CHENNAI MATHEMATICAL INSTITUTE

M.Sc. Data Science Entrance Examination

15^{th} May 2019

Enter your *Admit Card Number*: **D** – ☐☐☐ – ☐☐☐☐

IMPORTANT INSTRUCTIONS

- This booklet has 9 printed sheets, including this cover page. Sheet 2 (page 1) contains space to answer part A, sheets 3 to 5 contain *eighteen* questions in part A and sheets 6 to 9 contain *eighteen* questions in part B.
  - For questions in part (A), you only have to write your answer on the designated page on the appropriate line. For example, if the answer is parts (a) and (c), write only (a) and (c) on the line and if the answer is 2781, write this number on the line.
  - For questions in part (B), you have to write your answer with short explanation in the space provided below the question.
  - For numerical answers, the following forms are acceptable: fractions, decimals, symbolic e.g.: $\binom{n}{r}$, ${}^nP_r$, $n!$ etc.

  For rough work use the blank pages at the end.

- **Time allowed is 3 hours. Total points: 90** = 36 for part A + 54 for part B.

- **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 office use only

	Points	Remarks
Part A		
Part B		
Total		

B1	B2	B3	B4	B5	B6	B7	B8	B9	Total
B10	B11	B12	B13	B14	B15	B16	B17	B18	Total

Part A - Answers

This is the only place that will be seen for grading part A. So carefully and clearly write the answers to each question on the designated line below. Write only the final answers, do not show any intermediate work. Illegible/unclear answers will not be considered.

A1.____________________

A2.____________________

A3.____________________

A4.____________________

A5.____________________

A6.____________________

A7.____________________

A8.____________________

A9.____________________

A10. ____________________

A11. ____________________

A12. ____________________

A13. ____________________

A14. ____________________

A15. ____________________

A16. ____________________

A17. ____________________

A18. ____________________

Notation

- 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)** if 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 $x \in A$ and then g is said to be inverse of f and is denoted by f^{-1}.
- For a matrix A, $|A|$ denotes the determinant of A and A^T denotes the transpose of A.

Part A - Questions

This section consists of some questions requiring a single answer and some multiple choice questions. For questions that ask you to calculate a value, you will be assessed based on the answer you provide. No explanation is required. In 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. Write the correct options / answer in the space provided on page 1. Only page 1 will be seen for grading part A.

1. Let $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2\}$. The number of surjective functions from X to Y equals
   (a) 2^n
   (b) $2^n - 1$
   (c) $2^n - 2$
   (d) $2^{n/2}$

2. If $P(A \cup B) = 0.7$ and $P(A \cup B^c) = 0.9$ then find $P(A)$.

3. Which of the following statements are true for all $n \times n$ matrices A, B:
   (a) $(A^T)^T = A$;
   (b) $|A^T| = |A|$;
   (c) $(AB)^T = A^T B^T$;
   (d) $(A + B)^T = A^T + B^T$.

4. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 5 & 5 \\ 0 & 10 & 10 \\ 0 & 0 & 15 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 0 & 0 \\ 3 & 6 & 0 \\ 3 & 6 & 9 \end{bmatrix}$. Which of the following statements are true?

   (a) $|A| = |B|$;
   (b) $|B| = 125|A|$;
   (c) $|C| = 27|A|$;
   (d) $|C| = \dfrac{|A|}{3}$.

5. Consider the polynomials $p(x) = (5x^2 + 6x + 1)(x + 1)(2x + 3)$ and $q(x) = (5x^2 - 9x - 2)(2x^2 + 5x + 3)$. The set of common divisors of $p(x)$ and $q(x)$ is:

   (a) $\{2x + 3,\ x + 1,\ 5x + 1\}$
   (b) $\{2x + 3,\ x - 1,\ 5x + 1\}$
   (c) $\{x + 3,\ 2x + 1,\ x - 2\}$
   (d) $\{2x - 3,\ x + 1,\ 5x + 1\}$

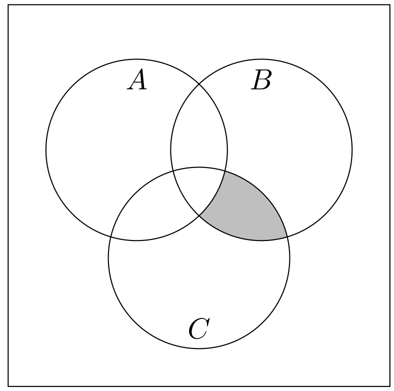
6. Let R denote the set of real numbers and let $A = \{x \in R : x \neq 3\}$. For $x \in A$, let $f(x) = \frac{2x+1}{x-3}$. Let B denote the range of f. Then

   (a) $B = \{x \in R : x \neq -2\}$ and $f^{-1}(x) = \frac{3x-1}{x+2}$;
   (b) $B = \{x \in R : x \neq 2\}$ and $f^{-1}(x) = \frac{3x+1}{x-2}$;
   (c) $B = \{x \in R : x \neq 2\}$ and $f^{-1}(x) = \frac{3x-1}{x-2}$;
   (d) $f^{-1}(x)$ does not exist because f is not injective.

7. We need to choose a team of 11 from a pool of 15 players and also select a captain. The number of different ways this can be done is:

   (a) $\binom{15}{11}$
   (b) $11 \cdot \binom{15}{11}$
   (c) $15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5$
   (d) $(15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5) \cdot 11$

8. Consider the following Venn diagram. The universal set U is the set of all natural numbers from 1 to 1000.



The sets A, B, C contain integers in U that are multiples of $6, 7, 8$ respectively. The number of elements in the shaded region is:

(a) 12
(b) 15
(c) 16
(d) 17

9. In the code fragment below, `start` and `end` are integer values and `prime(x)` is a function that returns `True` if `x` is a prime number and `False` otherwise.

```
i = 0;
j = 0;
k = 0;
for m = start to end {
  if prime(m) == True {
     i = i + 1;
     # Statement 1
  }else{
     j = j - 1 ;
     # Statement 2
  }
}
```

We wish to maintain the invariant `k == i - j` after each iteration of the `for` loop. What should we insert at `Statement 1` and `Statement 2`?

(a) Statement 1: `k = k + 1` and Statement 2: `k = k + 1`
(b) Statement 1: `k = k + 1` and Statement 2: `k = k - 1`
(c) Statement 1: `k = k - 1` and Statement 2: `k = k + 1`
(d) Statement 1: `k = k - 1` and Statement 2: `k = k - 1`

Description for the following two questions:

The following table gives the budget allocation (in Rupees Crores) to 5 different departments and their quarterly expenditure for a particular year.

Department	Allocation	Q1	Q2	Q3	Q4
D1	240	25	60	70	45
D2	120	10	40	30	20
D3	200	30	40	60	20
D4	125	20	35	20	25
D5	180	40	20	60	30

10. As a percentage of total allocation, the maximum quarterly expenditure (in any quarter) was shown by

    (a) D2
    (b) D5
    (c) D1
    (d) D3

11. Looking at the combined expenditure of all the five departments, which of the following gives the correct order as far as expenditure in each quarter is concerned?

    (a) $Q4 < Q1 < Q3 < Q2$
    (b) $Q1 < Q4 < Q2 < Q3$
    (c) $Q4 < Q1 < Q2 < Q3$
    (d) $Q1 < Q4 < Q3 < Q2$.

12. Three boxes are presented to you. At most one of them contains some gold. Each box has printed on it a clue about its contents.The clues are:

    (Box 1) *The gold is not here.*
    (Box 2) *The gold is not here.*
    (Box 3) *The gold is in Box 1.*

    Only one clue is true; the other two are false. Which box has the gold?

    (a) Box 1
    (b) Box 2
    (c) Box 3
    (d) None of them has the gold

13. Abha and Vibha both have white and yellow handkerchieves. To distinguish them, their mother has marked Abha's handkerchieves with the letter A and Vibha's handkerchieves with the letter V. There are 8 white handkerchieves of which 3 belong to Abha, and 11 yellow handkerchieves

of which 4 belong to Abha. All 19 handkerchieves were packed together in a bag for a trip. Their mother pulled out a handkerchief from the bag without looking and found it was marked V. What is the probability that the handkerchief was yellow?

(a) 5/12
(b) 7/19
(c) 7/12
(d) 11/19

14. We need to choose a team of 11 from a pool of 15 players and also select a captain. The number of different ways this can be done is:

(a) $\binom{15}{11}$
(b) $11 \cdot \binom{15}{11}$
(c) $15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5$
(d) $(15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5) \cdot 11$

15. The sum of the diagonal elements of a matrix A is called the *trace* of A and is denoted by $\mathrm{tr}(A)$. Which of the following statements about the trace are true?

(a) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$;
(b) $\mathrm{tr}(2A) = 2\mathrm{tr}(A)$;
(c) $\mathrm{tr}(A^T) = \mathrm{tr}(A)$;
(d) $\mathrm{tr}(A^{-1}) = \mathrm{tr}(A)$.

16. An *upper triangular* matrix is a square matrix with all entries below the diagonal being zero. Suppose A and B are upper triangular matrices. Which of the following statements are true?

(a) The matrix $A + B$ is upper triangular.
(b) The matrix A^T is upper triangular.
(c) The matrix A^{-1} is upper triangular.
(d) The matrix AB is upper triangular.

Description for the following 2 questions: If Z is a continuous random variable which follows a Gaussian distribution with mean = 0 and standard deviation = 1, then

$$\mathbb{P}(Z \leq a) = \int_{-\infty}^{a} \frac{\exp\{-z^2/2\}}{\sqrt{2\pi}} dz = \Phi(a),$$

where $\Phi(a = -2) = 0.02$, $\Phi(a = -1.5) = 0.067$, $\Phi(a = -1) = 0.16$, $\Phi(a = -0.5) = 0.31$, $\Phi(a = 0) = 0.50$, $\Phi(a = 0.5) = 0.69$, $\Phi(a = 1) = 0.84$, $\Phi(a = 1.5) = 0.933$, and $\Phi(a = 2) = 0.98$. Note that $X = \mu + \sigma Z$ follows Gaussian distribution with mean μ and standard deviation σ.

Suppose the score distribution of an exam has a Gaussian distribution with mean μ and standard deviation σ. A candidate fails if she/he obtains less than 35% marks. However she/he must

Result	Percentage of Students
Passed with distinction	2
Passed without distinction	82

Table 1: Result of a particular exam taken by 225 students.

obtain more that 80% marks in order to pass with distinction. The exam is taken by a group of 225 students and the results are given in Table 1..

You may use the fact that the sample mean $\bar{X}$ follows a Gaussian distribution with mean μ and standard deviation $\sigma/\sqrt{n}$.

17. Which of the following statements are true?

   (a) $\mu - \sigma = 35$ and $\mu + 2\sigma = 82$;

   (b) $\mu - 2\sigma = 16$ and $\mu + \sigma = 82$;

   (c) $\mu = 53.33$ and $\sigma = 13.33$;

   (d) $\mu = 50$ and $\sigma = 15$.

18. Which of the following statements are true?

   (a) The probability that the average score of the group of 225 students is less than 48 is 2%.

   (b) The probability that the average score of the group of 225 students is less than 42.5 is 6.7%.

   (c) The probability that the average score of the group of 225 students is greater than 49 is more than 50%.

   (d) The probability that the average score of the group of 225 students is greater than 57.5 is more than 16%.

Part B

For questions in part (B), you have to write your answer with a short explanation in the space provided below the question. For numerical answers, the following forms are acceptable: fractions, decimals, symbolic e.g.: $\binom{n}{r}$, ${}^{n}P_{r}$, $n!$ *etc.*

1. For positive numbers a, b, c, show that

$$\frac{a}{b} + \frac{b}{c} + \frac{c}{a} \geq 3.$$

Description for the following question:
Suppose X is the number of successes out of n trials, where the trials are independent of each other. The probability of success at every trial is p. The probability that there will be exactly k successes out of n trials is

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

The expected number of successes is $\mathbb{E}(X) = np$. If $n \to \infty$ and $p \to 0$

$$\binom{n}{k} p^k (1-p)^{n-k} \to e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots,$$

where $\lambda = np$.

2. Each time a gambler plays a game, there is a one-in-million chance of winning. The gambler plays the game one million times. Find the probability of winning the game zero times, i.e. find the probability of the event that the gambler will lose all one million times that she/he will try. Note : 1 million = 10^6

3. Suppose X is a continuous distribution with probability density function

$$f(x) = k(x - x^2), \quad 0 \leq x \leq 1,$$

where k is the normalizing constant. Find the value of k and the expected value of the distribution.

4. Four friends attending an opera leave their coats at the checkroom. When they return each one is handed a coat that does not belong to her. In how many ways can this happen?

5. If the lines $2x + y = 2, -5x + 3y = 4$ and $ax + by = 1$ are concurrent then prove that the line $9x - 10y = 1$ passes through (a, b).

6. Let $\theta = \log_e(2)$. For $-\theta \leq x \leq \theta$, let $f(x) = \frac{\exp(x)}{1+\exp(x)}$. Compute α and β where

$$\alpha = \max_{-\theta \leq x \leq \theta} f(x) \quad \text{and} \quad \beta = \min_{-\theta \leq x \leq \theta} f(x).$$

Justify your answers.

7. Ani is training for the olympics with Usain Bolt. After a few days of training Usain challenges Ani to catch him. Usain sets off running very slowly with a view to encourage Ani. He covers 70m the first minute, 100m the next minute, then 130m the minute afterwards and so on. Ani is told to start 3 minutes later. Having trained hard he runs 100m the first minute, 150 m the second Minute, then 200m and so on. Ani catches Usain at an integral multiple of a minute. How many minutes did Ani run before catching up with Usain. What were their respective speeds?

8. A small circular fire is spreading with its radius increasing at the rate of 1.5 metres per minute. When the radius of the fire is 5 metres, how fast is the burned area growing?

9. Show that among any set of 7 distinct integers there must exist 2 integers whose sum or difference is divisible by 10.

10. Let $p(x)$ be a polynomial with integer coefficients. Let n be a positive integer and suppose a and b are two integers such that $a \equiv b(\text{mod } n)$. Is it true that $p(a) \equiv p(b)(\text{mod } n)$? Justify your answer.

11. A thin piece of metal of length 20 cm and width 16 cm is to be used to construct an open-topped box. A square will be cut from each corner and the sides will be folded up. What size corner should be cut so that the volume of the box is maximized?

12. Let n, k be positive integers. The expansion of $(x_1 + \cdots + x_k)^n$ is given by

$$(x_1 + \cdots + x_k)^n = \sum \frac{n!}{n_1!n_2!\ldots n_k!} x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k},$$

where the sum is taken over all sequences $n_1, n_2, \ldots, n_k$ of non-negative integers such that $n_1 + n_2 + \cdots + n_k = n$. What is the coefficient of x^5 in the expansion of $(1 + 3x + 2x^2)^4$?

13. How many positive integers are factors of 2310?

14. A thief picks a wallet and starts running at a speed of 5 m/s with zero acceleration. In 5 seconds, a policewoman notices the alarm and starts following the thief with a speed of 3 m/s with a uniform acceleration of 1 m/s^2. How many seconds will it take the policewoman to catch the thief, assuming she follows the same path as that of the thief? (A body moving with initial velocity v_0 m/s and uniform acceleration a m/s^2 will cover a distance of $(at^2/2 + v_0t)$ meters in t seconds.)

15. Consider the following pseudocode for a function that operates on an N element array $A[1], A[2], \ldots, A[N]$ of integers.

```
function mystery (A[1..N]) {
  int i,j,position,tmp;
  for i = 1 to N {
    position = i;
    for j = i+1 to N {
      if (A[j] < A[position]) {
        position = j;
      }
    }
    tmp = A[i];
    A[i] = A[position];
    A[position] = tmp;
  }
}
```

(a) Explain what effect the function has on the input array A.

(b) If $N = 100$, how many times is the comparison `A[i] < A[position]` checked?

Description for next three questions:

The break-up of the colour of cars sold by an Indian company in a given year is provided in the pie-chart, see the Figure (1). The bar chart shows the number of grey coloured cars sold in different cities, see the Figure (2).

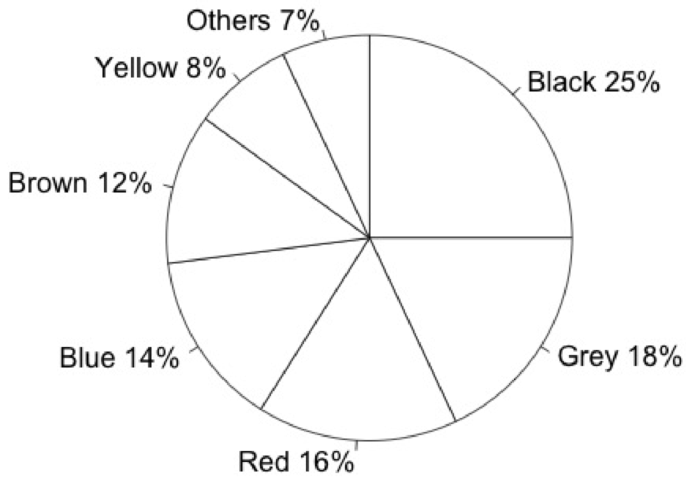


Figure 1: The break-up of the colour of cars sold in a given year.

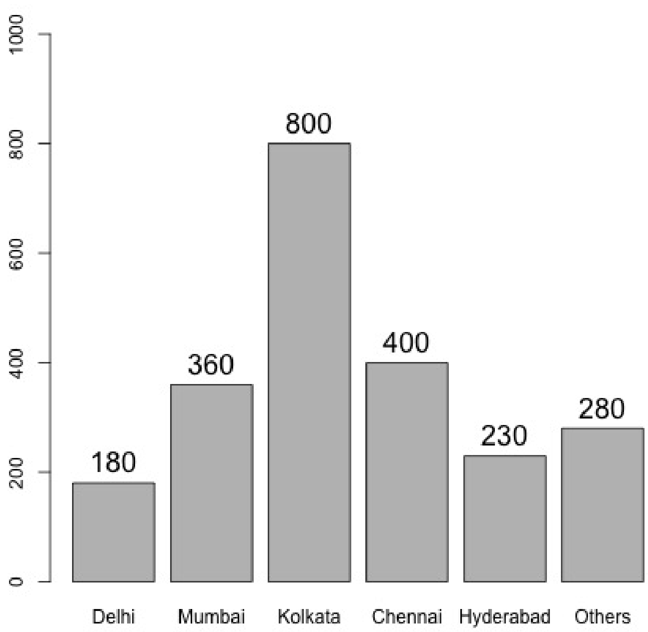


Figure 2: The number of grey cars sold in cities.

16. How many blue cars were sold in that year?

17. If the proportion of grey cars sold in Delhi is the same as the proportion of red cars sold in Delhi, how many red cars were sold in Delhi?

18. 12% of all cars were sold in Chennai. What is the largest possible percentage of brown cars sold in Chennai, rounded off to the nearest integer?