alpha)$, show that the power of the multivariate run test (with nominal level α) based on the minimum spanning tree converges to 0 as d tends to infinity. [6]

(c) Assuming the prior probabilities of the two distributions to be equal, a person constructs a kernel discriminant analysis rule using these $m+n$ observations. The person uses the Gaussian kernel with a common bandwidth matrix $h^2\mathbf{I}$ for both classes. If h is of the order $O(\sqrt{d})$, find the asymptotic misclassification probability of this classifier when d grows to infinity. [6]

(d) If the average linkage method (based on Euclidean distance) is used to find two clusters in this data set, what will be the configurations of the estimated clusters when the dimension is large? Justify your answer. [6]

7. Computer Assignment [20]

INDIAN STATISTICAL INSTITUTE

Semestral Examination (2018–2019)

M Stat II

Advanced Functional Analysis

Date : 23.04.2019 Maximum Marks : 75 Time : 3 hrs.

This is an open note exam.

This paper carries 80 marks. Maximum you can score is 75.

Precisely justify all your steps. Carefully state all the results you are using.

1. For a Banach space X, show that the following are equivalent : [15]

   (a) X has the RNP.

   (b) Every closed bounded convex set in X is the norm closed convex hull of its denting points.

   (c) Every closed bounded convex set in X has a denting point.

2. Let (Ω, Σ, μ) be a probability space and X^* be a dual Banach space.

   (a) Show that $f : \Omega \to X$ is μ-measurable if and only if f is almost separably valued and for each $x \in X$, $\widehat{x} \circ f$ is measurable. [15]

   (b) If each separable subspace Y of X has Y^* separable, show that X^* has the RNP. [15]

3. (a) For a continuous convex function ϕ on a Banach space X, define the subdifferential $\partial\phi(x)$ of ϕ at some $x \in X$. [5]

   (b) If $\phi(x) = \|x\|$ for all $x \in X$, show that [10]

   $$\partial\phi(x) = \{x^* \in S_{X^*} : x^*(x) = \|x\|\}.$$

   (c) Find all points of ℓ^1 where the norm is Gâteaux differentiable. [10]

   (d) Show that the norm on ℓ^1 is not Fréchet differentiable at any point. [10]

M. Stat II year (2018-19)
Mathematical Biology 23
Date of Examination: April ~~18~~, 2019: Time: 2-5PM
Total Marks: 60

(Calculators will be allowed)

1. a) Write down Holling - Tanner model for predator-prey interaction and point out the drawback of the model.

b) Consider the following assumptions: -

(i) prey population grow with logistic fashion
(ii) predator-prey interaction follows Holling - Type I functional response
(iii) density-dependent mortality rate of predator.

Based on the above assumptions, write down a predator-prey model. Show that the solutions which initiate in R_2^+ are uniformly bounded. Prove that the system around the positive interior equilibrium point is globally asymptotically stable.

(3+8)

2. Population of the US in 1800 and 1850 was 5.3 and 23.1 million respectively. Predict its population in 1900 and 1950 using the exponential model of population growth. Then consider that the population of the US in 1900 was actually 76 million. Correct your prediction for 1950 using the logistic model of population growth (Hint: with this data, the rate of population growth =0.031476 in the logistic model). What is the carrying capacity of the US according to this model?

(5)

3.a) Stating clearly the basic assumptions, write down the general epidemic model.

b) Show that the spread of the disease will not stop due to lack of susceptible population and hence deduce Kermack and McKendrick's threshold theorem. (3+5+8)

4. State the basic deterministic model of recurrent epidemics and determine the stability of the equilibrium. Discuss the drawback of the model (7+1)

5. a) Write down the underlying assumptions of the model

$$\frac{dS}{dt} = -kSI + bS + pb'I - \gamma S + gI$$

$$\frac{dI}{dt} = kSI + qb'I - r'I + gI$$

with S(0)≥0, I(0)≥0 and $p + q = 1$

b) Find out the basic reproductive ratio of the system.

c) Find out the conditions for which the zero equilibrium will be locally asymptotically stable (LAS)

d) Show that the endemic equilibrium is LAS whenever it exists

e) Sketch a schematic diagram of the system. (2+2+2+2+2=10)

6. Pseudo-recovery is a relapse phenomenon whereby signs and symptoms of a disease are reverted after a period of improvement and it is due to incomplete treatment of the disease. Write down a simple deterministic epidemiological model with pseudo -recovery by dividing the total human population into four categories, namely, susceptible, exposed, infected and recovery population. Find out the basic reproductive ratio by using next generation matrix approach and show that the disease-free equilibrium is globally asymptotically stable by using a suitable Lyapunov function.

(2+4+4=10)

Indian Statistical Institute

Second Semestral Examination 2018-19

M. Stat. II yr

Statistical Inference III

Date: April 24, 2019 | Maximum marks: 100 | Duration: 3 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 deficiencies $\delta(\mathcal{E}, \mathcal{F}) = \delta(\mathcal{F}, \mathcal{E}) = 0$, 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 $u : \mathcal{X} \longrightarrow \mathbb{R}$.

[10+20 =30]

2. Let $X_1, X_2, \ldots, X_m$ be iid F and $Y_1, Y_2, \ldots, Y_n$ be iid G respectively, $N = m + n$. Show that for sufficiently small $\theta > 0$ the Wilcoxon test at level $\alpha = k/\binom{N}{n}$, k a positive integer, maximises the power (among rank tests) against the alternatives (F, G) with $G = (1 - \theta)F + \theta F^2$.

[15]

3. (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) Let X, Y be independently normally distributed as $\mathrm{N}(\theta, 1)$ and let $V = Y - X$. Define $W = V\,[I(X + Y > 0) - I(X + Y \leq 0)]$. Show that both V and W are ancillary, but neither is a function of the other. Further show that (V, W) is not ancillary.

[8+12=20]

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.

[20]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2018-19

MStat II year

Survival Analysis

Date: 26 April 2019 | Maximum marks: 100 | Duration: 3 hours

This examination is closed book, closed notes. Non-programmable calculators are allowed. The entire question paper is for 110 marks. The maximum you can score is 100.

1. If the mean residual life of a lifetime random variable at time t is $\frac{1}{2}e^{-t}$, what is its survival function? [8]

2. Derive a confidence interval (with asymptotic coverage probability 0.95) for the survival probability beyond time t_0, from randomly right-censored data $(T_1, \delta_1), \ldots, (T_n, \delta_n)$, where the underlying uncensored lifetimes are assumed to have the survival function $S(t) =$ ~~e^{-t^2}~~. $e^{-\lambda t^2}$ [9]

3. Assuming that nonparametric MLE (NPMLE) of the survival function for randomly right-censored survival data consists of point masses allocated to the times of observed failure, show that the NPMLE is the Kaplan-Meier estimator. [8]

4. Describe Tarone and Ware's family of two-sample tests. Show that for a suitable choice of the weights, the test reduces to Gehan's test. Give an approximate expression for the null variance of Gehan's test statistic, with justification. [3+2+4=9]

5. Let (T_{ih}, δ_{ih}), $i = 1, \ldots, n_h$, $h = 1, \ldots, k$ be survival times and censoring indicators from k samples. You have to test whether the cumulative hazards $A_1(t), \ldots, A_k(t)$ of the underlying populations are identical. Let $\hat{A}_1(t), \ldots, \hat{A}_k(t)$ be the Nelson-Aalen estimators of $A_1(t), \ldots, A_k(t)$, respectively, and $\hat{A}(t)$ be the corresponding estimator computed from the pooled sample. Suppose, for $h = 1, \ldots, k$, $\tilde{A}_h(t) = \int_0^h J_h(s) d\hat{A}(s)$, where $J_h(t) = I(Y_h(t) > 0)$, $Y_h(t) = \sum_{i=1}^{n} I(T_{ih} \geq t)$ and $I(\cdot)$ is an indicator function.

   (a) Show that the difference $\hat{A}_h(t) - \tilde{A}_h(t)$ is a martingale under appropriate assumptions.

   (b) Suppose, for $h = 1, \ldots, k$, $Z_h(t) = \int_0^t L(t) Y_h(t) d(\hat{A}_h - \tilde{A}_h)(s)$, where $L(s)$ is a predictable process. Show that $\sum_{h=1}^{k} Z_h(t) = 0$ for all t.

   (c) Show that the variance-covariance matrix of $Z_1(t), \ldots, Z_k(t)$ is singular.

   (d) Describe a consistent estimator of the variance-covariance matrix of $Z_1(t), \ldots, Z_k(t)$, with justification.

   (e) Devise a test statistic for the hypothesis of homogeneity of the samples by suitably scaling $Z_1(t), \ldots, Z_k(t)$. [5+2+1+4+3=15]

6. Assuming that covariates do not depend on time, show that the relative risk regression model and the accelerated failure time regression model can hold simultaneously if and only if the life distribution for any given covariate profile is Weibull. [10]

7. You have censored data on time to event of two groups, male and female, along with a single real-valued covariate. It is believed that for a given value of the covariate, the ratio of the hazards of the male and female groups is a linear function of time.

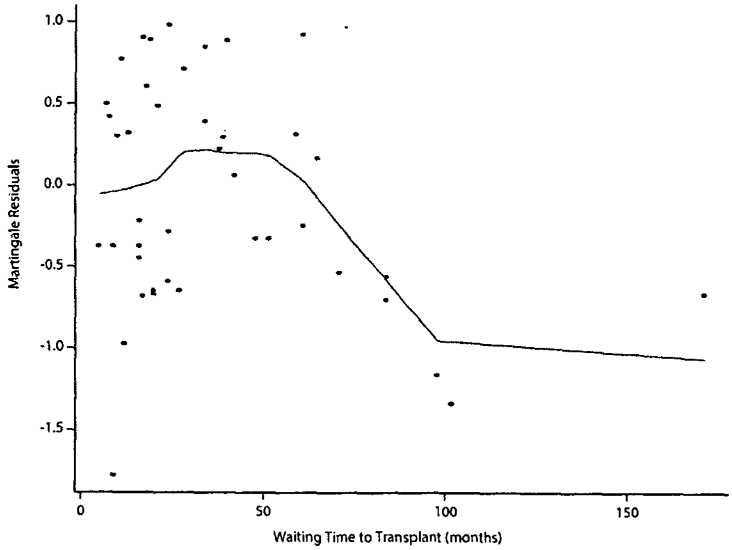
   (a) Explain how the significance of the effect of the covariate can be tested in the framework of Cox's relative risk regression model.

   (b) Using suitable notations, give an explicit expression for the partial likelihood.

   (c) Describe large sample distribution of the test statistic, with justification.

   (d) If one allows for arbitrary hazard ratio (disregarding the information of linear hazard ratio) but still believes the covariate effect would be identical in the two groups, how would the answer to part (a) change? [4+2+4+3=13]

8. This question concerns times between first treatment and either death or end of the study for a group of 90 patients diagnosed with either Hodgkins disease (HL) or non-Hodgkin's lymphoma (NHL) of the larynx, who were treated with allogeneic (allo) bone marrow transplants from an HLA-matched sibling or an autogeneic (auto) bone marrow transplant. There are several covariates, including the type of transplant (1-allo, 0-auto), disease (1-HL, 0-NHL), a disease-transplant type interaction (1 if allo and HL, 0 otherwise), the patient's initial Karnofsky score (a measure of the patient's condition at transplant) and the waiting time from diagnosis to transplant in months. A Cox model is fitted with the first four covariates, and the martingale residuals are plotted against the fifth one, along with a smooth, in the following figure.



   (a) What is a martingale residual?

   (b) What does the plot indicate?

   (c) Suggest a reasonable course of further analysis. [2+4+3=9]

9. In a particular data set of randomly right censored failure times with fixed covariates, it so happens that there are exactly two failure times tied at every observed occasion of failure. Write down the partial likelihood for this data, by breaking ties in all possible ways and averaging over the relevant terms. What would be Efron's approximate likelihood in this situation? [4+3=7]

10. A university student believes that the time taken from graduation till getting the first job (T) is a random variable with log-normal distribution, but the time also depends on fraction of daily time (x) spent on studies. He conducts a survey of his seniors to collect data on T and x.

    (a) Describe the Accelerated Failure Time (AFT) regression model for this data.

    (b) How would you interpret the hypothesis that the regression coefficient of x is zero?

    (c) Derive suitable scores for a linear rank test for testing the hypothesis of part (b), assuming that all graduates get a job, and the time of getting the first job is observed.

    (d) Explain how the regression coefficients can be estimated by using rank scores.

    (e) Explain the advantages and disadvantages of using linear regression for this problem. [2+1+4+2+2=11]

11. Consider the relative risk regression model

$$\lambda_j(t; Z) = \lambda_{0j}(t)e^{Z'\beta}, \quad j = 1, \ldots, m$$

for competing risks data $(T_i, \delta_i, j_i, Z_i)$, $i = 1, \ldots, n$, where T_i, δ_i, j_i, Z_i are the time to event, censoring indicator, failure type index and covariate vector for the ith case.

    (a) Write down the partial likelihood for this data and explain how you would estimate the regression coefficients.

    (b) Explain, using the counting process theory for the partial likelihood process, the large sample distribution of the estimators of part (a).

    (c) How would you estimate the baseline cause-specific hazards? [2+7+2=11]

Indian Statistical Institute
Semestral Examination

April 26, 2019

Weak Convergence and Empirical Processes, M2

Total points: 40 Time: 3 hours

Note: This is a **closed notes/closed book** examination. Do any **five** of the following problems. Notations, if not explicitly explained, are to be interpreted as defined in class.

1. (a) Show that $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ if and only if $\limsup_{n\to\infty} \mathbb{P}_n f \leq \mathbb{P} f$ for any upper semicontinuous f that is bounded from above.

   (b) Let ξ_i be standardized random variables defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$, $S_n := \sum_{i=1}^n \xi_i$, and $X_n(t) := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} + (nt - \lfloor nt \rfloor)\frac{\xi_{\lfloor nt \rfloor + 1}}{\sqrt{n}}$. Assume that $X_n \xrightarrow{w} B$. Show that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{\delta} \mathbb{P}\left(\max_{1\leq i\leq n} \frac{|S_i|}{\sigma\sqrt{n}} \geq \frac{1}{\sqrt{\delta}} \right) = 0.$$

[4 + 4]

2. (a) Let X be a random element of $D[0,1]$ such that

$$\lim_{\delta\to 0} \sup_{0\leq t\leq 1-\delta} \frac{1}{\delta}\mathbb{P}(|X(t+\delta) - X(t)| \geq \epsilon) = 0$$

   for any $\epsilon > 0$. Show that $\mathbb{P}(X \in C[0,1]) = 1$.

   (b) Suppose $\xi \overset{i.i.d.}{\sim} \text{Beta}(\alpha, \beta)$, $\alpha, \beta \in (0, \infty)$. Let $F_{\alpha,\beta}$ denote the CDF of a $\text{Beta}(\alpha, \beta)$ random variable. Find the limit distribution of

$$T_{\text{KS}}^{(\alpha,\beta)} := \sqrt{n} \sup_{t\in\mathbb{R}} \left|\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty,t]}(\xi_i) - F_{\alpha,\beta}(t)\right|.$$

[5 + 3]

3. (a) Show that $J_t := \{x \mid x(t) \neq x(t-)\}$ is Borel measurable (in the Skorohod topology).

   (b) Suppose that $\mathbb{P}_n \circ \pi_{t_1,\ldots,t_k}^{-1} \xrightarrow{w} \mathbb{P} \circ \pi_{t_1,\ldots,t_k}^{-1}$ for all $t_i \in T_{\mathbb{P}}, k \in \mathbb{N}$, where $T_{\mathbb{P}}$ is the set of all points $t \in [0,1]$ for which π_t is continuous $\mathbb{P}$-a.e. If $\mathbb{P}(J_1) = 0$, show that

$$\lim_{\delta\to 0} \limsup_{n\to 0} \mathbb{P}_n(w_x[1-\delta, 1) \geq \epsilon) = 0 \text{ for any } \epsilon > 0.$$

[3 + 5]

4. (a) Suppose that $x_n \in D[0,1]$ and $\|x_n - x\|_\infty \to x$ for some measurable function x. Does $x \in D[0,1]$?

(b) Let

$$x_n(t) = \begin{cases} 0 & \text{if } t \le \frac{1}{2} - \frac{1}{n}, \\ 1 & \text{if } t \ge \frac{1}{2}, \text{ and} \\ \text{linear} & \text{on } [\frac{1}{2} - \frac{1}{n}, \frac{1}{2}]. \end{cases}$$

Does the sequence $\{x_n\}$ converge in $(D[0,1], d^\circ)$? If yes, what is the limit function?

[3 + 5]

5. (a) Define the concept of VC-dimension. Why are Boolean function classes of finite VC-dimension useful in empirical process theory?

(b) Let $\mathcal{F}_{\text{ball}} = \{B[\mu; r] \mid \mu \in \mathbb{R}^2, r \ge 0\}$ be the class of all 2-dimensional closed Euclidean balls. Compute the VC-dimension of $\mathcal{F}_{\text{ball}}$.

[(2 + 1) + 5]

6. Let $\xi_i \overset{i.i.d.}{\sim} \xi, i = 1, \ldots, n$. Ignoring measurability issues, show that, for any convex nondecreasing function Φ and any class of measurable functions $\mathcal{F}$, one has

$$\mathbb{E}_{\xi,\epsilon}\Phi(\frac{1}{2}\sup_{f\in\bar{\mathcal{F}}}|\frac{1}{n}\sum_{i=1}^{n}\epsilon_i f(\xi_i)|) \le \mathbb{E}\Phi(\|\mathbb{P}_n - \mathbb{P}\|_{\mathcal{F}}) \le \mathbb{E}_{\xi,\epsilon}\Phi(2\sup_{f\in\mathcal{F}}|\frac{1}{n}\sum_{i=1}^{n}\epsilon_i f(\xi_i)|),$$

where $\bar{\mathcal{F}} = \{f - \mathbb{P}f \mid f \in \mathcal{F}\}$ and $(\epsilon_i)_{1\le i\le n}$ is a vector of i.i.d. Rademacher variables, independent of the ξ_i's.

[(4 + 4)]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2018 – 19

MStat (2nd Year)

Financial Econometrics

Date: 26 April 2019 Maximum Marks: 100 Duration: 3 Hours

Be brief and to the point, use examples whenever applicable.
You may use standard textbook notation automatically.

1. What is the Efficient Market Hypothesis? What are the different versions of this? Are these testable – answer with a specific example of model.

[3 + 3 + 5 = 11]

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

(b) How is the variance ratio (VR) used in testing for Random Walk? Why do the weights decline linearly in

$$VR(q) = \frac{Var[r_t(q)]}{q\,Var[r_t]} = 1 + 2\sum_{k=1}^{q-1}(1 - \tfrac{k}{q})\rho(k)?$$

[(3 + 4) + (3 + 3) = 13]

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

Under the non-trading process defined by (assuming independence)

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

$$X_{it}(k) \equiv (1-\delta_{it})\delta_{it-1}\delta_{it-2}\dots\delta_{it-k}, \quad k > 0$$
$$= \begin{cases} 1 & with\ probability\ (1-\pi_i)\pi_i^k \\ 0 & with\ probability\ 1-(1-\pi_i)\pi_i^k \end{cases}$$

and assuming that virtual returns have a linear one factor structure

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

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

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

[(2 + 4) + (3 + 4) = 13]

4. (a) For measuring normal performance of the market, explain the role of the following models:

   (i) Constant mean return model

   (ii) Market model

   (b) What is the exact factor pricing model? What are the alternative versions of this model?

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

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

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

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

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

   will satisfy (i) $\beta_0 = 0$ and (ii) $\beta_p + \beta_{op} = 1$. [4 + 6 = 10]

6. (a) What is Brownian Motion?

   (b) Explain the Maximum likelihood method of parameter estimation for an option pricing model – use a specific example. [4 + 5 = 9]

7. You are told that an 8-year nominal zero-coupon bond has a log yield to maturity of 9.1%, and a 9-year nominal zero-coupon bond has a log yield of 8.0%.

(a) Can the pure expectations theory of the term structure describe these data?

(b) A year goes by, and the bonds in part (a) still have the same yields to maturity. Can the pure expectations theory of the term structure describe these new data?

(c) How would your answers change if you were told that the bonds have an 8% coupon rate per year, rather than zero coupons? [3+ 4 + 4 = 11]

8. Assume that the homoscedastic lognormal bond pricing model given by

$$x_{t+1} = (1 - \phi)\mu + \phi x_t + \xi_{t+1}$$

and

$$-m_{t+1} = x_t + \beta\xi_{t+1}$$

holds with $\phi < 1$. ξ_{t+i} is a normally distributed shock with constant variance.

(a) Suppose you fit the current term structure of interest rates using a random walk model augmented by deterministic drift terms,

$$x_{t+i} = x_{t+i-1} + g_{t+i} + \xi_{t+i},$$

where g_{t+i} is a deterministic drift term that is specified at time t for all future dates $(t + i)$ in order to fit the time t term structure of interest rates.

Derive an expression relating the drift terms to the state variable x_t and the parameters of the true bond pricing model.

(b) Compare the expected future log short rates implied by the true bond pricing model and the random walk model with deterministic drifts. [5 + 5 = 10]

9. Consider the GARCH (1, 1) model where the volatility at t is governed by $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\eta_t^2$ where $\eta_{t+1} \sim N(0, \sigma_t^2)$

What is the condition for weak stationarity of η_t ? Find the conditional and unconditional variance of η_{t+1} when stationary.

[3 + 4 + 4 = 11]

INDIAN STATISTICAL INSTITUTE

M.Stat. II Year

End-Semester Examination : Semester II : 2018-2019

BROWNIAN MOTION AND DIFFUSIONS

Date : ~~03.05.2019~~ 29/4/19 | **Maximum Score : 60** | Time : **3 Hours**

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

1. (a) State Reflection Principle for a Standard Brownian Motion.
   (b) Define what is meant by the Resolvent Operator of a Markov Process.
   (c) For a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t, t \geq 0), P)$, define the class $\mathcal{L}^2_{\text{loc}}$.
   (d) State Ito's Formula for $\varphi\left(\int_0^t f_s dB_s, \int_0^t g_s ds\right)$, stating the assumptions on φ, f and g.

   (4×2)=[**8**]

2. Let $\{B_t, t \geq 0\}$ be a SBM on $(\Omega, \mathcal{A}, P)$ and $\{\mathcal{A}_t, t \geq 0\}$ be its natural filtration.
   (a) Prove that, for any $s \geq 0$ and $t > 0$, a regular conditional distribution of $|B_{s+t}|$, given $\mathcal{A}_s$, is given by $p(t, |B_s|, \cdot)$, where $p(t, x, A) = P[|N(x,t)| \in A]$.
   (b) Deduce that $\{|B_t|, t \geq 0\}$ has the Markov property.

   (8+7)=[**15**]

3. Let $\{B_t, t \geq 0\}$ be a SBM and let $M_t = \max\{B_s : 0 \leq s \leq t\}$ for $t > 0$.
   (a) For $x \geq 0$ and $y \leq x$, find the probability $P[M_t \geq x, B_t < y]$.
   (b) Using (a) or otherwise, show that, for $x > 0$, the random variable T_x, the hitting time of x, has an absolutely continuous distribution.
   (c) Show that $\{T_x, x \geq 0\}$ is a stochastic process with almost surely right-continuous paths and stationary, independent increments.

   (4+4+7)=[**15**]

4. (a) Assuming usual notations for a Markov Process, show that, for all $\alpha > 0, t > 0$, $R_\alpha T_t = T_t R_\alpha$ and hence show that $g \in \mathcal{D}_A \Rightarrow T_t g \in \mathcal{D}_A$ and $AT_t g = T_t A g$.
   (b) Representing Poisson process with intensity $\lambda > 0$ as a Markov process with the set of non-negative integers as the state space S, find the generator $(A, \mathcal{D}_A)$.

   (8+7)=[**15**]

5. Let $(\Omega, \mathcal{A}, P)$ be a probability space with a right-continuous filtration $(\mathcal{A}_t, t \geq 0)$ and let $\{B_t, t \geq 0\}$ be a $\{\mathcal{A}_t, t \geq 0\}$-Brownian motion. Assume also that $\mathcal{A}$ is complete and that $\mathcal{A}_0$ contains all P-null sets in $\mathcal{A}$.
   (a) For each $n \geq 1$ and $s \in [0, \infty), \omega \in \Omega$, define $f_n(s, \omega) = \sum_{i \geq 0} B(\frac{i}{n}, \omega)\ \mathrm{I}_{(\frac{i}{n}, \frac{i+1}{n}]}(s)$. Show that $f_n(\cdot,\cdot) \longrightarrow B(\cdot,\cdot)$ in $\mathcal{L}^2$ and hence deduce that $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$, $t \geq 0$.
   (b) Let $f \in \mathcal{L}^2_{\text{loc}}$ and let τ be a stopping time. Define f^τ as $f^\tau(s, \omega) = f(s, \omega)\mathrm{I}_{s \leq \tau(\omega)}$. Show that $f^\tau \in \mathcal{L}^2_{\text{loc}}$ and that $\int_0^{t \wedge \tau} f_s dB_s = \int_0^t f_s^\tau dB_s$, $t \geq 0$.

   (8+7)=[**15**]

INDIAN STATISTICAL INSTITUTE

Semestral Examination (2nd Semester): 2018-19

M. STAT. 2nd year

Subject Name: Clinical Trials

Date: 29/4/2019, Maximum Marks: 50. Duration: 2 hrs. 30 min.

Group A: Answer all questions.

1. (a) How a randomized play-the-winner rule can be adjusted in the presence of delayed response? [4]

   (b) Obtain the transition probability matrix of the delayed response indicator where the response time follows an exponential distribution with mean a^{-1}, and the interarrival time follows another exponential distribution with mean b^{-1}, independently of each other. Show that the entries of the transition probability matrix are functions of $b/(a+b)$. [8]

2. (a) Discuss how optimal safe dose can be obtained in a phase II clinical trial set up. What is the role of a bivariate binary distribution for this purpose? [7+3]

   (b) Discuss and compare cumulative cohort design and the continual reassessment method in the context of phase I clinical trial set up. [4+4+2]

3. Suppose we have a two-treatment set up with binary responses where p_A and p_B be the success probabilities. The prior distributions of p_A and p_B are same and independent of each other. The prior for p_k is: For $0 < a < b < 1$, p_k can take the values a or b with probabilities 1/4 and 3/4; $k = A, B$. Find a suitable Bayesian response-adaptive design for allocating the $(n+1)$st patient based on the data from the first n patients. [8]

4. We need to analyze the data from a clinical trial to compare three alternative dose regimens of haloperidol for schizophrenia patients. Sixty-five patients with diagnosis of schizophrenia were randomly assigned to receive 5, 10, or 20 mg/day of haloperidol for 4 weeks. After randomization, we had 22, 25 and 18 participants for three dose groups respectively. 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).

   Twenty-nine patients dropped out of the study at $j = 3$, with 10, 10 and 9 dropouts from three dose groups respectively. 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.

   Develop a suitable selection model for the above dataset. Considering appropriate prior distributions for the model parameters, develop a Bayesian estimation method for the above problem. [10]

INDIAN STATISTICAL INSTIUTE

M. Stat (2nd year) 2018-19

Subject: Theory of Games and Decisions

Date: 30.04.2019 Full Marks: 60 Duration: 2 hrs 30 mins

Attempt all questions

1. a) State Nash's Theorem on finite non-cooperative games.

   b) Give example of a 3-player game with number of strategies 3, 4, 5 respectively, such that the game has more than one equilibrium situations each with different pay-off for the same player.

   Show that rectangular property for the set of equilibrium situations does not hold in the above context.

   [2+5]

2. Let A be a 2x2 real matrix.
   Consider mixed extension of matrix game A.
   Discuss graphical solution of the above game with optimal strategies and values, considering different cases (7 cases).

   [12]

3. a) Define co-operative Games.

   b) Let v be a characteristic function of a co-operative game deduced from a constant-sum non-cooperative game. Show that the following holds:

   $v(K \cup L) \geq v(K) + v(L) \quad \forall K, L$ such that $K \cap L = \Phi$

   c) Define inessential co-operative games.
   Show that for those games, the characteristic function satisfies

   $v(K) + v(L) = v(K \cup L), \quad \forall K, L$ such that $K \cap L = \Phi$

   d) Show that the 3-player, constant –sum, essential, co-operative game is unique in 0-1 reduced form.

   [1+11+6+6]

4. Let A be an mxn real matrix.
   Consider the matrix game A (not mixed extension).
   Show that the game has a saddle point iff all the 2x2 submatrices of it have a saddle point.
   [Consider all entries of A are different.]

   [8]

5. a) Find the value of the mixed extension of the matrix game I_{nxn}.

   b) For each m,n in N (m,n>1) give example of a matrix A_{mxn}, such that the mixed extension of matrix game A , does not have equilibrium situation in pure strategies.

   [4+5]

-----------------xxxxx--------------------

INDIAN STATISTICAL INSTITUTE, KOLKATA
FINAL EXAMINATION: SECOND SEMESTER 2018 -'19
M. STAT II YEAR

03.05.2019

Subject: **Representation theory of finite groups** Date: **April ??, 2019**
Duration: **3 hours** Time: **??:?? ?M to ??:?? ?M**
Maximum score: **60**

Attempt all the problems. Justify every step in order to get full credit of your answers. All arguments should be clearly mentioned on the answerscript. Points will be deducted for missing or incomplete arguments.

Standing assumption: Cardinality of all groups and dimension of all vector spaces are assumed to be finite unless mentioned otherwise. For a finite group G and a subring R of $\mathbb{C}$,
(i) $\mathrm{Fus}_R(G)$ denotes the ring of R-linear combination of isomorphism classes of irreducible representations of G with multiplication given by the tensor product, and
(ii) $R\mathrm{Char}(G)$ is the set of all R-linear combination of elements in $\mathrm{Char}(G)$ (= the set of characters of G).

(1) If H is a subgroup of G, then $\mathrm{Ind}_H^G(\mathbb{K}\mathrm{Char}(H))$ is an ideal in $\mathbb{K}\mathrm{Char}(G)$ for any $\mathbb{K} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

[5 marks]

(2) For a cyclic group H, let $\alpha_H : H \to \mathbb{C}$ denote the function $\alpha_H(a) = \delta_{\langle a\rangle = H}\ |H|$. Show that $\sum_{H\in CS(G)} \mathrm{Ind}_H^G(\alpha_H) = |G|$ where $CS(G)$ denotes the set of all cyclic subgroups of G.

[8 marks]

(3) Let G be a finite group and α be a class function of G.
(a) If R is a subring of $\mathbb{C}$ such that $\mathbb{Q} \subset R$ and $\mathrm{Res}_H^G(\alpha) \in R\mathrm{Char}(H)$ for every cyclic subgroup H of G, then $\alpha \in R\mathrm{Char}(G)$.
(b) If R is a subring of $\mathbb{C}$ such that $\mathbb{Z} \subset R$ and $\mathrm{Res}_H^G(\alpha) \in R\mathrm{Char}(H)$ for every elementary subgroup H of G, then $\alpha \in R\mathrm{Char}(G)$.

[6+6 marks]

(4) Show that H is normal in G if and only if $\mathrm{Ind}_H^G(1)$ is zero outside H.

[5 marks]

(5) For a conjugacy class C of G, let δ_C be the class function given by $\delta_C(g) := \delta_{g\in C}$. If Φ denotes the set of all irreducible characters of G, then show that

$$\delta_C = \sum_{\varphi\in\Phi} \frac{\overline{\varphi}(g)}{|\{a \in G : ag = ga\}|}\, \varphi \quad \text{for any } g \in C .$$

[10 marks]

(6) Let $\rho : G \to \mathrm{Aut}(V)$ be a representation of the group G.

(a) If the function $F_\wedge$ from G to power series with complex coefficient is given by

$$F_\wedge(g)(t) := \sum_{k \geq 0} \chi_{\wedge^k V}(g)\ t^k\ ,$$

then show that $F_\wedge(g)(t) = \det(1 + \rho(g)t)$.

(b) If the function F_{Sym} from G to power series with complex coefficient is given by

$$F_{\mathsf{Sym}}(g)(t) := \sum_{k \geq 0} \chi_{\mathrm{Sym}^k V}(g)\ t^k\ ,$$

then show that $F_{\mathsf{Sym}}(g)(t) = \dfrac{1}{\det(1 - \rho(g)t)}$.

[10+15 marks]

INDIAN STATISTICAL INSTITUTE

ROBUST STATISTICS

M-Stat II, Second Semester 2018-2019

Semester Examination

Date: 02.05.2019 Duration: 3 hours

[Answer as many as you can. Total points 57. Maximum you can score is 50.]

1. Define qualitative robustness. Show that the sample mean is not qualitatively robust. [2+7]

2. (a) Define a statistical functional $T(F)$, and its von-Mises derivative.

   (b) Define the influence function. Discuss its dual role in indicating the robustness and the efficiency of an estimator. [3+6]

3. Derive the influence function of the 10% trimmed mean. [8]

4. Suppose that we have a random sample $X_1, \ldots, X_n$ from the true unknown distribution G which is modeled by a discrete parametric family $\{F_\theta\}$, and without loss of generality let $\{0, 1, 2, 3, \ldots,\}$ be the common support of G and $\{F_\theta\}$. We are interested in estimating the unknown θ using an appropriate disparity. Let d_n represent the observed relative frequencies obtained from the data. Let $\{\xi_j : j = 1, 2, \ldots\}$ be a sequence of elements of the sample space and let $\epsilon \in (0, 1)$ be the contaminating proportion. Consider the contaminated data density

$$d_j(x) = (1 - \epsilon)d_n(x) + \epsilon\chi_{\xi_j}(x),$$

where $\chi_y(x)$ is the indicator function at y.

   (a) When do we say that $\{\xi_j\}$ constitutes an outlier sequence for the model density f_θ and the data density d_n?

   (b) Let $d_\epsilon^*(x) = (1 - \epsilon)d_n(x)$. Show that under suitable conditions (to be stated by you) on the $C(\cdot)$ function of the disparity,

$$\rho_C(d_j, f_\theta) \to \rho_C(d_\epsilon^*, f_\theta) \text{ as } j \to \infty$$

   for any outlier sequence $\{\xi_j\}$.

   (c) What is the effect of the convergence in part (b) in case of the estimator and the test based on the Hellinger distance? [1+7+2]

5. (a) Consider the parametric model $\{F_\theta\}$, and suppose we have i.i.d. data from the true distribution G (which is modeled by F_θ). We are interested in estimating the parameter θ. Let G_n represent the empirical distribution based on the data. Suppose

   (i) θ_0 is an isolated root of $\int \psi(x,\theta)dG(x) = \lambda_G(\theta) = 0$.

   (ii) $\psi(x,\theta)$ is monotone and continuous in θ at $\theta = \theta_0$.

   (iii) $\int \psi^2(x,\theta)dG(x) < \infty$ for θ in a neighborhood of θ_0.

   (iv) $\lambda_G(\theta)$ is differentiable at $\theta = \theta_0$, with $\lambda'_G(\theta_0) \neq 0$.

   Then θ_0 is unique, and any solution sequence $\{T_n\}$ of $\lambda_{G_n}(\theta) = 0$ converges to θ_0 with probability 1. Further,

$$n^{1/2}(T_n - \theta_0) \to Z\ N(0, \sigma^2(T,G)),$$

   with $\sigma^2(T,G) = \int \psi^2(x,\theta_0)dG(x)/[\lambda'_G(\theta_0)]^2 = 0$. [12]

6. Consider the problem of testing parametric hypothesis using the deviance test based on the Hellinger distance (rather than the likelihood ratio test).

   (a) Derive the power breakdown bound of the deviance test based on the Hellinger distance.

   (b) Consider the problem of testing the hypothesis $H_0 : \lambda = 1$ under data which are assumed to be generated by the Poisson(λ) model. Compute the power breakdown bounds for the deviance test based on the Hellinger distance when the true distributions are (i) Poisson with mean 4 and (ii) Poisson with mean 8. [5+4]

MSTAT II - Computational Finance

Final Exam. / Semester II 2018-19

Date - April 30, 2019 / Time - 4 hours

Maximum Score - 50

<u>NOTE</u> : **SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.**

1. (4+8+8=20) Consider the followig Black-Schole-Merton PDE $\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} = rf$ where f is price of an American Put Option.

   (a) Transform the B-S-M equation using $Z = \log S$. Is the trinomial probability model corresponding to the explicit difference method stable for some values of the parameters? Justify your answer showing your calculations.

   (b) Find the price of the American Put option numerically, dividing (3 months maturity) time into 30 parts and the space unit into 40 parts. Here you can assume, $r = 6\%$ per annum, $S_0 = 60$, $K = 60$, $S_{\max} = 160$ and choose any two values of σ for which the trinomial probability model is stable and find the price for each and compare.

   (c) Use Monte Carlo method to simulate the path of asset for the data given in part (b) above, and find the American Put option price and while reducing its variation using (i) antithetic method, and (ii) control variate method. Also provide the corresponding variances.

2. (4+8=12) Consider the CIR model for the interest rate, $dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$ where k, σ and θ are positive constants and $\{W_t\}$ is a standard Brownian motion.

   (a) Can you give a condition on k, θ and σ for which r_t remains positive whenever $r_0 > 0$? Justify your answer in terms of financial reasoning.

   (b) Let P_s^T be the price of a zero coupon bond at time s that matures at time $T(> s)$ and given by $E[\exp\{-\int_s^T r_t dt\} \mid \mathcal{F}_s]$ (under risk-neutral measure), where $\mathcal{F}_s = \sigma(\{W_u : 0 \leq u \leq s\} \cup \{r_0\})$. Evaluate P_s^T, using Monte Carlo method simulating the corresponding path for the given data, while reducing its variation using (i) antithetic method, and (ii) control variate method. Also provide the corresponding variances.

   [Assume: $k = 2$, $\theta = 0.04$, $T = 6$ months, $\sigma = 0.3$, $s = 1$ month and $r_1 = 0.03$.]

3. (2+8=10) Consider the following Black-Schole-Merton PDE $\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} = rf$ for the evaluation of the price of the up-and-out European Call option with the barrier H.

   (a) Give the boundary and maturity conditions in this case.

(b) Find the price of the option numerically using implicit or explicit method, dividing (4 months maturity) time and the space units appropriately near and away from the barrier. Here you can assume, $r = 4\%$ per annum, $\sigma = 0.25$, $K = 50$, $H = 70$ and do the problem for various values of initial price $S_0 = 44, 48, 52, 56$.

4. (3+5+8=16)

(a) Give an example of a random number generated by Mersenne Twister.

(b) Let $R_t \sim \epsilon_t \sigma_t$ and $\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \epsilon_{t-1} + \gamma_1(|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)) + \beta_1 \ln(\sigma_t^2)$, where $\{\epsilon_t\}$ are i.i.d. $N(0,1)$. Would this model for R_t be showing clustering effect or correlation with σ_t^2? Justify your answer by calculating kurtosis of R_t and $Cov(R_t, \sigma_t^2)$, assuming stationarity.

(c) Define VaR and Expected Shortfall for a given price data set. Find the VaR and Expected Shortfall for the minimum variance portfolio of the given data set of asset prices (to be supplied during the exam). Give an estimate of variability of your estimate.

All the best.

Indian Statistical Institute

M.Stat. Second year, Second Semester Examination: 2018-19

Topic: Statistical Computing II

Maximum Marks: 50, Duration: 3 hours DATE - 07.05. 2019

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

   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). [4+6=10]

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

   (b) Define the term "exchangeable random variables". State De Finetti's theorem of exchangeability for an infinite sequence of random variables. [5+5=10]

3. (a) Let a random probability measure G on $(\Omega, \mathcal{B})$ be assigned a Dirichlet Process (DP) prior, i.e. we assume that $G \sim DP(\alpha, G_0)$, where α is the concentration parameter, and G_0 is the base measure on $(\Omega, \mathcal{B})$. For all $B \in \mathcal{B}$, show that
   (i) $E(G(B)) = G_0(B)$; and (ii) Variance$(G(B)) = \frac{G_0(B)(1-G_0(B))}{1+\alpha}$.

   (b) Consider the following Dirichlet Process Mixture model for n observations: $X_i|\theta_i \sim N(\theta_i, 10)$; $\theta_i|G \sim G$; $G \sim DP(\alpha, G_0)$. The goal is to sample from the joint posterior distribution $P(\theta_1, \theta_2, \ldots, \theta_n | X_1, X_2, \ldots, X_n)$. Considering truncation approximation to Dirichlet Process, develop a Blocked Gibbs sampler for sampling from the joint posterior distribution. [4+8]

4. Consider the following regression model:
   $Y_i = f(X_i) + e_i$; for $i = 1, 2, \ldots, n$; where e_is are iid $N(0, 1)$. Assuming a zero-mean Gaussian Process prior (with squared exponential covariance kernel) for the random function $f(.)$ how can you predict the response Y_{n+1} for the $(n+1)$-th subject given the relevant covariate X_{n+1}? Write down the steps and the algorithm clearly. [8]

5. Consider a cluster-based network with 5 clusters; and n_i denotes the size of the i-th cluster. For each node of each cluster, binary state values (1/0) are recorded at 10 different discrete time points. Our goal is to develop a semi-parametric approach for modeling the network so that information on model parameters is shared across the clusters.

   Write down an appropriate linear statistical model, and the prior distributions on model parameters. State the Markov Chain Monte Carlo (MCMC) computation steps for estimating the model parameters. [10]

Date: 10.7.2019

INDIAN STATISTICAL INSTITUTE

M.STAT SECOND YEAR

Second Semester, 2018-19

Inference for High Dimensional Data

Time: $3\frac{1}{2}$ Hours | Semestral (Backpaper) Examination | Full Marks: 100

1. (a) Give an example to show that Sidak's multiple testing procedure may fail to control the family-wise error rate at the desired level when the test statistics corresponding to different tests are not independent. [8]

   (b) Give an example of a multiple testing procedure which provides a weak control on family-wise error rate but fails to provide a strong control on it. [6]

   (c) Suppose that $X_1, X_2, \ldots, X_n$ jointly follow a multivariate normal distribution with the dispersion matrix $\mathbf{\Sigma} = ((\sigma_{ij}))$. Show that they are positively regression dependent if and only if $\sigma_{ij} \geq 0$ for all $i \neq j$. [6]

2. (a) Show that in a multiple linear regression problem, the LASSO estimate of regression coefficients can be viewed as the mode of a posterior distribution for a suitable choice of the prior distribution. [4]

   (b) Write down the soft thresholding operator used in LASSO. Give an example of a multiple linear regression problem, where the LASSO solution can be obtained by using this operator on the least squares estimates. Give justification for your answer. [2+2+4]

   (c) Why do we need group LASSO algorithm? Write down the objective function used in grouped LASSO. [2+2]

   (d) Describe how coordinate descent method can be used to estimate the model parameters when a fused LASSO algorithm is used. [4]

3. Suppose that we have two sets of independent observations $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ from two d-variate normal distributions $N_d(\mathbf{0}, \sigma_1^2\mathbf{I})$ and $N_d(\boldsymbol{\mu}, \sigma_2^2\mathbf{I})$, respectively, where $\boldsymbol{\mu} = (\nu, \ldots, \nu)^\top$.

   (a) Find the high dimensional behavior of the multivariate run statistic based on the minimum spanning tree and that based on the shortest Hamiltonian path when (i) $\nu^2 > |\sigma_1^2 - \sigma_2^2|$ (ii) $\nu^2 < |\sigma_1^2 - \sigma_2^2|$. [5+5]

   (b) Consider the k-nearest neighbor test with $k < \min\{m, n\}$. If $n/(m-1) < (1-\alpha)/2$ and $\nu^2 < \sigma_1^2 - \sigma_2^2$, show that the power of this test converges to 0 as d tends to infinity. How will you modify this test to take care of this problem? [6+4]

INDIAN STATISTICAL INSTITUTE

M.Stat. II Year
Backpaper Examination : Semester II : 2018-2019
BROWNIAN MOTION AND DIFFUSIONS

Date : ~~12.07.2017~~ 10.09.2019 | **Maximum Score : 45** | Time : 3 Hours

Note : This paper carries questions worth a **total** of **100 marks**. Answer as much as you can. The **maximum** you can score is **45**.

1. (a) State clearly Kolmogorov Consistency Theorem and Kolmogorov Continuity Criterion.

   (b) Show that there is a zero-mean Gaussian process $\{X_t, t \geq 0\}$ with continuous paths and $\text{Cov}(X_s, X_t) = \exp(-|s-t|)$.

   (c) Show that it is not possible to have a process $\{X_t, t \geq 0\}$ with continuous paths where the X_t, $t \geq 0$, are i.i.d. $N(0,1)$ random variables. $[(2+2)+8+8] = [\mathbf{20}]$

2. (a) Let $\{B_t, t \in [0,1]\}$ be a SBM. Show that $X = \int_0^1 B_s^2\, ds$ is a random variable and find its mean and variance.

   (b) Let $\{B_1(t)\}, \ldots, \{B_k(t)\}$ be independent standard Brownian motions. Show that for every $\mathbf{x} \in \mathbb{R}^k$ with $\|\mathbf{x}\| = 1$, $\{B(t) = \langle \mathbf{x}, (B_1(t), \ldots, B_k(t)) \rangle\}_{t \geq 0}$ is a standard Brownian motion. Here, $\|\cdot\|$ denotes the usual euclidean norm on $\mathbb{R}^k$.

   (c) Let $\{B_t, t \in [0,\infty)\}$ $\{\beta_t, t \in [0,\infty)\}$ be two SBMs independent of each other. Denoting $\tau_a = \inf\{t \geq 0 : B_t \geq a\}$ for $a > 0$, find the distribution of the random variable β_{τ_a}.

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

   (i) Denote $\alpha = \frac{m+1}{2m-2}$, where $m \geq 2$ is an integer. For integers $j \geq 1, k \geq 1$, denoting $D_{j,k}$ to be the event $\bigcap_{n>(m+1)k} \bigcup_{1 \leq i < n-m} \bigcap_{l=1}^{m} \{\omega : |B(\frac{i+l}{n}, \omega) - B(\frac{i+l-1}{n}, \omega)| \leq j[(l+1)^\alpha + l^\alpha]/n^\alpha\}$, show that $P(D_{j,k}) = 0$.

   (ii) Deduce that, there is a P-null set N such that, if $\omega \notin N$, then for all $t \geq 0$ and for all $\alpha > \frac{1}{2}$, the set $\left\{\frac{B_s(\omega) - B_t(\omega)}{|s-t|^\alpha},\ s \neq t\right\}$ remains unbounded, as $s \to t$. $[4+4+4+(6+6)] = [\mathbf{24}]$

3. Let C denote the set of all real continuous functions on $[0,\infty)$ and $\{X_t,\ t \geq 0\}$ denote the co-ordinate process on C. Denoting $\{\mathcal{C}_t\}_{t \geq 0}$ to be the natural filtration of the process $\{X_t\}$ and $\mathcal{C} = \vee\{\mathcal{C}_t, t \geq 0\}$, show the following:

   (a) τ is a $\{\mathcal{C}_t\}$-stopping time if and only if for every $\omega, \omega' \in \mathrm{C}$, $\tau(\omega) \leq t$ and $X_s(\omega) = X_s(\omega')$ for $s \leq t$ imply $\tau(\omega) = \tau(\omega')$.

   (b) For any open set $G \subset \mathbb{R}$, the random variable $\tau(\omega) = \inf\{t \geq 0 : X_t(\omega) \in G\}$ is a $\{\mathcal{C}_t\}$-optional time, but need not be a $\{\mathcal{C}_t\}$-stopping time.

   (c) A real random variable Z on $(\mathrm{C}, \mathcal{C})$ is $\mathcal{C}_\tau$-measurable if and only if for every $\omega \in \mathrm{C}$, $Z(\omega) = Z(\omega^\tau)$, where $\omega^\tau(t) = \omega(t \wedge \tau(\omega))$. $[8+6+8] = [\mathbf{22}]$

4. Let $S = [0,\infty)$ and, for each $x \in S$, denote P_x to be the distribution of the process $\{X_t = |x + B_t|,\ t \geq 0\}$, where $\{B_t, t \in [0,\infty)\}$ be a SBM.

   (a) Show that the family $\{P_x,\ x \in S\}$ of probabilities on the appropriate path space is a Markov process with state space $S = [0,\infty)$ and find its transition probabilities.

   (b) Show that the above Markov process has the Feller property.

   (c) Taking $C_b(S)$ as the underlying Banach space, find the generator of the above Markov process. [May use: $\int_0^\infty e^{-(\alpha u - \beta/u)^2} du = \frac{\sqrt{2\pi}}{2\alpha}$, for any $\alpha > 0, \beta > 0$.] $[8+4+8] = [\mathbf{20}]$

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

   (a) Show that, for any $f \in \mathcal{L}^2$, $M_t = \left[\int_0^t f_s dB_s\right]^2 - \int_0^t f_s^2 ds$, $t \geq 0$ is a martingale.

   (b) For a $C_{2,1}$ function φ, state and prove Ito's formula for $\varphi(B_t, t)$. $[6+8] = [\mathbf{14}]$

INDIAN STATISTICAL INSTITUTE

M.STAT. (2nd Year), 2018-2019

Back Paper Examination

Subject: Theory of Games and Decisions

Date:11.07.2019 F.M.-100 Duration: 3hrs

Attempt All Questions

1. State and prove Nash's Theorem on finite non-cooperative Games. [20]

2. Let A be an $m \times 2$ matrix, consider mixed extension of matrix Game A, Discuss how to solve the above game graphically, highlighting different cases with values and optimal strategies. [20]

3. Let $\{I, v_1\}$ and $\{I, v_2\}$ denote two co-operative games on the same set of players. When are they called strategically equivalent? Show that all essential co-operative games are strategically equivalent to some unique co-operative game in 0-1 reduced form. Show that all inessential games are strategically equivalent to zero games. [20]

4. Let $A_{m \times n}$ be a real matrix. Consider mixed extension to the matrix game A. Prove the following-

   a) $MAX_X \, Min_Y \, XAY^T$ and $Min_Y \, Max_X \, XAY^T$ exist.

   b) $Max_X \, Min_Y \, XAY^T = Min_Y \, Max_X \, XAY^T$. [10+15]

5. Let $A_{m \times n}$ be a diagonal matrix with diagonal elements $d_1, d_2, d_3, \ldots, d_n$. Consider mixed extension to the matrix game A. Considering different cases, derive the value of the game and optimal strategies of the players. [15]

****

INDIAN STATISTICAL INSTITUTE

Second Semester Back`paper Examination: 2018 – 19

MStat (2nd Year)

Financial Econometrics

Date: 15 July 2019 Maximum Marks: 100 Duration: 3 Hours

Attempt ALL questions

You may use all standard notation automatically

1. How are the Random Walk Hypotheses RW1, RW2 and RW3 related? Use a Venn diagram for your answer and provide specific examples. [15]

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

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

Derive the unconditional mean, variance and first order auto covariance of δ_{it} as functions of π_i and π_i'. [4 + 5 + 7 = 16]

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. [5 + 8 + 7 = 20]

4. What is the exact factor pricing model? What are the alternative versions of this model? How are they tested in practice? [4 + 4 + 8 = 16]

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

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

[6 + 11 = 17]

INDIAN STATISTICAL INSTITUTE

Backpaper Examination (2018–2019)

M Stat II

Advanced Functional Analysis

Date : 15.07.2019 Maximum Marks : 100 Time : 3 hrs.

Precisely justify all your steps. Carefully state all the results you are using.

1. (a) Show that the space c_0 is not isometric to the dual of any Banach space. [10]

   (b) Show that the space c_0 is not even isomorphic to the dual of any Banach space. [10]

2. (a) Let X be a Banach space. Define a denting point of a closed bounded convex set $K \subseteq X$. [5]

   (b) Show that a denting point is always an extreme point. [10]

   (c) Show that the converse holds if K is compact. [10]

3. For a Banach space X, show that the following are equivalent : [15]

   (a) X has the RNP.

   (b) Every closed bounded convex set in X is the norm closed convex hull of its denting points.

   (c) Every closed bounded convex set in X has a denting point.

4. (a) For a continuous convex function ϕ on a Banach space X, define the subdifferential $\partial\phi(x)$ of ϕ at some $x \in X$. [5]

   (b) If $\phi(x) = \|x\|$ for all $x \in X$, show that [10]

$$\partial\phi(x) = \{x^* \in S_{X^*} : x^*(x) = \|x\|\}.$$

   (c) Find all points of ℓ^1 where the norm is Gâteaux differentiable. [10]

   (d) Show that the norm on ℓ^1 is not Fréchet differentiable at any point. [15]

Indian Statistical Institute

Backpaper Examination

July 16, 2019

Weak Convergence and Empirical Processes, M2

Total points: 100 | Time: 3 hours

Note: This is a **closed notes/closed book** examination. Notations, if not explicitly explained, are to be interpreted as defined in class.

1. Suppose $(\mathcal{M}_1, \rho_1), (\mathcal{M}_2, \rho_2)$ are two metric spaces and let $\mathcal{M} = (\mathcal{M}_1 \times \mathcal{M}_2, \rho_1 \vee \rho_2)$ be their product. Show that

   (a) $\mathcal{M}$ is separable if and only if $\mathcal{M}_1$ and $\mathcal{M}_2$ are so.

   (b) Let $\mathbb{P}, \mathbb{P}_n, n \geq 1$ be probability measures on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$. Assuming $\mathcal{M}$ to be separable, show that $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ if and only if $\mathbb{P}_n(A_1 \times A_2) \to \mathbb{P}(A_1 \times A_2)$ for all $\mathbb{P} \circ \pi_1^{-1}$ continuity sets A_1 and $\mathbb{P} \circ \pi_2^{-1}$ continuity sets A_2, where π_i's are the co-ordinate projections.

   (c) Let $\mathcal{M}_i = [0, 1]$ equipped with the Euclidean metric. Let $\mathbb{P}_1$ be the uniform distribution on $\mathcal{M}$, and $\mathbb{P}_2$ be the uniform distribution on the diagonal $\{(x, x) \mid x \in [0, 1]\}$. Show that $\mathbb{P}_1$ and $\mathbb{P}_2$ have identical marginals.

   (d) Using (c) or otherwise, show, in the context of (b), that $\mathbb{P}_n \circ \pi_1^{-1} \xrightarrow{w} \mathbb{P} \circ \pi_1^{-1}$ and $\mathbb{P}_n \circ \pi_2^{-1} \xrightarrow{w} \mathbb{P} \circ \pi_2^{-1}$ together do not imply that $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

   [4 + 10 + 3 + 3]

2. (a) Show that $\lambda : [0, 1] \to [0, 1]$ is a homeomorphism with $\lambda(0) = 0, \lambda(1) = 1$ if and only if λ is a continuous strictly increasing surjection on $[0, 1]$. Show that such λ's form a group with respect to function composition.

   (b) Define the Skorohod metric d on $D[0, 1]$.

   (c) Let $\phi_n(t) = \mathbf{1}_{[\frac{1}{10} + \frac{1}{2n}, 1]}(t)$ and $\phi(t) = \mathbf{1}_{[\frac{1}{10}, 1]}(t)$. Show that $d(\phi_n, \phi) \to 0$.

   (d) Show that if $x \in C[0, 1]$ and $x_n \in D[0, 1], n \geq 1$, then $d(x_n, x) \to 0 \implies \|x_n - x\|_\infty \to 0$.

   (e) Show that if $x \in D[0, 1]$ and $x_n \in C[0, 1], n \geq 1$, then $d(x_n, x) \to 0 \implies \|x_n - x\|_\infty \to 0$.

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

3. Let $\xi_i \overset{i.i.d.}{\sim} F, i = 1, \ldots, n$. Consider the empirical process

$$X_n(t) = \sqrt{n}\left(\frac{1}{n} \sum_{1 \leq i \leq n} \mathbf{1}_{\{\xi_i \leq t\}} - F(t)\right).$$

   Derive the limit distribution of $(X_n(t_1), \ldots, X_n(t_k))$, where $t_i \in (0, 1)$. State Donsker's theorem for X_n.

   Now suppose that F is the CDF of the uniform distribution on $[0, 1]$. Then X_n is the uniform empirical process. State with proper justification if X_n is measurable with respect to

PTO

(i) the Borel σ-field on $(D[0,1], \|\cdot\|_\infty)$,

(ii) the ball σ-field on $(D[0,1], \|\cdot\|_\infty)$, and

(iii) the Borel σ-field in the Skorohod topology on $D[0,1]$.

[(3 + 2) + (5 + 5 + 5)

4. (a) Suppose $\mathcal{F}, \mathcal{F}_1, \mathcal{F}_2$ are $\mathbb{P}$-Glivenko-Cantelli classes. State with proper justification which of th following are $\mathbb{P}$-Glivenko-Cantelli classes:

   (i) $\mathcal{F}_1 \cup \mathcal{F}_2$,

   (ii) $\mathcal{F}_1 \cap \mathcal{F}_2$,

   (iii) $\{a_1 f_1 + a_2 f_2 \mid f_i \in \mathcal{F}_i, |a_i| \le 1\}$,

   (iv) $\{f \mid \exists$ a sequence $(f_m)_{m\ge1}$ in $\mathcal{F}$ such that $f_m \to f$ pointwise as well as in $L_1(\mathbb{P})\}$.

   (b) Suppose $X_i, i = 1, \ldots, n$ are independent random variables defined on some probability spac $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$. Let $\|\cdot\|_{\mathcal{H}}$ be the corresponding norm Suppose that $\|X_i\|_{\mathcal{H}} \le b_i$ almost surely, for constants $b_i > 0, i = 1, \ldots, n$. Consider the randor variable $S_n = \|\sum_{i=1}^n X_i\|_{\mathcal{H}}$. Show that

$$\mathbb{P}(|S_n - \mathbb{E}S_n| \ge n\delta) \le 2\exp\left(-\frac{n^2\delta^2}{8\sum_{i=1}^n b_i^2}\right).$$

[(2 + 2 + 4 + 6) + 6

5. (a) Define covering and bracketing numbers. Give an example where these are equivalent (up t constant factors).

   (b) State and prove the classical Glivenko-Cantelli theorem using L_1-bracketing numbers.

   (c) The δ-packing number $M(\delta; \mathcal{M}, \rho)$ of a (pseudo-)metric space $(\mathcal{M}, \rho)$ is the maximum cardinalit of a set of points in $\mathcal{M}$ that are more than δ apart in ρ. If $N(\delta; \mathcal{M}, \rho)$ denotes the δ-coverin number of $(\mathcal{M}, \rho)$, show that

$$M(2\delta; \mathcal{M}, \rho) \le N(\delta; \mathcal{M}, \rho) \le M(\delta; \mathcal{M}, \rho).$$

[(3 + 2) + 7 + (4 + 4