INDIAN STATISTICAL INSTITUTE

Mid-Semester (Supplementary) Examination
B. Stat I year, 1st Sem: 2016–2017

Introduction to Programming and Data Structures

Date: 28. 10. 2016, Maximum Marks: 50, Time: 2 Hours (4:30 PM to 6:30 PM)

Please try to write all the part answers of a question at the same place.

1. (a) With 15 bits, how many distinct integers can you represent in signed magnitude, 1's complement and 2's complement respectively?

   (b) If you multiply two n-bit numbers, what is the maximum possible length (in bits) of the product? [(2 + 2 + 2) + 4 = 10]

2. (a) Write a recursive C function that takes as inputs two positive integers and returns the GCD of the two integers.

   (b) In Euclid's algorithm to compute the GCD of two positive integers, if the quotients are always 1, then what can you say about the sequence of the remainders? [4 + 6 = 10]

3. What will be the output of the following C program? If you think it will give a runtime error, you need to mention it with justification. If you think there is no error, then does the output resemble enumeration of any known sampling procedure?

```
#include<stdio.h>
main()
{
    unsigned char i, j, k, a[] = {'a', 'b', 'c', 'd', 'e'};
    for(i = 0; i < 5; i ++)
       for(j = i+1; j < 5; j ++)
          for(k = j+1; k < 5; k ++)
             printf("%c, %c, %c\n", a[i], a[j], a[k]);
```

[8 + 2 = 10]

4. Write a C function that takes as inputs base addresses of two strings $s1$ and $s2$ and returns 1 if the reverse of $s1$ is a substring of $s2$; and 0 otherwise. [10]

5. Consider a two-dimensional integer array a of 8 rows and 12 columns. Assuming that it takes 4 bytes to represent an integer, if the address of a[2][3] is α, what is the address of a[3][2]? [10]

6. If we want to combine two already-sorted arrays of sizes m and n respectively to produce a bigger sorted array using the "merge" algorithm, what are the minimum and the maximum possible values of the number of comparisons needed? [5 + 5 = 10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examinations: 2016-17

B. Stat. I Year Analysis-I

Date: 29/08/2016 Maximum Marks: 40 Duration: 3 Hours

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

---

1. (a) Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two bounded sequences of real numbers such that $\{b_n\}_{n=1}^{\infty}$ converges to b. Show that

$$\limsup(a_n + b_n) = \limsup a_n + b.$$

(b) Give an example of a sequence $\{a_n\}_{n=1}^{\infty}$ that is not a Cauchy sequence, but that satisfies $\lim_{n\to\infty} |a_{n+1} - a_n| = 0$.

(c) Consider the series

$$\frac{1}{1.2} - \frac{1}{2.3} + \frac{1}{3.4} - \frac{1}{4.5} + (-1)^{n+1}\frac{1}{n(n+1)} + \cdots.$$

Does there exist a rearrangement of this series which is divergent?

(d) Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and $f(a) > 0$ for some $a > 0$. Show that there exists $\delta > 0$ such that $f(x) > 0$ for all $x \in (a - \delta, a + \delta)$.

(e) Let $g(x) = \sin\frac{1}{x}$. Show that g is not uniformly continuous on $(0, \infty)$. [10]

2. Let the sequence $\{x_n\}_{n=1}^{\infty}$ be defined by

$$x_1 = 1, \quad x_2 = 2, \quad \text{and} \quad x_n = \frac{x_{n-2} + x_{n-1}}{2} \quad \text{for } n \geq 3.$$

(a) Show that $|x_n - x_{n+1}| = \frac{1}{2^{n-1}}$ for all $n \geq 1$ and conclude that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

(b) For all $n \geq 1$, show that

$$x_{2n+1} = 1 + \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots + \frac{1}{2^{2n-1}}.$$

(c) Find the limit of the sequence $\{x_n\}_{n=1}^{\infty}$. [8]

3. Discuss the convergence of the infinite series

$$\sum_{n=1}^{\infty}(-1)^{\left[\frac{n^2+7}{2n}\right]}\frac{\sqrt{n+1}}{n},$$

where $[x]$ denotes the greatest integer less than or equal to x. [6]

4. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence such that $\lim_{n\to\infty} a_n = a \neq 0$ with $a_n \neq 0$ for all $n \in \mathbb{N}$. Show that $\sum_{n=1}^{\infty}(a_{n+1} - a_n)$ is absolutely convergent if and only if $\sum_{n=1}^{\infty}\left(\frac{1}{a_{n+1}} - \frac{1}{a_n}\right)$ is absolutely convergent. [5]

5. Let S be a nonempty proper subset of $\mathbb{R}$. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, \text{if } x \in S, \\ 0, \text{if } x \notin S. \end{cases}$$

Prove that f has at least one point of discontinuity. You can use the fact that the only subsets of $\mathbb{R}$ which are both closed and open are $\emptyset$ and $\mathbb{R}$.

[Hint. If $a \in S$ is a limit point of S^c, then f is discontinuous at a (Why?).] [8]

6. Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are both uniformly continuous and are both bounded functions on $\mathbb{R}$. Prove that the product fg is uniformly continuous on $\mathbb{R}$. [6]

INDIAN STATISTICAL INSTITUTE

Midsemester Examination : (2016-2017)

B. Stat 1st Year

Statistical Methods - I

Date: 30. 08. 2016 Maximum marks: 60 Duration: 2 hours.

Note: This paper carries 65 points. Maximum you can score is 60.

1. (a) In a (hypothetical) study, it was found that the percentage of incidence of hypertension was less for people who regularly exercised. The study concluded that regular exercise prevents hypertension. Give reasons for your agreement or disagreement.

   (b) Dividing the two groups (people who regularly exercise and who don't) into young and old, illustrate with hypothetical numbers how the percentage of incidence of hypertension can be higher for people who regularly exercise whether young or old, even when the overall percentages agree with the study.

   (c) Explain briefly how will you design a statistical study to check their conclusion? [3+10+5]

2. Suppose there are two sets of values of a variable x with n_1 and n_2 values respectively.

   (a) Let the geometric means of the two sets be G_1, and G_2 respectively. Show that the geometric mean of all $n_1 + n_2$ values taken together will lie between G_1 and G_2.

   (b) Suppose the harmonic means of the two sets are H_1 and H_2 respectively. Will the harmonic mean of all $n_1 + n_2$ values taken together will lie between H_1 and H_2? Justify your answer. [5+7]

3. Using the Cauchy-Schwarz inequality or otherwise, prove the following results:

   (a) Gini's mean absolute difference cannot be greater than $\sqrt(2)$ times the standard deviation of a variable.

   (b) If the measure of Kurtosis $\beta_2 = \frac{m_4}{m_2^2}$ and the measure of skewness $\beta_1 = \frac{m_3^2}{m_2^3}$ for a variable, then $\beta_2 \geq \beta_1 + 1$. [5+10]

4. The first four moments of a distribution about the value 4 (i.e. $\frac{1}{n}\sum_{i=1}^{n}(X_i - 4)^r, r = 1, \ldots, 4$) are -1.5, 17, -30 and 108, respectively. Find the first four central moments. [10]

5. Assignments [10]

Indian Statistical Institute
Semester 1 (2016-2017)
B. Stat 1st Year
Midsemestral Examination
Probability Theory 1

Date: 31.8.16 , Time: 2:30-4:30 PM, Total Points $5 \times 6 = 30$

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or result proved in class state it explicitly.

1. The pigeonhole principle says that if $(m+1)$ or more objects are placed in m cells, then there will be a cell with more than one object. As an application of this consider n distinct primes $p_1, p_2, \dots, p_n$ all greater that 1. Suppose a selection of N items is made from these primes with replacement and the selection is written as a sequence (in a random order)

$$a_1, a_2, \dots, a_N.$$

If $N > 2^n$ then show that there is a subsequence $a_l, a_{l+1}, \dots, a_{l+k}$, $1 \leq l, l+k \leq N$, such that the product of the terms of the subsequence is a perfect square. (Hint: Consider a counter $(b_1, b_2, \dots, b_n)$ which tracks the power of each of the above primes as you read the sequence $a_1, a_2, \dots, a_N$. For example if $a_1, a_2, \dots, a_N$ is say $p_1, p_2, p_1, p_3, \dots$ then the successive values of the counter are $(1,0,0,\dots,0), (1,1,0,\dots,0), (2,1,0,\dots,0), (2,1,1,\dots,0)$, etc.)

2. One number is chosen at random from all the five digit numbers in the decimal scale (i.e. each digit can be one of $0, 1, 2, \dots, 9$). What is the probability that the digits of the number chosen are in nondecreasing order? (Hint: $0 \leq a_1 \leq a_2 \leq a_3 \leq a_4 \leq a_5 \leq 9$, can be converted to $0 \leq a_1 < a_2 + 1 < a_3 + 2 < a_4 + 3 < a_5 + 4$, in a one one manner where the digits are strictly increasing.)

3. N men run out of a men's club after a fire and each takes a coat and a hat. Prove that: $3 + 3 = 6$ pts.

   (a) the probability that no one will take his own coat and hat is

$$\sum_{k=0}^{N} (-1)^k \frac{(N-k)!}{N!k!}.$$

   (b) the probability that each man takes a wrong coat and a wrong hat is

$$\left[\sum_{k=2}^{N} (-1)^k \frac{1}{k!}\right]^2.$$

4. At a counter there are $2n$ people standing in a queue, n of them have only Rs 5 coupon each and the rest n have only Rs. 10 coupon each. Each customer takes a food plate worth Rs. 5. The transactions are to be done with coupons only. If the server has no change to begin with and the customers arrive in a random order, **find the probability that no customer will have to wait for change**. Note that each customer with Rs. 5 coupon increases the server's worth by the same amount that each customer with Rs. 10 coupon decreases it by. This amount can be taken as a unit to get a random walk for the server's worth.

5. A parent particle can be divided into 0, 1 or 2 particles with probabilities $1/4, 1/2, 1/4$ respectively. It disappears after splitting. Beginning with one particle, the progenitor, let us denote by X_i, the number of paticles in the i th generation. Find (a) $P(X_2 > 0)$, (b) the probability that $X_1 = 2$ given that $X_2 = 1$. 3 +3 = 6 pts.

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination
B. Stat I year, 1st Sem: 2016–2017

Introduction to Programming and Data Structures

Date: 01. 09. 2016, Maximum Marks: 100, Time: 3 Hours (2:30 PM to 5:30 PM)

Please try to write all the part answers of a question at the same place.

1. (a) Subtract 12_{10} from 7_{10} in binary representation using two separate methods: 1's complement and 2's complement arithmetic and check if they yield the same results.

   (b) What is the difference between *infinity* and *non-a-number* in computer arithmetic? How are they represented in floating point number representation?

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

2. (a) Describe Euclid's algorithm to compute the GCD of two positive integers.

   (b) Prove that the algorithm always halts for arbitrary inputs!

   (c) Prove the correctness of the algorithm.

   [4 + 4 + 6 = 14]

3. Derive the minimum number of moves required to complete the *Towers of Hanoi* game for $n(\geq 0)$ disks. [8]

4. What will be the output of the following C program? If you think it will give a runtime error, you need to mention it. In either case, your answer must include proper justifications without which no credit will be given.

```
#include<stdio.h>
main()
{
        unsigned char i, j, a[] = {1, 2, 3, 4, 5};
        int n;
        i = j = n = 5;
        while(i-- != 0) n += a[i];
        while(j++ != 0) n += 2;
        printf("i = %d, j = %d, n = %d\n", i, j, n);
}
```

   [(2 × 3) + 4 = 10]

5. Write a C function that takes as inputs base addresses of two strings and produces a third string (whose base address is also one of the arguments of the function) by interleaving the strings, taking one character from the first string followed by one character from the second string and so on, alternately. When the end of one string is reached, the remainder of the other string should be concatenated to the resulting third string. [8]

6. (a) Consider drawing 3 samples with replacement from a population of size 5 consisting of the characters 'a', 'b', 'c', 'd', 'e'. Write a complete C program to generate all possible samples.

   (b) Solve the same problem for sampling without replacement.

   (c) Discuss how would you modify your program so that you can read the population size N and the sample size n from the user at runtime and generate the samples accordingly.

[4 + 6 + 8 = 18]

7. (a) Consider the following piece of code.

```
int a[10][15];
printf("%p %p %p %p %p\n", a, a[0], &a[0][0], a[1], &a[1][0]);
```

   If the first value printed is X, what are the other four values in terms of X?

   (b) Consider a d-dimensional array a in C with the sizes of the dimensions as $n_1, n_2, \ldots, n_d$. For an arbitrary element $a[i_1][i_2]\ldots[i_d]$, how is its address in the main memory computed? Conversely, given an address of an array element in the main memory, how are the corresponding array indices computed? Justify your answers.

[4 + (4 + 6) = 14]

8. Write a C function that will take as inputs a square matrix M, the size of its row (or column) n, and an integer t and will compute M^t. [8]

9. (a) For performing *binary search* of an element in a sorted array of n elements, what are the minimum and the maximum possible values of the number of comparisons needed?

   (b) For *insertion sort* on n real numbers, what are the minimum and the maximum possible values of the number of comparisons needed?

[(2 + 4) + (3 + 5) = 14]

Indian Statistical Institute

Vectors and Matrices I

B-I, Midsem

Date: Sep 05, 2016 Duration: 3hrs.

Attempt all questions. The maximum you can score is 50. Justify all your steps. This is an open book, open notes examination.

If copying is detected in the solution for any problem, all the students involved in the copying will get 0 for that problem. Also an additional penalty of 10 will be subtracted from the overall aggregate of each of these students.

1. Let V, W be both vector spaces over a common field F. Let $f : V \to W$ be any function. For any subset $S \subseteq V$ we define $f(S) \subseteq W$ as

$$f(S) = \{f(s) \ : \ s \in S\}.$$

Also, for any subset $T \subseteq W$ we define

$$f^{-1}(T) = \{v \in V \ : \ f(v) \in T\}.$$

Then prove or disprove each of the following statements.

(a) If f is a linear transformation, then for any subspace S of V, the set $f(S)$ is a subspace of W.

[5 marks]

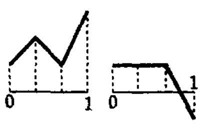
(b) If f is a linear transformation, then for any subspace T of W, the set $f^{-1}(T)$ is a subspace of V.

[5 marks]

(c) If for every subspace S of V, the set $f(S)$ is a subspace of W, and for every subspace $T \subseteq W$ the set $f^{-1}(T)$ is a subspace of V, then f must be a linear transformation.

[5 marks]

2. Let V be the vector space (over $\mathbb{R}$) of all continuous functions from $[0, 1]$ to $\mathbb{R}$ under pointwise operations. Consider $S \subseteq V$ consisting of all $f \in V$ such that the graph of f is a straight line over each of the intervals $\left[0, \frac{1}{3}\right]$ $\left[\frac{1}{3}, \frac{2}{3}\right]$ and $\left[\frac{2}{3}, 1\right]$. Here are the graphs of two elements of V :



Show that S is a subspace of V. Find (with justification) its dimension.

3. Consider $\mathcal{P}_4$, the vector space of all real polynomials with degree ≤ 3. Show that

$$S = \{p \in \mathcal{P}_4 \ : \ p(0) = p(1) = p(2)\}$$

is a subspace of $\mathcal{P}_4$. Find a direct complement of S in $\mathcal{P}_4$.

[10 marks]

4. Let A be an $n \times n$ upper triangular matrix with all diagonal entries zero. Show that

$$\forall k \in \{1, 2, ..., n\} \ \ \forall i \in \{1, ..., n-k\} \ \ \forall j \in \{i, ..., i+k-1\} \ \ ((A^k))_{i,j} = 0.$$

Here $((B))_{i,j}$ means the (i, j)-th entry of a matrix B. (Intuitively, the zeros spread diagonally upwards.)

[10 marks]

5. Let V be a vector space and $S_1, ..., S_k$ be its subspaces. If $V = S_1 \cup \cdots \cup S_k$, then show that $V = S_i$ for some i.

[10 marks]

Indian Statistical Institute
Semester 1 (2016-2017)
B. Stat 1st Year
Supplementary Midsemestral Examination
Probability Theory 1

Date: 20.10.16 , Time: 4:30-6:30 PM , Total Points $5 \times 6 = 30$

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or result proved in class state it explicitly.

1. From an urn containing $(2n+1)$ tickets numbered serially, three tickets are drawn at random without replacement. Find the probability that the numbers on them are in arithmetical progression.

2. Ten manuscripts are arranged in 30 files (3 files for each manuscript). Find the probability that no 6 files selected at random contain an entire manuscript.

3. A person tosses an unbiased coin repeatedly and moves on the upper right quadrant as follows: if the result of a toss is H he moves 1 unit to the right, if it is T he moves 1 unit up. Thus after the kth toss his coordinates are the number of heads and the number of tails upto that point. $3+3=6$ pts.

   (a) If he tosses $2n$ rimes, show that the probability that he does not touch or cross the diagonal is
   $$\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots 2n}.$$

   (b) If after $2n$ tosses he has reached (n, n), find the probability that he did not touch or cross the diagonal before reaching (n, n).

4. Show that in $2n-1$ trials of tossing a fair coin the probability of a run of n consecutive heads is $(n+1)/2^{n+1}$. (Consider where the run can start and decompose into disjoint events.)

5. An urn contains n balls, each of different color, of which one is white. Two independent observers, each with probability 0.1 of telling the truth, assert that a ball drawn at random from the urn is white. Prove that the probability that the ball is, in fact, white is $(n-1)/(n+80)$.

Indian Statistical Institute

Vectors and Matrices I

B-I, Midsem (Supplementary)

Date: Oct ~~10~~ 26 2016 Duration: 3hrs.

This paper carries 55 marks. Attempt all questions. The maximum you can score is 30. Justify all your steps. This is an open book, open notes examination.

1. Let V be a vector space over $\mathbb{R}$. Let $v_0 \in V$ be any fixed non-null vector. Let $f : V \to V$ be a linear transformation. Consider the sets $S, T \subseteq V$ defined as
$$S = \{v \in V \ : \ f(v) = 2v\},$$
and
$$T = \{f(v) - v_0 \ : \ v \in V\}.$$
Then prove or disprove each of the following statements.
   (a) S is a subspace of V.
   (b) T is a subspace of V.

[7.5×2 marks]

2. If $V = S \oplus T$ and $f : V \to W$ is an onto linear transformation, then is it always true that $W = f(S) \oplus f(T)$? [You can assume that $f(S), f(T)$ are both subspaces of W.] Justify your answer with a proof or a counterexample.

[10 marks]

3. Let $A = \{1, 2, 3\}$ and let $F = \{0, 1\}$ be the binary field (operations modulo 2). Let $V = \mathcal{P}(A)$ be the power set vector space (with symmetric difference as addition). Find the dimension of V. Find a direct complement of the subspace $W = \mathcal{P}(\{1, 2\})$ in V.

[5+10 marks]

4. Let $V = S \oplus T$. Let $\{s_1, ..., s_m\}$ and $\{t_1, ..., t_n\}$ be any two bases of S and T, respectively. Then is it true that $\{s_1, ..., s_m, t_1, ..., t_n\}$ must be a basis of V? Justify your answer with a proof or a counterexample.

[10 marks]

5. Let $f : V \to V$ be a linear transformation such that for all $v \in V$ we have $f(f(v)) = f(v)$ Then show that $Range(f)$ and $ker(f)$ are direct complements of each other in V. [No need to prove that these are subspaces of V.]

[5 marks]

INDIAN STATISTICAL INSTITUTE

End Semester Examinations: 2016-17

B. Stat. I Year Analysis-I

Date: 07/11/2016 Maximum Marks: 60 Duration: 3 Hours

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

---

(1) Find the supremum and infimum of the set $\left\{\frac{(n+1)^2}{2^n} : n \in \mathbb{N}\right\}$. [5]

(2) Let $\{a_n\}$ be a sequence of positive real numbers such that

$$\left(\limsup a_n\right)\left(\limsup \frac{1}{a_n}\right) = 1.$$

Show that the sequence $\{a_n\}$ is convergent. [5]

(3) Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be continuous functions such that

$$f(g(x)) = g(f(x)) \quad \text{for all } x \in \mathbb{R}.$$

Suppose that the equation $f(f(x)) = g(g(x))$ has a solution. Show that $f(x) = g(x)$ also has a solution. [5]

(4) Let $f : [0, \infty) \to \mathbb{R}$ be continuously differentiable on $(0, \infty)$ and let $f(0) = 1$. Suppose that $|f(x)| \leq e^{-x}$ for all $x \geq 0$. Show that there exists $x_0 > 0$ such that $f'(x_0) = -e^{-x_0}$. [10]

[Hint. Consider $g(x) = f(x) - e^{-x}, x \geq 0$. Observe that $g(x) \leq 0$ and $\lim_{x\to\infty} g(x) = 0$.]

(5) Suppose that $f : [-1, 1] \to \mathbb{R}$ is thrice differentiable and such that

$$f(-1) = 0, \quad f(0) = 0, \quad f(1) = 1, \quad f'(0) = 0.$$

Show that there exists a $c \in (-1, 1)$ such that $f'''(c) \geq 3$. [10]

[Hint. Use Taylor's theorem to show that there exists $s \in (-1, 0)$ and $t \in (0, 1)$ such that $f'''(s) + f'''(t) = 6$.]

(6) Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \frac{1}{1+|x|} + \frac{1}{1+|x-1|}.$$

Find the maximum value of f on $\mathbb{R}$. [10]

(7) Let $f : [0, \infty) \to \mathbb{R}$ be defined by

$$f(x) = x^{1/3}(1-x)^{2/3}.$$

Determine all local extreme points of f. [10]

(8) Let $f : (0, \infty) \to \mathbb{R}$ be a convex function such that

$$\lim_{x\to 0+} f(x) = 0.$$

Let $g(x) = \frac{f(x)}{x}$. Show that g is increasing on $(0, \infty)$. [10]

[Hint. Let $x_1, x_2 \in (0, \infty)$ with $x_1 < x_2$. Take $x \in (0, x_1)$.]

INDIAN STATISTICAL INSTITUTE

First Semestral Examination : Semester I (2016-17)

B. Stat 1st Year

Statistical Methods - I

Date: 9. 11. 2016 Maximum marks: 100 Duration: 3 hours.

Note: This paper carries 110 points. Maximum you can score is 100.

1. Show that the standard deviation of the scores obtained in this examination cannot exceed 50. Prove any result that you need to use. [5 + 10]

2. Using the Cauchy-Schwarz inequality or otherwise, prove that the measure of Kurtosis $\beta_2 = \frac{m_4}{m_2^2}$ is always greater than or equal to 1. Discuss in detail the case when equality holds. [5 + 5]

3. Consider a 2 x 2 contingency table with two attributes A and B each having two forms A and α, and B and β, respectively. Suppose the total frequency is n.

   Writing the cell frequencies as $f_{AB}, f_{A\beta}, f_{\alpha B}$ and $f_{\alpha\beta}$, show that if $\delta = f_{AB} - \frac{f_A f_B}{n}$, then

   $$f_{AB}^2 + f_{\alpha\beta}^2 - f_{\alpha B}^2 - f_{A\beta}^2 = (f_A - f_\alpha)(f_B - f_\beta) + 2n\delta,$$

   where f_A, f_B, f_α and f_β are the frequencies of the corresponding forms. [10]

4. Let there be k groups of data on two variables x and y, with means $\bar{x}_i$ and $\bar{y}_i$, variances s_{xi}^2 and s_{yi}^2, correlations r_i and number of observations n_i for $i = 1, \ldots, k$.

   (a) Express the corration coefficient of the combined data in terms of the above quantities. [15]

   (b) Using part(a) or otherwise, give a concrete example with specific numbers where even though each r_i is positive, the correlation coefficient of the combined data is negative. [10]

[P.T.O.]

5. (a) For the special case when there are two members in each class, express the intra-class correlation coefficient in terms of the product-moment correlation coefficient. Using this, show that when the intra-class correlation coefficient attains its extreme values, so does the product-moment correlation coefficient. [10 +8]

   (b) Consider data r_i and s_i, $i = 1, \ldots, n$ on ranks assigned (without ties) by two judges to n participants in a competition. Suppose $r_i = s_i + 1$ for $1 = 1, 2, \ldots, n-1$, and $r_n = 1$ and $s_n = n$. If the Spearman's rank correlation coefficient is 0.5, find the value of n. [10]

6. The average height and weight of a group of students turned out to be 5 ft 6 inches and 65 kilograms respectively. The standard deviations were 3 inches and 5 kilograms respectively. Using the regression equation for predicting weight based on height, the estimated weight of a 6 ft tall student was calculated to be 72 kilograms. Predict the height of a student whose weight is 60 kilograms. [12]

7. Assignments [10]

Indian Statistical Institute
Semester 1 (2016-2017)
B. Stat 1st Year
Semestral Examination
Probability Theory I

Friday 11.11.2016, 2:30-5:30 Total Points: $5 \times 14 = 70$

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or result proved in class state it explicitly.

1. A pack of cards contains s identical series, each containing n cards numbered $1, 2, \cdots, n$. A random sample of $r \geq n$ cards is drawn from the pack *without replacement.* (a) Calculate the probability u_r that each number is represented in the sample. (b) Determine the limit of u_r as $s \to \infty$.

   $7 + 7 = 14$ pts.

2. (a) n letters marked $1, 2, \ldots, n$ are to be placed in n envelopes marked $1, 2, \ldots, n$, letter i is meant for envelope i, $i = 1, 2, \ldots, n$. Suppose the letters are placed randomly (one in each envelope) and let P_n be the probability that no letter goes into the correct envelope. Find a recursion expressing P_n in terms of P_{n-1} and P_{n-2}, formulating the following possibilities through conditional probabilities: suppose letter 1 goes to the jth envelope, $j \neq 1$. There are two cases; (i) jth letter goes to envelope 1, (ii) jth letter goes to an envelope other than envelope 1 (and obviously not to j th envelope). All the steps must be justified.

   (b) In the same set up, find the expected number of letters which go to the correct envelope. $7 + 7 = 14$ pts.

3. A bag contains n balls, each ball can be black or white. The probability that the number of white balls is k is $1/(n+1)$, for $k = 0, 1, 2, \ldots, n$ (this is called a uniform prior distribution). A ball is drawn at random and found to be white. It is replaced and another ball drawn at random is found to be white. If this ball is replaced then find the conditional probability (given the previous observations) that another ball drawn at random will be found to be black. The answer must be given in a closed form simplifying any summations that may arise in the way and the answer will come out as a ratio of two polynomials of degree one in n.

4. Suppose (X_1, X_2, X_3, X_4) follow a multinomial distribution with parameters (n, p_1, p_2, p_3, p_4) where $p_1 + \cdots + p_4 = 1$.

   $7 + 7 = 14$ pts.

   (a) Find $E(X_1 X_2)$ in a closed form.

P.T.O.

(b) Find $E(X_1|X_1+X_2=k)$ in a closed form where $1<k<n-1$.

5. Suppose N is Poisson(λ) and $X_1, X_2, \ldots$ are independent and identically distributed Bernoulli(p) random variables which are independent of N. Denote $S_N = X_1 + X_2 + \cdots + X_N$.

   (a) For $0 \le t \le 1$ compute $E\left[t^{S_N}\right]$ by computing the RHS of

$$E\left[t^{S_N}\right] = E\left[E(t^{X_1+\cdots+X_N}|N)\right].$$

   (b) Also compute the mean and variance of S_N using $E\left[t^{S_N}\right]$.

$8+6=14$ pts.

INDIAN STATISTICAL INSTITUTE

Semester Examination
B. Stat I year, 1st Sem: 2016–2017
Introduction to Programming and Data Structures

Date: 15. 11. 2016, Maximum Marks: 50, Time: 2 Hours (2:30 PM to 4:30 PM)

Answer all questions. Please try to write all the part answers of a question at the same place.

1. (a) Given any sorted sequence of k numbers, we know that there are $k + 1$ positions for placing a new number, including the two end-positions. Assume that for all integers $k \geq 1$, the random variable denoting the position of the new number has uniform distribution over the $k+1$ values. With this model, find the *expected* number of comparisons needed to perform insertion sort with n integers using a linked list.

   (b) How is the result different in the case of an array implementation? Justify.

   [8 + 2 = 10]

2. Convert the postfix expression "$xyz - +uv - ab - c + /*$" to infix. Show the content of the stack at each step. [10]

3. (a) Can there be two paths between any two vertices of a tree? (Recall that typically trees are considered to be undirected).

   (b) Write the algorithm for breadth-first-search from a given node in a graph.

   (c) How will the algorithm change for depth-first-search?

   [2 + 6 + 2 = 10]

4. Build a binary search tree considering that the numbers to be stored in the nodes arrive in the following order: 13, 3, 4, 14, 12, 10, 1, 18, 2. Draw the partial tree constructed after inserting each node. [10]

5. Assume the following structure for a binary tree.

```
struct binTreeNode {
   int data;
   struct binTreeNode *leftChild;
   struct binTreeNode *rightChild;
};
```

   (a) Write recursive functions in C for preorder and postorder traversals of binary trees.

   (b) What is the inorder traversal output for the tree you constructed in Question no. 4?

   [(4 + 4) + 2 = 10]

6. (a) When is an adjacency list better than an adjacency matrix for implementation of graphs?

   (b) Let $G = (V, E)$ be any graph. For $S \subseteq V$, the induced subgraph is $G[S] = (S, E')$ such that E' consists of all of the edges in E that have both endpoints in S. How many (induced) subgraphs are there for a complete graph of n vertices? (Recall that a complete graph is considered simple and undirected by convention).

   (c) What is a planar graph? Describe an application of planar graphs.

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

Indian Statistical Institute
Vectors and Matrices I
B-I, First semestral examination, 2016-2017

Date: Nov 18, 2016 Duration: 3hrs.

This paper carries 55 marks. Attempt all questions. The maximum you can score is 50. Justify all your steps. This is an open book, open notes examination.

If copying is detected in the solution for any problem, all the students involved in the copying will get 0 for that problem. Also an additional penalty of 10 will be subtracted from the overall aggregate for this subject of each of these students.

1. Let V be a vector space, and $w \in V$ is some given vector. You are given $\{v_1, ..., v_n\} \subseteq V$ such that this particular w can be expressed uniquely as a linear combination of the v_i's. Then prove or disprove the statement: $\{v_1, ..., v_n\}$ must be linearly independent.

   [5 marks]

2. Find the reduced row echelon form of the following matrix. Justify your answer.

$$\begin{bmatrix} 2 & 9 & -2 & 4 & 7 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 2 & 1 & 1 \\ 3 & 6 & 10 & 6 & 2 \\ 0 & 0 & 0 & 6 & 5 \\ 4 & 8 & 12 & 16 & 21 \end{bmatrix}$$

   [Hint: Don't get nervous! Think!!]

   [10 marks]

3. Prove or provide a counterexample for each of the following two statements.

   (a) If A is a 10×10 invertible matrix, then it is always possible to convert it to a singular matrix by changing a single suitably chosen entry.

   (b) If A is a 10×10 singular matrix, then it is always possible to convert it to an invertible matrix by changing a single suitably chosen entry.

   [5+5 marks]

4. Let $u \in \mathbb{R}^n$ be the $n \times 1$ vector with all entries equal to 1. Consider the matrix $A = I + \alpha uu'$ where $\alpha \in \mathbb{R}$. For which values of α will A be a projection operator? *For each of these values*

   (a) find the rank of A. [5 marks]

   (b) express $\mathbb{R}^n$ as $S \oplus T$ such that A projects $\mathbb{R}^n$ onto S along T. Justify your answer, and describe S, T as explicitly as possible. [10 marks]

5. Find a rank factorisation of J, which is the 5×5 matrix with all entries equal to 1. Is the rank factorisation unique? Justify your answers.

[5+3 marks]

6. Consider the system of linear equations $Ax = b$ where A is the 15×9 partitioned matrix

$$\begin{bmatrix} 1_5 & 1_5 & 0_5 & 0_5 & I_5 \\ 1_5 & 0_5 & 1_5 & 0_5 & I_5 \\ 1_5 & 0_5 & 0_5 & 1_5 & I_5 \end{bmatrix}$$

Here 1_5 is the 5×1 vector of all 1's. Similarly, 0_5 is the 5×1 vector of all 0's. Also I_5 is the 5×5 identity matrix. Will the system be consistent for all $b \in \mathbb{R}^{15}$? What is the rank of A? [3+4 marks]

First Semester Examinations (2016-2017)
B. Stat. – 1st yr
Remedial English
100 Marks
1 ½ hours

Date : 21·11·2016

1) Write an essay on any one of the following topics. Five paragraphs are expected : 60 marks
   a) Travels
   b) Space adventures and modern technology
   c) The relevance of history

2) Fill in the blanks with appropriate prepositions (Write the full sentence): 20 marks
   a) The dog ran ___ the thief.
   b) The river flowed ___ the hill.
   c) She jumped ___ the pool.
   d) Trees are planted ___ the road.
   e) The cat climbed ___ ___ the blackboard.
   f) I will complete this assignment ___ evening.
   g) The school will remain closed ___ Monday.
   h) I will meet you ___ Monday.
   i) My brother was born ___ 1986.
   j) I prefer tea ___ coffee.
   k) I am obliged ___ you ___ your kindness.
   l) The reasons ___ choosing him are different.
   m) Try ___ think ___ it !
   n) I was travelling ___ Kolkata ___ Mumbai.
   o) The train suddenly stopped ___ Kharagpur.
   p) Are you aware that you are a part ___ this culture ?

3) Fill in the blanks with appropriate words (Write the full sentence): 20 marks

I stood ___ out at the sea. The grey-blue ___ came onto the wet sand and I felt the cold water around my bare ___. The ___ was near the western horizon, a dull red glare behind the white ___. Seagulls soared through the darkening blue ___, struggling hard against the strong ___ that blew in from the ___. Their shrill ___ rented the air.

I ___ someone call me and ___ around to face the beach. It was a little boy, his face somewhat familiar. He stood a little away ___ me, a tattered cloth bag hanging from his ___ and a large sea shell held up high in his right ___. It was the latter that he was waving at me, a grin on his ___.

"Very good shell, Sir, very good", he ___ in broken English. "I ___ it up today from this beach. Will you ___ it, Sir ? Not expensive, not expensive, only forty ___."

I ___ my head, urging him to go away.

INDIAN STATISTICAL INSTITUTE

Semester (Supplementary) Examination
B. Stat. I year, 1st Sem: 2016–2017
Introduction to Programming and Data Structures

Date: 26/12/2016 Maximum Marks: 50. Time: 2 Hours (2:30 PM to 4:30 PM)

Answer all questions. Please try to write all the part answers of a question at the same place.

1. (a) Discuss with an example how to delete a node in a doubly linked list.

   (b) How many pointers do we need to refer to a circularly linked list uniquely? Justify.

   [6 + 4 = 10]

2. Write the algorithm for converting a postfix expression into infix. [10]

3. Let n be the number of nodes in a binary tree and n_i be the number of nodes in it with exactly i children, $i \geq 0$. Let h be the height of the binary tree.

   (a) Prove that $h \geq \lceil \log_2 n_0 \rceil$.

   (b) Prove that $h \geq \lfloor \log_2 n \rfloor$.

   [5 + 5 = 10]

4. Draw the binary search tree considering that the numbers to be stored in the nodes arrive in the following order: 23, 13, 14, 24, 22, 20, 11, 28, 12. Explain how should the tree change when we delete 13. [4 + 6 = 10]

5. (a) Assume the following structure for a binary tree.

```
struct binTreeNode {
    int data;
    struct binTreeNode *leftChild;
    struct binTreeNode *rightChild;
};
```

   Write a recursive function in C for inorder traversal of a binary tree.

   (b) What are the preorder and postorder traversal outputs for the tree you constructed in Question no. 4?

   [4 + (3 + 3) = 10]

6. (a) Give the definition of a bipartite graph with an example.

   (b) Give some applications of stable marriage problem.

   [(2 + 4) + 4 = 10]

INDIAN STATISTICAL INSTITUTE

Back Paper Examination : Semester I (2016-17)

B. Stat 1st Year

Statistical Methods - I

Date: 27.12.2016 Maximum marks: 100 Duration: 3 hours.

1. (a) For the Salk vaccine trial, briefly discuss the difference between the two adopted designs and the results that were reported for them (without the actual numbers). Comment on the difference of the strength of evidence for the vaccine as was found by the two designs. What result would you have expected if the vaccine was given to every student whose parents consented? Explain why. [8 + 4 + 3]

   (b) Suppose that male and female students are applying to a particular college for admission in three different streams. Illustrate Simpson's paradox with hypothetical data for this scenario. [10]

2. For a 2 x 2 contingency table with cell frequencies a, b, c and d respectively, find the coefficients of contingency C and T in terms of the cell frequencies. [15]

3. Suppose the daily demand for petrol in a petrol pump averages about 1500 liters with a standard deviation of 200 liters. The pump gets a regular daily supply of 1800 liters of petrol. Any unused petrol is returned from the pump daily. Let p be the proportion of days when the demand exceeds the supply. Find an upper bound for p. Prove any result that you use for obtaining this upper bound. [5+5]

4. Using the Cauchy-Schwarz inequality, prove the following results

   (a) The covariance between two variables cannot exceed $\frac{1}{4}$th of the product of the ranges of the two variables. [10]

   (b) Gini's mean absolute difference cannot be greater than $\sqrt{2}$ times the standard deviation of a variable. [10]

   (c) The difference between the mean and median of a variable never exceeds the standard deviation for any variable. [10]

[P.T.O.]

5. The average height and weight of a group of students turned out to be 5 ft 6 inches and 65 kilograms respectively. The correlation coefficient was found to be 0.6. Using the regression equation for predicting weight based on height, the estimated weight of a 6 ft tall student was calculated to be 80 kilograms. Predict the height of a student whose weight is 60 kilograms. [12]

6. Derive the intra-class correlation coefficient when there are equal number of members in each group. [8]

Indian Statistical Institute

Vectors and Matrices I

B-I, First back paper/examination, 2016-2017

Date: 28 · 12 · 2016 /Supplementary Duration: 3hrs.

This paper carries 100 marks. Attempt all questions. The maximum you can score is 45. Justify all your steps. This is an open book, open notes examination.

1. Let A, B be two matrices of the same size and having the same reduced row echelon form. Then prove or disprove (with a counterexample) each of the following statements:

   (a) $\mathcal{N}(A) = \mathcal{N}(B)$.

   (b) $\mathcal{C}(A) = \mathcal{C}(B)$.

   (c) $\mathcal{R}(A) = \mathcal{R}(B)$.

   [5+5+5=15 marks]

2. Let V be a vector space over a field F. Let $\{v_1, ..., v_n\} \subseteq V$ be such that for any $\alpha_1, ..., \alpha_n \in F$ (not necessarily distinct), there is a unique linear transform $f : V \to F$ with $f(v_i) = \alpha_i$ for $i = 1, ..., n$. Then prove or disprove each of the following statements:

   (a) $\{v_1, ..., v_n\}$ must be linearly independent.

   (b) $\{v_1, ..., v_n\}$ must generate V.

   [5+5=10 marks]

3. Suggest an algorithm to find a basis of the nullspace of a matrix in reduced row echelon form. Apply your algorithm on the following matrix (if you think it is not in reduced row echelon form, then first bring it to that form).

$$\begin{bmatrix} 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

   [10+5=15 marks]

4. Prove or provide a counterexample for each of the following two statements.

   (a) If A is a 10×10 invertible matrix, then it is always possible to convert it to singular matrix by changing a single suitably chosen entry.

   (b) If A is a 10×10 singular matrix, then it is always possible to convert it to an invertible matrix by changing a single suitably chosen entry.

[5+5 marks]

5. (a) Give an $A_{m\times n}$, $b_{n\times 1}$ and $c_{n\times 1}$ (for suitable m, n of your choice) such that the linear system of equation $Ax = b$ has unique solution but $Ax = c$ has at least two distinct solutions. If you think that this is not possible, then prove it.

   (b) Give an $A_{m\times n}$, $b_{n\times 1}$ and $c_{n\times 1}$ (for suitable m, n of your choice) such that the linear system of equation $Ax = b$ has unique solution but $Ax = c$ has no solution. If you think that this is not possible, then prove it.

   You can assume that the underlying field is $\mathbb{R}$.

[10+10=20 marks]

6. Prove that, for any given idempotent matrix A_1, there is an idempotent matrix A_2 such that $A_1 + A_2$ is again idempotent. Is it true that given any two idempotent matrices B_1, B_2, there must exist an idempotent matrix B_3 such that $B_1 + B_2 + B_3$ is idempotent? Justify your answer.

[5+9 marks]

7. A matrix $A_{5\times 5}$ is such that $Ax = 1_5$ has unique solution, where 1_5 is the 5×1 vector with all entries equal to 1. Then obtain a rank factorisation of A, or prove that it is not possible to obtain a rank factorisation based on just this information.

[10 marks]

8. Let $f : V \to V$ be a linear transformation such that $\forall v \in V \;\; f(f(v)) = f(v)$. Show that $Range(f)$ and $ker(f)$ are direct complements of each other in V. Is the converse true? Justify your answer.

[3+3 marks]

Indian Statistical Institute
Semester 1 (2016-2017)
B. Stat 1st Year
Backpaper Examination
Probability Theory I

Date: 29.12.16 , Time: Total Points: $5 \times 20 = 100$

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or result proved in class state it explicitly.

1. (a) Prove geometrically that there are exactly as many paths ending at $(2n+2, 0)$ and having all interior vertices strictly above the axis as there are paths ending at $(2n, 0)$ and having all vertices above or on the axis. (b) Hence compute the value of

$$P(S_1 \geq 0, \cdots, S_{2n-1} \geq 0, S_{2n} = 0).$$

$10 + 10 = 20$ pts.

2. Suppose that the probability that the weather (fine or wet) of one day will be the same as that of the preceding day is a constant p (this model works near the coasts where season changes have little effect). If, on the basis of past records, it is assumed that the probability of the first of July, 2006, being fine is θ, a constant, then (a) determine θ_n, the probability of it being fine on the following nth day. (b) Find the limiting value of θ_n as $n \to \infty$.

$10 + 10 = 20$ pts.

3. Suppose we have two decks of n cards each numbered $1, 2, \cdots, n$. The two decks are shuffled and the cards are matched against each other. We say that a match occurs at position i if the ith card drawn from each deck has the same number. Let S_n denote the number of matches. Find (a) ES_n and (b) $Var(S_n)$.

$10 + 10 = 20$ pts.

4. The random variables $X_1, X_2, \cdots, X_n$ have the same expectation 0 and $Var(X_i) = 1, \forall i$, $Cov(X_i, X_j) = \rho, \forall i \neq j = 1, 2, \cdots, n$. Prove that $\rho \geq -\frac{1}{n-1}$.

5. A fair coin is tossed repeatedly and independently. Let T denote the time from the begining until the observation of the first HH. Find the expectation of T in a closed from. (Conditional expectation may be used to do this computation.)

First Semester Examination (2016-2017)
Back Paper
B. Stat. - 1 yr
Remedial English
100 Marks
$1 \frac{1}{2}$ hours
30.12.2016

1. Write an essay on *any one* of the following topics. Five paragraphs are expected.

   (a) Pollution

   (b) An Unforgettable Character

   (c) School Days

   (60 marks)

2. Fill in the blanks with appropriate prepositions:
He was a student __________ Mathematics and an expert __________ Indo-European languages. I met him __________ Chennai __________ a conference. He was __________ a white suit and __________ lunch preferred non-vegetarian food __________ the vegetarian fare. __________ enquiry I found he had diverse interests ranging __________ poetry __________ music. The conference lasted __________ three days __________ __________ the last day was set aside __________ interaction. I stayed back. He left __________ Kolkata, his home town, but before he left I thanked him __________ __________ __________ the team __________ being the life and soul __________ the meet.

   (20 marks)

3. Fill in the blanks with appropriate words:
A letter is a substitute for __________ contact. It __________ __________ a way of making __________ that __________ __________ __________ record of __________ which __________ proof. This is why a business telephone call is __________ followed by __________ __________ of information.

   (20 marks)

INDIAN STATISTICAL INSTITUTE

End Semester Examinations: 2016-17 (Backpaper)

B. Stat. I Year Analysis-I

Date: 30/12/2016 Maximum Marks: 100 Duration: 3 Hours

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

---

(1) Find the supremum and infimum of the following sets:

$$A = \left\{ \frac{m}{|m| + n} : m \in \mathbb{Z}, n \in \mathbb{N} \right\}, \quad B = \left\{ \frac{mn}{1 + m + n} : m, n \in \mathbb{N} \right\}.$$

[8]

(2) Let $\{a_n\}$ be a sequence such that

$$a_1 = \frac{3}{2} \quad \text{and} \quad a_n = \sqrt{3a_{n-1} - 2}, \quad n \geq 2.$$

Show that $\{a_n\}$ is convergent and find its limit. [8]

(3) Suppose that a sequence $\{a_n\}$ satisfies the condition

$$|a_{n+1} - a_{n+2}| < r|a_n - a_{n+1}|,$$

where $0 < r < 1$. Show that $\{a_n\}$ is convergent. [8]

(4) Find all real numbers a such that the series $\sum_{n=2}^{\infty} \frac{1}{n(\log n)^a}$ is convergent. [8]

(5) Determine whether the following series is absolutely convergent, conditionally convergent or divergent:

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

[8]

(6) Determine the set of continuity of f given by

$$f(x) = \begin{cases} x^2 - 1, & x \text{ is irrational}, \\ 0, & x \text{ is rational}. \end{cases}$$

[8]

(7) Let $f : \mathbb{R} \to \mathbb{R}$ satisfies the condition $f(f(x)) = -x$ for all $x \in \mathbb{R}$. Show that f cannot be continuous. [8]

(8) Examine the uniform continuity of the following functions on $(0, \infty)$:

$$f(x) = x \sin(1/x), \quad g(x) = e^{-1/x}.$$

[8]

(9) Let f be differentiable on $[a, b]$ and let $f'(a) = f'(b) = 0$. Suppose that f'' exists on (a, b). Show that there exists $c \in (a, b)$ such that

$$|f''(c)| \geq \frac{1}{(b-a)^2} |f(b) - f(a)|.$$

[8]

(10) Let $a > 0, a \neq 1$. Evaluate the limit:

$$\lim_{x \to \infty} \left(\frac{a^x - 1}{x(a-1)} \right)^{1/x}.$$

[8]

(11) Let $f : [1, \infty) \to \mathbb{R}$. f is continuously differentiable on $(1, \infty)$ and $f(1) = 1$. Suppose that $|f(x)| \leq \frac{1}{x}$ for all $x \geq 1$. Show that there exists $x_0 > 0$ such that $f'(x_0) = -\frac{1}{x_0^2}$. [10]

(12) Let $m, n \in \mathbb{N}$. Find all local extreme points of the function $f(x) = x^m(1-x)^n$. [10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016-17

B. Stat. I Year Analysis-II

Date: 23/02/2017 Maximum Marks: 40 Duration: 3 Hours

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

---

(1) Give an example of a continuous function $f : [0, \infty) \to \mathbb{R}$ with the property that $\sum_{n=1}^{\infty} f(n)$ converges but $\int_0^\infty f(x)\, dx$ diverges. [4]

(2) Let $f \in \mathcal{R}[a, b]$ and $f(x) = g(x)$ except for a countable number of points $x \in [a, b]$. It it always true that $g \in \mathcal{R}[a, b]$?. [4]

(3) Is the improper integral $\int_0^{\pi/2} \frac{1}{\sin x}\, dx$ convergent?. [4]

(4) Find the values of $K > 0$ such that the improper integral

$$\int_0^\infty \left(\frac{1}{\sqrt{2x^2 + 1}} - \frac{K}{x+1} \right) dx$$

is convergent. [6]

(5) Let $E \subset \mathbb{R}$. Suppose that $\{f_n\}_{n=1}^\infty$ converges uniformly to f on E and $\{g_n\}_{n=1}^\infty$ converges uniformly to g on E. If f and g are bounded on E, then show that $\{f_n g_n\}_{n=1}^\infty$ converges uniformly to fg. [6]

(6) Let $f_n(x) = \frac{nx+1}{n}$ and $f(x) = x$, $n \in \mathbb{N}$, $x \in \mathbb{R}$. Show that $f_n \to f$ uniformly on $\mathbb{R}$. Does f_n^2 converge uniformly to f^2 on $\mathbb{R}$? [7]

(7) Show that the series

$$\sum_{n=1}^{\infty} \frac{nx^2}{n^3 + x^3}$$

is uniformly convergent in any finite interval $[0, a]$. [4]

(8) For $n \in \mathbb{N}$, consider the functions $f_n : [0, \infty) \to \mathbb{R}$ defined by

$$f_n(x) = \frac{x^n}{x^n + 1}.$$

Investigate pointwise and uniform convergence of the series $\sum_{n=1}^{\infty} f_n$. If the series is uniformly convergent only on a subset of $[0, \infty)$, then find that subset. [7]

INDIAN STATISTICAL INSTITUTE

Midsemester Examination : (2016-2017)

B. Stat 1st Year

Statistical Methods -II

Date: 24. 2. 17 Maximum marks: 40 Duration: 2 hours.

Note: This paper carries 65 points. Maximum you can score is 60. You may use any results proved in the class by stating the results clearly. If you use other results not discussed in class, they need to be proved. You may use calculators.

1. In the multiple regression model with an intercept, assume that there are m predictor variables which have mean 0.

   (a) Derive the expression for different sum of squares in the ANOVA decomposition and find their expected values. [10]

   (b) Assuming further that the independent variables are uncorrelated, derive the relationship between (i) the least-square estimates and (ii) the sum of squares due to regression with the corresponding quantities in the m simple linear regressions with one independent variable taken at a time. [10]

2. Consider data on three variables X_1, X_2 and X_3. Suppose a multiple regression of X_1 on X_2 and X_3 is performed to yield the equation $\hat{X}_1 = \hat{\beta}_0 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$. Suppose a simple regression of X_2 on X_3 is performed and then another simple linear regression of X_1 is performed on the residuals $x_{2.3}$ of the first equation to get the equation $\hat{X}_1 = \beta_0^* + \beta_2^* x_{2.3}$. Derive an inequality involving β_2^* and $\hat{\beta}_2$. [15]

3. The following table gives the correlation coefficients between 4 x-variables. Using them, find (a) $r_{14.23}$ and (b) $R^2_{1(234)}$. [12 + 8]

1	.8	.6	.6
.8	1	.8	.6
.6	.8	1	.8
.6	.6	.8	1

4. Assignments [10]

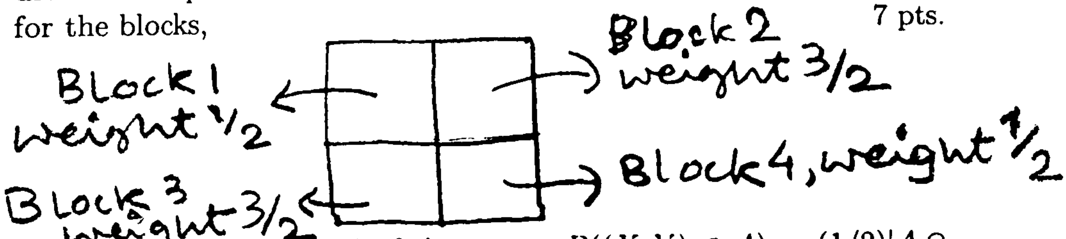
Indian Statistical Institute
Semester 2 (2016-2017)
B. Stat 1st Year
Mid-Semestral Examination
Probability Theory 2

Monday, 20.2.2017, 2:30-4:30 PM Total Points: 30

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or lemma proved in class state it explicitly.

1. A point is uniformly chosen from the square $[0,1] \times [0,1]$ (which means that the probability that it belongs to a subset A of the square is the area of A divided by the area of the square) and let (X, Y) be the coordinates of this point. Consider the random variable $|X - Y|$ and note that its minimum value 0 is attained when (X, Y) is on the diagonal connecting $(0,0)$ and $(1,1)$, and the maximum value 1 is attained when $(X,Y) = (0,1)$ or $(1,0)$. Find the cumulative distribution function of $|X - Y|$, i.e. find $P(|X - Y| \leq x)$ for $0 \leq x \leq 1$. 8 pts.

2. In the above example it is easily verified that $X \sim U(0,1)$ and $Y \sim U(0,1)$. Instead of choosing uniformly, if one chooses with different probabilities from different parts of the unit square, even then it is possible that $X \sim U(0,1)$ and $Y \sim U(0,1)$, but the distribution of $|X - Y|$ is different from that in problem 1. As an instance one can divide the square in four equal blocks and have the following weights for the blocks, 7 pts.



so that for any subset A of the square $P((X,Y) \in A) = (1/2)|A \cap block\, 1| + (3/2)|A \cap block\, 2| + (3/2)|A \cap block\, 3| + (1/2)|A \cap block\, 4|$, $|B|$ denoting the area of the set B. Show that whenever the point is chosen in such a way that $X \sim U(0,1)$ and $Y \sim U(0,1)$, then $E|X - Y| \leq 1/2$.

P.T.O.

3. (a) Can the following function

$$g(t) = \frac{e^{\lambda e^t} - 1}{e^\lambda - 1},$$

where $\lambda > 0, t \in \mathbb{R}$, be a moment generating function? Why or why not?

(b) Consider the following series in even powers of $t \in \mathbb{R}$ such that near the origin the expansion is of the form $g(t) = 1 + 3t^2 + t^4 + r(t)$ where $r(t)/t^4 \to 0$ as $t \to 0$. Can this be a moment generating function? Why or why not?

$4 + 4 = 8$ pts.

4. Suppose $X \sim N(\mu, 1)$ where $\mu > 0$. Find $Eg(X)$ in closed form where

$$g(x) = e^{x^2/2} \int_x^\infty e^{-t^2/2} dt, x \in \mathbb{R}.$$

7 pts.

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
B. Stat I year, 2nd Sem: 2016–2017
Numerical Analysis

Date: 21. 02. 2017, Maximum Marks: 100, Time: 3 Hours (2:30 PM to 5:30 PM)

Please try to write all the part answers of a question at the same place.

1. (a) What is *Regula Falsi* method?

   (b) Does it always have better rate of convergence than the *bisection* method? If yes, prove it. If no, then explain how can you modify the *Regula Falsi* method to have provably better convergence than the bisection method?

   (c) Find a real root of $f(x) = 2^x - 3x = 0$ correct up to 3 decimal places using all the three above methods separately and compare the number of iterations.

   [2 + (2 + 8) + (4 + 4 + 6) = 26]

2. (a) For *Newton's divided difference* method, prove that

   $$f[x_0, x_1, \ldots, x_n] = f[x_{i_0}, x_{i_1}, \ldots, x_{i_n}]$$

   for any permutation $(i_0, i_1, \ldots, i_n)$ of $(0, 1, \ldots, n)$.

   (b) What is the simplified expression of $f[x_0, x_1, \ldots, x_n]$, when $x_0 = x_1 = \cdots x_n$?

   [10 + 6 = 16]

3. (a) What is the geometric interpretation of *Newton-Raphson* method?

   (b) Derive a recursive algorithm for finding the square root of a non-negative real number using Newton-Raphson method and write the pseudocode for your derived algorithm.

   (c) Extend the above algorithm for finding the n-th root of a non-negative real number.

   (d) Use your extended algorithm for finding the positive real cube-root of 10 correct up to eight decimal places.

   [4 + (6 + 4) + 4 + 8 = 26]

4. (a) If $V(x_1, \ldots, x_n)$ is an $n \times n$ *Vandermonde* matrix, show that its determinant is equal to $\Pi_{1 \leq j < i \leq n}(x_i - x_j)$.

   (b) Assume that x_i's are all distinct. Prove that $V\vec{b} = \vec{0}$ implies $\vec{b} = \vec{0}$, without using the non-zero determinant property of V.

   [8 + 6 = 14]

5. Consider interpolation with equally-spaced points. Prove the following result, where the notations have their usual meaning.

$$f_p = \sum_{i=0}^{n} \binom{-p+i-1}{i} \nabla^i f_n.$$

[10]

6. (a) Given the following table, approximate $f(x)$ as a polynomial in x using *Newton's divided difference formula.*

x	4	5	7	10	11	13
$f(x)$	48	100	294	900	1210	2028

(b) Estimate $f(8)$ for the above function.

[10 + 4 = 14]

7. (a) What is the motivation for *Hermite interpolation*?

(b) Determine the Hermite interpolating polynomial fitting the following data.

x	-1	0	1
y	-10	-4	-2
y'	10	3	2

[2 + 12 = 14]

Indian Statistical Institute
Vectors and Matrices II
B-I, Midsem

Date: Feb 22, 2017 Duration: 3hrs.

Attempt all questions. The maximum you can score is 50. Justify all your steps. This is an open book, open notes examination.

If copying is detected in the solution for any problem, all the students involved in the copying will get 0 for that problem. Also an additional penalty of 10 will be subtracted from the overall aggregate of each of these students.

1. State true or false with proof or counterexample, as appropriate.

   (a) Any symmetric (not necessarily Hermitian) matrix with complex entries must have all eigenvalues real.

   [5 marks]

   (b) The x-y plane and the y-z plane are mutually orthogonal subspaces in $\mathbb{R}^3$.

   [5 marks]

2. Let A be a real symmetric matrix. Let $S_{n-1} \subseteq \mathbb{R}^n$ be the unit sphere

$$S_{n-1} = \{(x_1, ..., x_n) \in \mathbb{R}^n \ : \ \sum x_i^2 = 1\}.$$

   Consider the function $f : S_{n-1} \to \mathbb{R}$ defined as

$$f(x_1, ..., x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

   where $A \equiv ((a_{ij}))$ is a real, symmetric matrix. Show that the maximum value of $f(x_1, ..., x_n)$ is attained when $(x_1, ..., x_n)$ is an eigen vector of A corresponding to the largest eigenvalue of A.

   [10 marks]

3. Show that for every linear trasformation from a finite dimensional real vector space V to itself, there is a basis $\{v_1, ..., v_n\}$ of V such that

$$\forall i \in \{1, ..., n\} \ \ f(v_i) \in span\{v_1, ..., v_i\}.$$

   [5 marks]

4. We have two types of balls: red and blue. All red balls weigh x grams and all blue balls weigh y grams. A number of approximate measurements have led to the following data:

   - 2 red plus 3 blue balls weigh approximately m_1 grams.

- 1 red plus 1 blue ball weigh approximately m_2 grams.
- 2 red plus 2 blue balls weigh approximately m_3 grams.
- 5 red plus 3 blue balls weigh approximately m_4 grams.

Based on these measurements it is required to estimate the true weights x and y as best as possible in the least squares sense, *i.e*, if, in the four weighings above, the actual weights are $w_i(x,y)$ for $i = 1,2,3,4$, then we want to minimise

$$\sum_1^4 (w_i(x,y) - m_i)^2.$$

Suggest how you can use QR decomposition to do this. [10 marks]

5. Consider the canonical inner product for the basis

$$\{ (1,0,0),\ (1,1,0),\ (1,1,1) \}.$$

Perform Gram-Schmidt Orthonormalisation of the basis $\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ w.r.t. this inner product.

[15 marks]

6. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation that denotes the orthogonal projection to a proper subspace S of $\mathbb{R}^n$. Take an ONB of S, and extend it to an ONB of $\mathbb{R}^n$. What will be the matrix of T w.r.t. this basis? Justify your answer.

[5 marks]

Indian Statistical Institute

Vectors and Matrices II

B. Stat. (Hons) first year
Semestral Examination

Date: Apr 24, 2017 Duration: 3hrs.

Attempt all questions. The maximum you can score is 50. Justify all your steps. This is an open book, open notes examination.

If copying is detected in the solution for any problem, all the students involved in the copying will get 0 for that problem. Also an additional penalty of 10 will be subtracted from the overall aggregate of each of these students.

1. If A is a positive definite matrix and B is a nonnegative definite matrix, where $det(A+B) = det(A)$, then show that B must be zero matrix.

   [5 marks]

2. Let A be a real symmetric matrix such that

   $$\forall i, j \in \{1, ..., n\} \quad (|i-j| > 1 \Rightarrow a_{ij} = 0).$$

   If the $n-1$ numbers $a_{2,1}, ..., a_{n,n-1}$ are all nonzero, then show that the eigenvalues of A must all be distinct. [10 marks]

3. Let G be the Moore-Penrose g-inverse of $A_{m \times n}$. Then show that for any $b \in \mathbb{R}^m$ the vector $x_0 = Gb \in \mathbb{R}^n$ is a minimum norm least squares solution of the system $Ax = b$. Is it true that a minimum norm least squares solution is unique in general? Justify your answer with a proof or a counter-example. [10+5 marks]

4. Let $A_{9 \times 9}$ be a nonnegative definite matrix. We partition it as

   $$A = \begin{bmatrix} B & C \\ C' & D, \end{bmatrix}$$

   where B is 5×5. Show that $\mathscr{C}(C) \subseteq \mathscr{C}(B)$. Is it true that $\mathscr{C}(C) \subseteq \mathscr{C}(D)$? Justify your answer. Here $\mathscr{C}(M)$ denotes the column space of a matrix M. [5+5 marks]

5. Construct a non-Hermitian square matrix A (over $\mathbb{C}$) such that there is an ONB of $\mathbb{C}^n$ consisting of eigenvectors of A. If this is impossible, prove the impossibility. [10 marks]

6. A complex matrix $A_{9 \times 9}$ is such that every 2-dimensional subspace of $\mathbb{C}^9$ is an eigenspace of A. Find the characteristic polynomial of A as explictily as possible. [5 marks]

INDIAN STATISTICAL INSTITTUE, KOLKATA

Second Semestral Examination 2016-17

B. Stat. (Hons.) II Year

Subject: Economic and Official Statistics and Demography

Date: 24.04.17 Maximum Marks 100 Duration: 4 hours

*(**Instructions**: This question paper has two Groups, A and B. Answer Group A and Group B on separate answer booklets. **Each Group carries a maximum of 50 marks**. Marks allotted to each question are given within parentheses. **The total duration is 4 hours**. Standard notations are followed.)*

Group A: Demography and Economic Statistics

(Answer as many as you can.)

1. Write short notes on the following. All used symbols must be clearly defined.

   a) Gross Reproduction Rate and Net Reproduction Rate;

   b) Important properties of Lorenz curve;

   c) Relation between Lorenz Ratio and Gini Mean Difference.

[4 + 4 + 4 = 12]

2. Answer the following.

   i) The equation

   $q_x = 2m_x / (2 + m_x)$

   shows how the m-type and q-type mortality rates are related to one another. Derive a similar equation for the more general case with age groups of width n years.

   ii) The standardized death rate for town A was 1.23 when the population of town B was used as the standard. What does this tell you about mortality in A to that in B?

   iii) Show that the crude birth rate in a stationary population corresponding to a life table is equal to $(1/e_0)$ where e_0 is the life expectation at birth. [4 + 2 + 4 = 10]

3. (a) Express ${}_nL_x$ in terms of l_x, ${}_na_x$ and ${}_nd_x$.

(b) Why q_x cannot be calculated directly? Describe how it gets estimated.

(c) If the crude birth rate in a country remains constant over a number of years, but the general fertility rate increases steadily, what does this tell you about the dynamism of the country's population? [3 + (1 + 3) + 3 = 10]

4. Explain how Chandra Sekar and Deming have estimated the number of vital events, missed by both the sample registration system and the sample survey. State clearly the conditions that are to be satisfied for validity of the estimate. [4 + 3 = 7]

5. What is the purpose of El-Badry's procedure? Explain the procedure. All symbols must be defined clearly. [1 + 9 = 10]

6. The table below gives the parity progression ratios for a number of birth cohorts in a country.

(i) Assuming that no woman in any of these birth cohorts had a fifth child, calculate

(a) the proportion of women in each birth cohort who had exactly 0, 1, 2, 3 and 4 children,

(b) the total fertility rate for women in each birth cohort.

(ii) Comment on your results.

Calendar years of birth	Parity Progression Ratios			
	0 -1	1 – 2	2 – 3	3 – 4
1931-33	0.861	0.804	0.555	0.518
1934-36	0.885	0.828	0.555	0.489
1937-39	0.886	0.847	0.543	0.455
1940-42	0.890	0.857	0.516	0.416
1943-45	0.892	0.854	0.458	0.378
1946-48	0.885	0.849	0.418	0.333

[(6 + 3) + 1 = 10]

Group B: Economic and Official Statistics

[Note: Answers should be to the point and brief. Group B has two parts – I and II]

Part I: ***Answer Question no. 1 and any two from the rest***

1. State the exact forms and elasticities of Engel Curve related to the following non-linear forms: Log inverse, Semi- log, Double- log. [3+3+3=9]

2. a) State the forms of Cobb Douglas and CES Production Functions along with all assumptions.

   b) Derive the elasticity of substitution for Cobb Douglas Production Function. [4+4=8]

3. a) Explain the properties of Price Index numbers.

   b) Briefly discuss about the types of errors observed while computing such index numbers.

   [5+3=8]

4. Explain the Homogeneity Hypothesis for effects of variation in household size in Engel curve analysis. Also state the implications of this hypothesis. [5+3=8]

Part II: ***Answer Question no. 1 and any two from the rest***

1. Explain the Sampling Registration Scheme followed by the Office of the Registrar General and Census Commissioner, India. [9]

2. Explain the types of Poverty Data published by the World Bank. [8]

3. Mention, in brief, the functions of Niti Aayog. [8]

4. State the key indicators of All India Debt and Investment Survey conducted by NSSO (70th Round). [8]

5. Explain, in brief, the type of Industrial Statistics compiled by CSO, Ministry of Statistics and Programme Implementation, Government of India. [8]

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination : (2016-2017)

B. Stat 1st Year

Statistical Methods -II

Date: 26. 4. 17 Maximum marks: 100 Duration: 3 hours.

Note: This paper carries 110 points. Maximum you can score is 100.

1. Consider data on three variables X_1, X_2 and X_3.

   (a) Suppose a simple regression of X_1 on X_2 is performed to get the equation $\hat{X}_1 = \hat{\beta}_0 + \hat{\beta}_2 X_2$. If a multiple regression of X_1 on X_2 and X_3 is performed now by how much will the regression coefficient of X_2 change? [10]

   (b) Suppose the simple regression equations $\hat{X}_1 = 3 + X_2$, $\hat{X}_1 = 2 + 4X_3$, $\hat{X}_2 = -1 + 2X_3$ and the pairwise correlations ($r_{12} = r_{23} = r_{13} = 0.5$) are available but the original data is lost. Using part (a) or otherwise, can you find the multiple regression equation of X_1 on X_2 and X_3 based on those information? [10]

2. For simulating a random observation X from Negative Binomial distribution with the parameters r and p, obtain three algorithms based on the approaches given below.

   (a) Use the relationship between a negative binomial random variable and a geometric random variable. [7]

   (b) Let $p_j = P[X = j]$. Use a recursive relationship between p_j and p_{j+1}. [7]

   (c) Use the interpretation that X is the waiting time for the r-th success in repeated Bernoulli trials. [6]

3. Suppose X has the following distribution function:

$$F(x) = 1 - exp(-2x^2), \quad 0 < x < \infty.$$

   (a) Give an algorithm based on the inverse transform method for simulating a random observation from this distribution. [5]

   (b) Give an algorithm based on the rejection method for the same problem. Choose an appropriate value of the parameter for the distribution that you choose for this purpose so that the efficiency of this algorithm is maximized. [15]

[P.T.O.]

4. Let $X_1, \ldots, X_n$ be a sample from the following distribution:

$$f(x \mid \theta) = \frac{x}{\theta^2} exp(-\frac{x^2}{2\theta^2}), \quad x \geq 0.$$

Compare the method of moments estimate and MLE for θ. [15]

5. In an ecological study of the feeding behaviour of birds, the number of hops between flights was counted for several birds. For the data given below, see whether an appropriate geometric distribution can give a good fit.

Number of hops	Frequency
1	48
2	31
3	20
4	9
5	6
6	5
7	4
8	2
9	1
10	1
11	2
12	1

(a) Check whether the method of moments and method of ML give same estimates of the parameter. [10]

(b) Calculate the chi-square statistic for goodness of fit. Comment on the results. [15]

6. Assignments. [10]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2016-17

B. Stat. I Year   Analysis-II

Date: 28/04/2017   Maximum Marks: 60   Duration: 3 Hours

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

---

(1) Let $\alpha > 0$. Discuss the improper integral $\int_1^\infty \frac{\sin x}{x^\alpha}\,dx$ for convergence and absolute convergence. [10]

(2) For each of the following power series, find the radius of convergence and determine whether the power series converges at the endpoints of the interval of convergence.

(a) $\sum_{n=1}^{\infty} \frac{2^n}{n^2} x^n$,   (b) $\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!}(x-1)^n$. [12]

(3) Let the radius of convergence of $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be equal to 1. Suppose $\lim_{n\to\infty} a_n = 0$. Prove that $\lim_{x\to 1} (1-x)f(x) = 0$. [6]

(4) For $x \in \mathbb{R}$, define

$$f(x) = \sum_{n=0}^{\infty} \frac{\cos n^2 x}{e^n}.$$

For $k \in \mathbb{N}$, let $f^{(k)}(x)$ denote the k-th derivative of f at x.

(a) Show that f is infinitely differentiable on $\mathbb{R}$.

(b) For $k \in \mathbb{N}$, show that $f^{(2k-1)}(0) = 0$ and $f^{(2k)}(0) = (-1)^k \sum_{n=0}^{\infty} \frac{n^{4k}}{e^n}$.

(c) For $x \neq 0$, show that there exists $K_x \in \mathbb{N}$ such that $\left|\frac{f^{(2k)}(0)}{(2k)!} x^{2k}\right| > 1$ for all $k > K_x$.

(d) Conclude that the Taylor series of f about 0 diverges for $x \neq 0$. [12]

(5) Show that the trigonometric series $\sum_{n=2}^{\infty} \frac{1}{\log n} \sin nx$ is convergent for each $x \in \mathbb{R}$, but is not the Fourier series of any Riemann integrable function. [10]

(6) Let $f(x) = |x|$, $x \in [-\pi, \pi]$. Use Parseval's identity to prove that [10]

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^4} = \frac{\pi^4}{96} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

(7) Let $p > 0$. Show that the unordered series $\sum_{(m,n)\in\mathbb{N}\times\mathbb{N}} \frac{1}{(m+n)^p}$ is summable if and only if $p > 2$. [10]

[**Hint.** $\sum_{m=1}^{n} \frac{1}{(m+n)^p} > \frac{n}{(2n)^p}$.]

Indian Statistical Institute
Semester 2 (2016-2017)
B. Stat 1st Year
Final Examination
Probability Theory 2

Date: 2.5.2017 **Time:** 2:30 PM - 5:30 PM Total Points: 70

Answers must be justified with clear and precise arguments. More than one uncrossed answers to the same question or part of a question will not be entertained and only the first uncrossed answer will be graded. If you use any theorem or lemma proved in class state it explicitly.

1. X and Y are i.i.d. exponential random variables with parameter λ (or mean $1/\lambda$). Find the moment generating function of $|X - Y|$ in closed form. 15 pts.

2. (X, Y) are jointly continuous with density $f(x, y) = c \cdot e^{\frac{y}{x+y}}, x, y > 0, x + y \leq 1$ and $f(x, y) = 0$ elsewhere. Find the constant c that makes $\int_0^\infty \int_0^\infty f(x, y)dxdy = 1$, in terms of known mathematical constants and find EY in closed form. 15 pts.

3. X and Y are i.i.d. random variables with density $f(x) = \frac{c}{1+x^4}, x \in \mathbb{R}$, where c is a constant that makes $\int_{\mathbb{R}} f(x)dx = 1$. Find the probability density of X/Y in closed form and also find c in terms of known mathematical constants. 15 pts.

4. (X, Y) follows a bivariate normal distribution with means μ_1, μ_2, variances σ_1^2, σ_2^2 and correlation coefficient ρ. Find the probability that

$$\frac{1}{(1-\rho^2)}\left\{\left(\frac{X-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{X-\mu_1}{\sigma_1}\right)\left(\frac{Y-\mu_2}{\sigma_2}\right) + \left(\frac{Y-\mu_2}{\sigma_2}\right)^2\right\} \leq r^2$$

in closed form in terms of r and known mathematical constants. 15 pts.

5. Suppose (X_1, X_2, X_3) has density $f(x_1, x_2, x_3) = x_1 x_2 e^{-x_1 x_2 x_3}, 0 < x_1, x_2 < 1, x_3 > 0$, and $f(x_1, x_2, x_3) = 0$ elsewhere. Find the conditional density of $Y_3 = X_1 X_2 X_3$ given $Y_1 = X_1, Y_2 = X_2$. 10 pts.

INDIAN STATISTICAL INSTITUTE

Semester Examination

B. Stat I year, 2nd Sem: 2016–2017

Numerical Analysis

Date: 05. 05. 2017, Maximum Marks: 50, Time: 2.5 Hours (2:30 PM to 5 PM)

Total marks is 60, but maximum attainable marks is 50.

Please try to write all the part answers of a question at the same place.

1. (a) For a given n, how can one choose the points $x_0, \ldots, x_n$ in $[a, b]$ such that $\max_{x\in[a,b]} |\pi_n(x)|$ is as small as possible, where $\pi_n(x) := (x - x_0)\cdots(x - x_n)$?

   (b) Define a cubic spline function. What is the difference between complete spline, natural spline and not-a-knot spline?

   [4 + (3 + 3) = 10]

2. (a) Suppose a numerical differentiation rule is of the form $f'(x_0) = a_0 f_0 + a_1 f_1 + a_2 f_2$, where f_i denotes $f(x_i)$ and $x_1 - x_0 = x_2 - x_1 = h$. Find the values of a_0, a_1 and a_2 so that the rule is exact for all polynomials of degree 2 or less in $[x_0, x_2]$. Also, find the error term.

   (b) Under what condition is $n! f[x_0, \ldots, x_n]$ a good approximation to $f^{(n)}(x)$?

   [(6 + 2) + 2 = 10]

3. (a) Evaluate $\int_0^1 e^{-x^2} dx$ using Trapezoidal rule and Simpson's one-third rule taking $h = 1/6$.

   (b) What are the order of the errors in each case?

   [(4 + 4) + 2 = 10]

4. (a) Let $P_0(x), P_1(x), \ldots, P_k(x)$ be a sequence of orthogonal polynomials in the interval (a, b), where $P_i(x)$ is of exact degree i. Show that $P_n(x)$ has exactly n simple real zeros, all of which lie in the interval (a, b).

   (b) What is the advantage of Gaussian quadrature over Newton-Cote's integration?

   (c) Find an approximate value of $\int_{-1}^1 \sqrt{1-x^2}\cos x dx$ using Gauss-Chebyshev and Gauss-Legendre three-point quadrature. Assume that the roots of the third degree Chebyshev polynomial are $-\sqrt{3}/2$, 0 and $\sqrt{3}/2$ and those of the third degree Legendre polynomial are $-\sqrt{0.6}$, 0 and $\sqrt{0.6}$ respectively.

   [4 + 2 + (4 + 4) = 14]

5. (a) Prove that the global error e_n in Euler's method of solving an initial value ODE is bounded by $\frac{hM}{2K}\left(e^{(x_n - x_0)K} - 1\right)$, where h is the step size, $K := \|f_y(x, y)\|_\infty$ and $M = \max_{x_0 \le x \le x_N} |y''(x)|$, $0 \le n \le N$.

   (b) Why is Runge-Kutta method more practical than Euler's method or Taylor's algorithm of higher orders for solving an ODE?

   [6+2 = 8]

6. (a) Find the Cholesky factorization of $\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$.

   (b) Which method of solving simultaneous linear equations between Gauss-Jacobi and Gauss-Seidel iterations is expected to converge faster, assuming that all conditions of convergence hold?

   [6 + 2 = 8]

Indian Statistical Institute
Vectors and Matrices II
B. Stat. (Hons), first year
Back paper examination

Date: 10.07.2017

Duration: 3hrs.

This paper carries 100 marks. Attempt all questions. The maximum you can score is 45. Justify all your steps. This is an open book, open notes examination.

1. Let A be an $n \times n$ matrix such that 0 is an eigenvalue of A with algebraic multiplicity n. Prove that A^k is zero matrix for some $k \in \mathbb{N}$.

   [10 marks]

2. Prove or disprove: If a matrix $A_{5\times 5}$ has 0 as an eigenvalue with algebraic multiplicity 3, then A must have rank 2.

   [10 marks]

3. Find an orthogonal matrix P such that $Px = y$ where $x = (1, 2, 3, 1, 1)'$ and $y = (a, 0, 0, 0, 0)$ for some number a. If this is impossible, then prove it.

   [10 marks]

4. Let A be a real symmetric matrix. Let S be the largest subspace such that for every nonzero $x \in S$ we have $x'Ax > 0$. Show that the dimension of S must be the number of positive eigenvalues of A.

   [10 marks]

5. A complex matrix $A_{9\times 9}$ is such that every 3-dimensional subspace of $\mathbb{C}^9$ is an eigenspace of A. Find the characteristic polynomial of A as explicitly as possible. Also find the Jordan Canonical Form of A. [5+5 marks]

6. Show that every real, symmetric matrix can be expressed as the difference of two negative definite matrices.

   [10 marks]

7. Let V be a vector space and S be a subspace of V. Let $W = u + S$ where $u \in V$ is not in S. Prove that for any vector x there is a unique $y \in W$ such that
   $$\forall w \in W \ \ \|x - y\| \leq \|x - w\|.$$
   Let the orthogonal projector onto S be given by the matrix P. Express this unique y in terms of x, u and P.

   [5+5 marks]

8. Let A be an $n \times n$ matrix and let

$$\alpha_i = \sum_{j \neq i}^{n} |a_{ij}|$$

for $i = 1, ..., n$. If λ is an eigenvalue of A, show that $|\lambda - a_{ii}| \leq \alpha_i$ for at least one i.

[10 marks]

9. Let S be a positive definite matrix. Show that A is an orthogonal projector w.r.t. the inner product $\langle x, y \rangle = x'Sy$ if and only if A is idempotent and $A^* S(I - A) = O$.

[10 marks]

10. Let $\|\cdot\|$ be a norm that satisfies the following condition:

$$\forall x, y \quad \|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Show that this norm must be induced by an inner product.

[10 marks]

INDIAN STATISTICAL INSTITUTE

Back Paper Examination : Semester II (2016-2017)

B. Stat 1st Year

Statistical Methods - II

Date: 12. 7. 2017 | Maximum marks: 100 | Duration: 3 hours.

1. In the multiple regression model with an intercept, assume that there are m independent variables which have mean 0.

   (a) Derive expressions for the means and variances of the (i) least-square estimates of the regression coefficients, (ii) the fitted values and (iii) the residuals. [10]

   (b) Show that the fitted values and the residuals are uncorrelated. [5]

   (c) Derive the expression of different sum of squares in the ANOVA decomposition and find their expected values. [20]

   (d) Assuming further that the independent variables are uncorrelated, derive the relationship between (i) the least-square estimates of the regression coefficients and (ii) the sum of squares due to regression with the corresponding quantities in the m simple linear regressions with one independent variable taken at a time. [20]

2. Derive the expression for the multiple correlation coefficient. Show that this correlation coefficient is necessarily between 0 and 1. [7 + 5]

3. Present a method to generate a value of the random variable X which has the following distribution:

$$P\{X = j\} = \left(\frac{1}{2}\right)^{j+1} + \frac{\left(\frac{1}{2}\right) 3^{j-1}}{4^j}, \quad j = 1, 2, \ldots$$

[13]

4. Suppose we want to generate a random number from the standard normal distribution.

   Give an algorithm based on the rejection method for this problem. Choose an appropriate value of the parameter for the distribution that you choose for this purpose so that the efficiency of this algorithm is maximized. [12 + 8]

INDIAN STATISTICAL INSTITUTE

Second Semester Examinations: 2016-17 (Backpaper)

B. Stat. I Year Analysis-II

Date: 14/7/2017 Maximum Marks: 100 Duration: 3 Hours

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

---

(1) Give an example of a non-negative continuous function $f : [1, \infty) \to \mathbb{R}$ with the property that $\sum_{n=1}^{\infty} f(n)$ diverges but $\int_0^\infty f(x)\, dx$ converges. [5]

(2) Consider the function $f(x) = \sin x$, $x \in [0, \pi/2]$. Let $P_n = \{0, \frac{\pi}{2n}, \frac{2\pi}{2n}, \frac{3\pi}{2n}, \ldots, \frac{n\pi}{2n}\}$, $n \in \mathbb{N}$.

(a) Show that $U(f, P_n) = \frac{\pi/4n}{\sin(\pi/4n)}\left[\cos\frac{\pi}{4n} - \cos(\frac{\pi}{2} + \frac{\pi}{4n})\right]$.

(b) Prove that $\lim_{n\to\infty} U(f, P_n) = 1$. [10]

[Hint. $2\sin(x/2)(\sin x + \sin 2x + \cdots + \sin nx) = \cos(x/2) - \cos((2n+1)x/2)$.]

(3) Find the limit

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{n}{n^2 + k^2}$$

by interpreting the sum as a Riemann sum of an appropriate function. [10]

(4) For $n \in \mathbb{N}$, let $f_n(x) = \frac{1}{1+(nx-1)^2}$, $x \in [0, 1]$. Find the function f such that $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in [0, 1]$. Does $\{f_n\}$ converge to f uniformly on $[0, 1]$? [10]

(5) Let $A = \{x \in \mathbb{R} : 1/2 \le |x| \le 2$. Does the series $\sum_{n=1}^{\infty} \frac{n^2}{\sqrt{n!}}(x^n + x^{-n})$ converge uniformly on the set A? [10]

(6) Find the radius of convergence of the power series $\sum_{n=1}^{\infty} \frac{(x-1)^{2n}}{2^n n^3}$. Is the power series convergent at the endpoints of the interval of convergence? [10]

(7) Let R_1 and R_2 be the radii of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$. Let R be the radius of convergence of $\sum_{n=0}^{\infty}(a_n + b_n)x^n$. Show that $R = \min\{R_1, R_2\}$ if $R_1 \ne R_2$. What can you say about R if $R_1 = R_2$? [10]

(8) Show that the trigonometric series $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \cos nx$ is convergent for each $x \in \mathbb{R}$, but is not the Fourier series of any Riemann integrable function. [10]

(9) Consider the 2π-periodic odd function defined on $[0, \pi]$ by $f(x) = x(\pi - x)$.

(a) Compute the Fourier coefficients of f.

(b) Show that

$$f(x) = \frac{8}{\pi} \sum_{k\ odd \ge 1} \frac{\sin kx}{k^3}.$$

(c) Use Parseval's identity to prove that [15]

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^6} = \frac{\pi^6}{960} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^4}{945}.$$

(10) Show that the double series $\sum_{(m,n)\in\mathbb{N}\times\mathbb{N}} \frac{1}{m^p n^q}$ is convergent if and only if $p > 1$ and $q > 1$. [10]