CHENNAI MATHEMATICAL INSTITUTE

M.Sc. Data Science Entrance Examination 2025

IMPORTANT INSTRUCTIONS

- Part (A) consists of multiple-choice questions. There may be multiple correct choices. You have to select all the correct options and no incorrect option to get full marks. **There is no partial credit.**
- For questions in part (A), you have to provide the answers on the computer. You only have to choose the appropriate answer(s) from the choices provided. If the answer is parts (a) and (c), choose only (a) and (c).
- **Important Note**: Please ignore Page 3 of your answer booklet that says *Answers to part A*. There is no need to write your answers there. Only answers entered in the computer will be considered.
- For questions in part (B), you have to write your answer with a short explanation in the space provided for the question.
- **Part(A) will be used for screening.** Part (B) will be graded only if you score a certain minimum in part (A). However your scores in both parts will be used while making the final decision.
- For numerical answers, the following forms are acceptable: fractions, decimals, symbolic e.g.: $\binom{n}{r}$, ${}^nP_r$, $n!$ etc.

Notation and terminology

- A function f from a set A to a set B is said to be **injective (or one-to-one)** if $f(x) = f(y)$ implies $x = y$ for all $x, y \in A$;
- f is said to be **surjective (or onto)** for every $y \in B$ there exists $x \in A$ such that $f(x) = y$;
- f is said to be **bijective** if it is both injective and surjective;
- f is said to be **invertible** if there exists a function g from B to A such that $f(g(y)) = y$ for all $y \in B$ and $g(f(x)) = x$ for all $x \in A$.
- For a matrix A, A^T denotes the transpose of A. For a square matrix A, $|A|$ denotes the determinant of A and trace(A) denotes the *trace* of A — namely the sum of the diagonal elements of A.
- A *diagonal matrix* is a square matrix D with all off-diagonal entries equal to zero i.e. $d_{ij} = 0$ for all $i \neq j$.
- An *upper triangular matrix* is a square matrix A for which all entries below the diagonal are zero, i.e. $a_{ij} = 0$ for $i > j$.
- A *symmetric matrix* is a square matrix S for which $s_{ij} = s_{ji}$ for all $i \neq j$.

Part (A) - Multiple-choice questions

1. Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c, d\}$ be two finite sets. Which of the following statements regarding functions from $X \to Y$ is/are true?

   (a) The number of one-to-one functions is 3^4.

   (b) There are as many one-to-one functions as onto functions.

   (c) The number of onto functions is strictly less than the number of one-to-one functions.

   (d) The number of one-to-one functions is 24.

2. The two arguments to the function `foo(A, n)` in the code below are: (i) an integer array `A` indexed from `0`, and (ii) the number `n` of elements in `A`.

```
function foo(A, n) {
     count = 0;
     for i from 0 to (n-1) {
          for j from (i+1) to (n-1) {
               if (A[i] > 2 * A[j]) {
                    count = count + 1;
               }
          }
     }
     return(count);
}
```

   Which of the following statements about the function `foo(A, n)` are correct?

   (a) `foo(A, n)` counts the number of index pairs (i, j) such that $i < j$ and $\texttt{A}[i] > 2 \times \texttt{A}[j]$.

   (b) For the input `A = [1, -4, 3, -5, -2], n = 5`, the function returns 7.

   (c) For the input `A = [1, 2, 3, 4, 5], n = 5`, the function returns 0.

   (d) For the input `A = [10, 5, 1], n = 3`, the function returns 2.

3. What is the domain of the following real valued function?

$$f(x) = \log_2(x^2 - 5x + 6).$$

   (a) $(-\infty, 2)$

   (b) $(3, \infty)$

   (c) $(-\infty, 2) \cup (3, \infty)$

   (d) $(-\infty, \infty)$

4. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real matrix which satisfies $A^{-1} = A$. Which of the following statements is/are *always* true?

   (a) $a + d = 0$

   (b) $ad - bc = 1$

   (c) $ad - bc \neq 0$

   (d) $a^2 + 2bc + d^2 = 2$

5. Let $\mathbb{Z}$ be the set of all integers. Let $A = \{n \in \mathbb{Z} \mid n^2 + 10n + 21 \text{ is divisible by } 7\}$. Which of the following statements is/are true?

   (a) $A = \{n \in \mathbb{Z} \mid (n \equiv 0 \mod 7)\}$

   (b) $A = \{n \in \mathbb{Z} \mid (n \equiv 0 \mod 7) \text{ or } (n \equiv 4 \mod 7)\}$

   (c) $A = \{n \in \mathbb{Z} \mid (n \equiv 1 \mod 7) \text{ or } (n \equiv 5 \mod 7)\}$

(d) $A = \{n \in \mathbb{Z} \mid (n \equiv 2 \mod 7) \text{ or } (n \equiv 6 \mod 7)\}$

6. Let $n \geq 3$ be an integer, and let $x_1, x_2, \ldots, x_n$ be variables which take real values with $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$. Let

$$\begin{aligned} A &= x_1 + x_2 + \ldots + x_n \\ B &= x_1x_2 + x_2x_3 + \ldots + x_{n-1}x_n + x_nx_1 \end{aligned}$$

Which of the following statements is/are true.

(a) $A \geq B$ is always true.
(b) $B > A$ is true for some values of the x_i's and $A > B$ is true for some values of the x_i's.
(c) $A = B$ has a finite number of solutions
(d) $A = B$ has an infinite number of solutions.

7. Let $B = ((b_{i,j}))$ be an $n \times n$ matrix. Let $p : \{1, 2, \ldots, n\} \mapsto \{1, 2, \ldots, n\}$ be a bijection (i.e. a one-to-one correspondence) and let a matrix $A = ((a_{i,j}))$ be defined by

$$a_{i,j} = b_{p(i),p(j)}, \quad 1 \leq i, j \leq n.$$

Which of the following statement(s) is/are true for all choices of B and P.

(a) A admits an inverse if and only if B admits an inverse.
(b) For any $x, y \in \mathbb{R}^n$, $Ax = y$ admits a solution if and only if $Bx = y$ admits a solution.
(c) A and B have the same trace.
(d) A and B have the same eigenvectors.

8. Let x be a variable that takes real values, and let $f(x) = x^3 - 3x$. Which of the following statements is/are true?

(a) $f(x)$ has a local maximum at $x = -\sqrt{3}$
(b) $f(x)$ has a local maximum at $x = -1$
(c) $f(x)$ has a local minimum at $x = \sqrt{3}$
(d) $f(x)$ has a global minimum at $x = 1$

9. A binary relation R defined on a set S is said to be **antisymmetric** if for $x, y \in S$, xRy **and** $yRx \implies x = y$. Let R_1, R_2 be two binary relations defined on a set S. The **union** of R_1, R_2 is the binary relation U defined on S as: For $x, y \in S$, $xUy \iff xR_1y$ **or** xR_2y. The **intersection** of R_1, R_2 is the binary relation I defined on S as: For $x, y \in S$, $xIy \iff xR_1y$ **and** xR_2y.

Which of the following statements is/are true?

(a) A binary relation cannot be both symmetric and antisymmetric.
(b) A binary relation can be both transitive and antisymmetric.
(c) The union of two equivalence relations is always an equivalence relation.
(d) The intersection of two equivalence relations is always an equivalence relation.

10. In how many ways can 10 identical chocolate bars be distributed among 5 children, in such a way that each child gets at least one chocolate bar?

(a) 50
(b) 126
(c) 252
(d) 3125

11. Which of the following is/are logically equivalent to $\neg(P \implies Q)$?

(a) $\neg P \vee Q$
(b) $\neg P \wedge Q$
(c) $Q \implies P$
(d) $P \wedge \neg Q$

12. Let $f(x) = \sqrt{x}$. We draw a tangent to the curve $y = f(x)$ at the point on the curve whose x coordinate is equal to 4. Where does this tangent intersect the X-axis?

(a) $x = 4$
(b) $x = -2$
(c) $x = -4$
(d) $x = 2$

13. Let n be an integer, $n \geq 4$. A is an $n \times n$ matrix with real entries. The matrix B is obtained by the following sequence of operations on A. First, multiply each entry of A by 2. Then add 3 times the second column to the third column. Finally, swap the first and the fourth columns. If $\det(A) = 5$, which of the following statements are true?

(a) 10 divides $\det(B)$
(b) $\det(B) = -5$
(c) 100 divides $\det(B)$
(d) $\det(B) = -2^n \cdot 5$

14. Let x, y, z be positive numbers such that $x^2 + y^2 = z^2$. Determine the value of the following expression:

$$\frac{\log_{y+z} x + \log_{z-y} x}{(\log_{y+z} x)(\log_{z-y} x)}.$$

(a) Undefined.
(b) $\frac{1}{2}$
(c) $\frac{1}{4}$
(d) 2

15. A game being offered in a casino consists of guessing the outcomes of two tosses of a fair coin. The gambler wins if she/he has correctly guessed at least one of the two tosses. To play a game, the gambler has to pay a fee of Rs. 80, and the winner gets a reward of Rs. 100 on winning the game (and nothing otherwise).

Which of the following statements are correct?

(a) In the first 10 minutes on a given day exactly three gamblers play the game, one after the other. The probability that the casino owner makes a profit in the first 10 minutes equals 1/4.
(b) One gambler plays the game three times. The probability that she wins exactly two of the three games is 27/64.
(c) Three friends go together and play the game with each playing once. The probability that all three win equals 27/64.
(d) If 1200 players play the game on a given day, the expected profit of the casino owner for the day equals Rs. 6000.

16. In the following code the operator `%` denotes the remainder after integer division. That is: for positive integers `a,b` the value `a % b` is the remainder obtained when `a` is divided by `b`.

```
function fizzbuzz(n) {
   count = 0;
   for i from 0 to (n-1) {
       if ((i % 3) == 0) and ((i % 5) != 0) {
           count = count + 1;
     }
   }
   return(count);
}
```

What does `fizzbuzz(100)` return?

(a) 27
(b) 33
(c) 45
(d) 60

17. Alpha and Beta are inhabitants of an island of knights and knaves, where knights always tell the truth and knaves always lie. Alpha and Beta are alone at a beach when Alpha says: "At least one of us is a knave." And Beta says: "We are both knaves." Which of the following is/are true?

(a) Alpha and Beta are both knights
(b) Alpha and Beta are both knaves
(c) Alpha is a knight and Beta is a knave
(d) Alpha is a knave and Beta is a knight

Questions 18 to 20 are based on the following code.

The two arguments to the function `Mystery(A, n)` in the code below are: (i) an integer array `A` indexed from `0`, and (ii) the number `n` of elements in `A`. Each element of `A` is an integer from the set $\{1, 2, \ldots, n\}$. The expression `[0] * (n+1)` creates an array, indexed from `0`, that contains $n + 1$ zeroes.

```
function Mystery(A, n) {
    found = False;
    value = None;
    B = [0] * (n+1);

    for i from 1 to n {
        B[A[i]] = B[A[i]] + 1;
    }

    for i from 1 to n {
        if (found == False) {
            if (B[A[i]] == 1) {
                found = True;
                value = A[i];
            }
        }
    }

    if (found == True) {
        return(value);
    } else {
        return(None);
    }
}
```

Answer the next three questions about this function.

18. What does the function call `Mystery([1, 2, 3, 3, 2, 1], 6)` return?
    (a) `None`
    (b) `1`
    (c) `2`
    (d) `3`

19. What does the function call `Mystery([1, 2, 4, 3, 2, 1], 6)` return?
    (a) `None`
    (b) `2`
    (c) `3`
    (d) `4`

20. What does the function call `Mystery([6, 5, 4, 3, 2, 5], 6)` return?
    (a) `3`
    (b) `4`
    (c) `5`
    (d) `6`

Part (B) - Short-answer questions

For questions in part (B), you have to write your answer with a short explanation in the space provided for the question in your answer sheet. If you need more space, you may continue on the pages provided for rough work. Any such overflows must be clearly labeled.

1. A girl writes five consecutive positive integers on a blackboard. She then erases one of them. The sum of the remaining four numbers is 2025. What number did she erase?

2. A toy company currently sells 1,000 toys each month at a price of ₹500 per toy. To increase their sales, the company is considering to lower the price. Market research shows that for every ₹10 decrease in the price, the number of toys sold increases by 100. However, the price cut applies to all the toys sold.

   (a) By how much should the company reduce the price, so as to maximize its monthly revenue?

   (b) What would be the maximum revenue per month that the company can achieve?

   Assume that the decrease in price is an integer multiple of ₹1.

3. 15 balls are placed independently and uniformly at random into 15 bins numbered from 1 to 15. The probability that a ball ends up in a particular bin is $\frac{1}{15}$.

   (a) What is the probability that the first 5 balls go into different bins, conditioned on the event that first four balls are in different bins?

   (b) What are the possible values of the expected number of balls in bin 1 conditioned on the event that the first 13 balls are in different bins?

4. There are 18 chocolates in a bag, of which 7 are green, 6 are blue, and 5 are red. We pick chocolates one at a time from the bag without replacement.

   (a) What is the probability that the first and the third chocolate are green?

   (b) What is the probability that after picking twelve chocolates, only chocolates of one colour remain in the bag?

5. Let M be an $n \times n$ matrix. We define an elementary row operation on M to be one of the following:

   i Interchanging some two rows of M.

   ii Multiplying a row in M by a non-zero scalar.

   iii Adding a scalar multiple of one row of M to another row of M.

   An $n \times n$ matrix is said to be elementary if it is the result of a single elementary row operation performed on the $n \times n$ identity matrix.

   Which of the following statements is/are true? Justify your answer with a short proof, if the statement is true, otherwise, provide a counterexample.

   a Every elementary operation on a $n \times n$ matrix A can be performed by multiplying A by an elementary $n \times n$ matrix on the right.

   b An elementary row operation on a $n \times n$ matrix A results in a matrix with the same determinant as that of A.

6. Consider the following polynomial of positive degree n

$$P_n(x) = 1 + 2x + 3x^2 + \ldots + (n+1)x^n.$$

   Show that there is no real number r such that $P_n(r) = 0$ when n is even.

7. How many positive integers less than 1000 are **neither** divisible by 3 **nor** divisible by 5? Explain how you arrived at your answer.

8. Three hostel friends Amar, Prem and Raj are suspected of breaking a window. They made the following statements when questioned by the warden:

   - Amar: *I did not break it. Prem is lying.*
   - Prem: *Amar is telling the truth. Raj broke the window.*
   - Raj: *I did not break it. Either Amar is telling the truth or Prem is telling the truth.*

   You know that exactly one of them lied and the other two told the truth. Then, who broke the window? Justify.

9. A cloth bag labeled X contains two apples, bag Y contains two oranges and bag Z one apple and one orange. You pick a bag at random and then remove one fruit from that bag at random. Suppose you removed an apple. What is the probability that the fruit remaining in the bag is also an apple? Justify.

Questions 10 and 11 are based on the following description.

The following question appeared in a quiz:

> "*Write the pseudocode for a function `Closest(A,n,x)` that takes an array `A`, a positive integer `n`, and an integer `x` as arguments. The elements of `A` are all integers less than 2^{64}, and `n` is the number of elements in `A`. The call `Closest(A,n,x)` should return an integer `y` such that: (i) $y \neq x$, (ii) `y` is present in `A`, and (iii) there is* no *$z \neq x$ in A where $|z - x| < |y - x|$ holds. If `A` has no such element `y`, then the function should return the special value `None`.*"

A student submitted the code below as the answer to this question. In the code the array `A` is indexed from `0`, and `MAXINT` $= 2^{64} - 1$. The call `abs(z)` returns the absolute value |`z`| of integer `z`.

```
function Closest(A, n, x) {
    minVal = MAXINT;

    for i from 0 to (n-1) {
        absDiff = abs(x - A[i]);
        if (absDiff < minVal) {
            minVal = absDiff;
            y = A[i];
        }
    }

    if (minVal != MAXINT) {
        return(y);
    } else {
        return(None);
    }
}
```

This answer turned out to be wrong; this function gives the correct answer for some valid inputs, and wrong answers for other valid inputs. Answer the next two questions about this function.

10. What do the following function calls return?

    (a) `Closest([-10,2,10], 3, 8)`

    (b) `Closest([0,-5,4], 3, 7)`

11. Give one example of (i) an input array `A` with exactly 3 elements and (ii) an integer `x` for which the call `Closest(A, 3, x)` returns a **wrong** answer. What is this wrong answer? What is the correct answer?

---

12. A sequence of five natural numbers $s_1 \leq s_2 \leq s_3 \leq s_4 \leq s_5$ satisfy the following conditions:

- $\sum_{i=1}^{5} s_i = 35$
- $\sum_{i=1}^{3} s_i = 15$
- $\sum_{i=3}^{5} s_i = 27$
- s_2 is even
- $s_4 - s_2 = 2$

Find all such sequences that satisfy the above conditions.

13. Five executives of a company namely CEO (chief executive officer), CFO (chief financial officer), COO (chief operating officer), CTO (chief technology officer), CMO (chief marketing officer) are to be seated around a circular table.

    - The CEO must sit next to the CFO.
    - The COO must not sit next to the CTO.
    - The CTO must not sit next to the CEO.

    In how many distinct ways can they be seated? (Rotations of the same arrangement are considered the same).

14. Evaluate the following limit:

$$\lim_{x\to 0} \frac{e^{2x} - 2x - \cos x}{x \sin x}.$$

15. Find the determinant of the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 4 & 3 \\ 5 & 8 & 26 & 21 \\ 2 & 3 & 10 & 8 \\ 3 & 5 & 16 & 13 \end{bmatrix}$$

16. Let $f : \mathbb{R} \to \mathbb{R}$ be a function which is twice differentiable everywhere and suppose that $f'(q) = 0$ for every rational number q. Find the value of $f(\pi^2) - f(\pi)$.

17. Consider the following 2×2 matrix

$$B = \begin{bmatrix} 0 & r \\ -1 & 0 \end{bmatrix},$$

where r is a nonzero real number. Find B^{2025}, i.e., the matrix obtained by multiplying B with itself 2025 times.

Instructions for Questions 18,19, and 20:

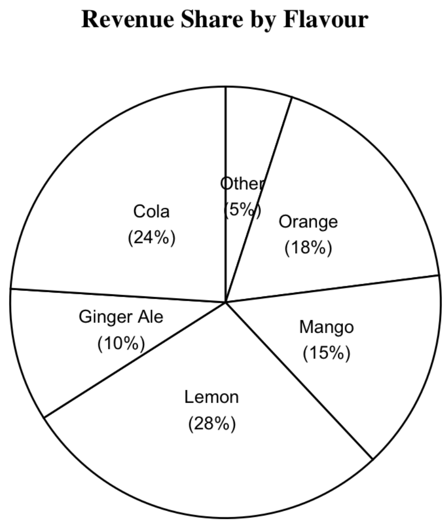
Read the following description carefully and answer the questions that follow. Use all the information provided. Clearly show all your calculations.

Description:

NorthCool Beverages Pvt. Ltd. is a leading beverage company that operates across six northern states of India. During the past summer the company launched an aggressive marketing campaign to promote its range of flavoured drinks. The charts below summarise the sales data collected during this campaign.

18. What is the revenue, in Lakhs, from the sales of the mango flavoured drink?

19. NorthCool Beverages plans to launch a new variant, "Lemon Max", in the two states with the highest sales of the lemon flavour. The company expects Lemon Max to generate additional revenue equal to 20% of the current lemon flavour sales in those two states. What would be the percentage revenue share of lemon flavour drinks if their expectations are met?



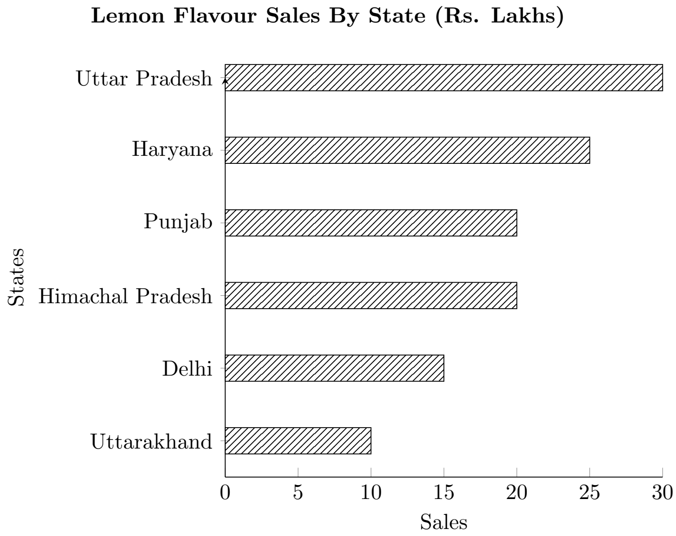


Figure 1: Revenue share by flavour, and sales of lemon-flavoured drink by state

20. On average, for every ₹1 lakh in total revenue from the lemon-flavoured drink, 10,000 litres of lemon-flavoured drink are sold across the six states. Estimate the total volume (in litres) of lemon-flavoured drink sold by NorthCool Beverages across all six states.