CHENNAI MATHEMATICAL INSTITUTE

M.Sc. Data Science Entrance Examination

15th May 2018

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

IMPORTANT INSTRUCTIONS

- This booklet has 8 printed sheets, including this cover page. Sheet 2 (page 1) contains space to answer part A, sheets 3 to 5 contain *twenty* questions in part A and sheets 6 to 8 contain *twenty* 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: 100** = 40 for part A + 60 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	B10	Total
B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	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.________________________________ A11. ________________________________

A2.________________________________ A12. ________________________________

A3.________________________________ A13. ________________________________

A4.________________________________ A14. ________________________________

A5.________________________________ A15. ________________________________

A6.________________________________ A16. ________________________________

A7.________________________________ A17. ________________________________

A8.________________________________ A18. ________________________________

A9.________________________________ A19. ________________________________

A10.________________________________ A20. ________________________________

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 $x \in A$.
- 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. If P is an invertible matrix and $A = PBP^{-1}$, then which of the following statements are necessarily true?

   (a) $B = P^{-1}AP$

   (b) $|A| = |B|$

   (c) A is invertible if and only if B is invertible

   (d) $B^T = A^T$.

2. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$, $C = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$ and $D = \begin{bmatrix} -1 & 2 & 3 \\ 4 & -5 & 6 \\ 7 & 8 & -9 \end{bmatrix}$.

   Which of the following statements are true?

   (a) $|A| = |B|$

   (b) $|C| = |D|$

   (c) $|B| = -|C|$

   (d) $|A| = -|D|$.

3. Let $x = \begin{bmatrix} 3 & 1 & 2 \end{bmatrix}$. Which of the following statements are true?

   (a) $x^T x$ is a 3×3 matrix
   (b) xx^T is a 3×3 matrix
   (c) xx^T is a 1×1 matrix
   (d) $xx^T = x^T x$

4. A $n \times n$ matrix A is said to be *symmetric* if $A^T = A$. Suppose A is an arbitrary 2×2 matrix. Then which of the following matrices are symmetric (here $\mathbf{0}$ denotes the 2×2 matrix consisting of zeroes):

   (a) $A^T A$
   (b) $\begin{bmatrix} \mathbf{0} & A^T \\ A & \mathbf{0} \end{bmatrix}$
   (c) AA^T
   (d) $\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A^T \end{bmatrix}$

Description for the following 3 questions: If Z is a continuous random variable which follows normal distribution with mean = 0 and standard deviation = 1, then

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

where $\Phi(a = -2) = 0.02$, $\Phi(a = -1) = 0.16$, $\Phi(a = 0) = 0.50$, $\Phi(a = 1) = 0.84$, and $\Phi(a = 2) = 0.98$. Note that $X = \mu + \sigma Z$ follows normal distribution with mean μ and standard deviation σ. The length of the life of an instrument produced by a machine has a normal distribution with a mean of 24 months and standard deviation of 2 months.

5. The probability that an instrument produced by this machine will last less than 18 months is

   (a) $\int_{-\infty}^{18} \frac{\exp\{-z^2/2\}}{\sqrt{2\pi}} dz.$
   (b) $\int_{18}^{\infty} \frac{\exp\{-z^2/2\}}{\sqrt{2\pi}} dz.$
   (c) is less than 0.02.
   (d) $\int_{-\infty}^{18} \frac{1}{\sqrt{2\pi . 2^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-24}{2}\right)^2\right\} dx.$

6. The probability that an instrument produced by this machine will die between 18 and 24 months is

   (a) $\Phi(0.5) - \Phi(-2)$
   (b) $\int_{18}^{24} \frac{1}{\sqrt{2\pi . 2^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-24}{2}\right)^2\right\} dx.$
   (c) more than 0.4.
   (d) less then 0.5.

7. Out of a large number of instruments produced by this machine, the percentage of instruments that will last more than 26 months approximately

   (a) equals 16%.
   (b) is more than 15%.
   (c) is less than 14%.
   (d) is between 10% and 15%.

Description for the following four questions: Given below is the time table for a trans-continental train cutting across several time zones. All times given below are local time in respective cities. It is given that the average speed between any two cities is the same for both way journeys.

local time arrival	local time departure		City		local time arrival	local time departure
-	06:00 AM	↓	Zut		12:40 PM	-
07:45 AM	07:50 AM	↓	Yag	↑	10:50 AM	10:55 AM
09:40 AM	09:45 AM	↓	Xum	↑	06:55 AM	07:00 AM
12:10 PM	12:15 PM	↓	Wip	↑	02:25 AM	02:30 AM
02:45 PM	02:50 PM	↓	Vaq	↑	11:50 PM	11:55 PM
04:05 PM	04:10 PM	↓	Uap	↑	08:30 PM	08:35 PM
05:40 PM	05:45 PM	↓	Tix	↑	04:55 PM	05:00 PM
07:40 PM	07:45 PM	↓	Sab	↑	12:55 PM	01:00 PM
10:40 PM	-		Raz	↑	-	08:00 AM

8. Which of the following pairs of cities are in the same time zone?

   (a) Zut and Yag
   (b) Xum and Wip
   (c) Vaq and Uap
   (d) Tix and Sab

9. Write down the total time taken in minutes by the train to go from Zut to Raz.

10. Write down time at Uap when it is 12:00 noon at Zut.

11. Write down time at Xum when it is 12:00 noon at Raz.

12. In an entrance examination with multiple choice questions, with each question having four options and a single correct answer, suppose that only 20% candidates think they know the answer to one difficult question and only half of them know it correctly and the other half get it wrong. The remaining candidates pick one option out of the four randomly and tick the same. If a candidate has correctly answered the question, what is the (conditional) probability that she knew the answer?

13. There are n songs segregated into 3 play lists. Assume that each play list has at least one song. The number of ways of choosing three songs consisting of one song from each play list is:

(A) $> \frac{n^3}{27}$ for all n

(B) $\leq \frac{n^3}{27}$ for all n

(C) $\binom{n}{3}$ for all n

(D) n^3 for all n

14. Consider the following functions defined from the interval $(0,1)$ to real numbers. Which of these functions attain their maximum value in the interval $(0,1)$?

(a) $f(x) = \frac{1}{x(1-x)}$

(b) $g(x) = -(x-0.75)^2$

(c) $u(x) = \sin(\frac{\pi x}{2})$

(d) $v(x) = x^2 + 2x$

15. A farmer owns 50 papaya trees. Each tree produces 600 papayas in a year. For each additional tree planted in the orchard, the output of each tree (including the pre-existing ones) drops by 5 papayas. How many trees should be added to the existing orchard in order to maximize the total production of papayas?

16. Let a, b be numbers between 1 and 2 and let c, d be numbers between 3 and 4. Let $u = a^{-1}, v = b^{-1}, w = c^{-1}$ and $x = d^{-1}$. Say which of the following inequalities are true:

(a) $(a+b+c+d)(u+v+w+x) > 16$

(b) $(a^4+b^4+c^4+d^4) \leq 4abcd$

(c) $(a^2+b^2)wx \leq (c^2+d^2)uv$

(d) $d(a^3+b^3+c^3) < 3abc$

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

```
i := 0;
j := 0;
k := 0;
for m in [start,start+1, ...,end]{
  if (square(m)=True){
     i := i + m*m;
     k := k + m*m;
  }else{
     j := j + m*2;
     k := k + m*2;
  }
}
```

At the end of the loop, which of the following are correct statements about the relationship between `i`, `j` and `k`?

(a) `k = i*i+j*2` if `(end - start)` is even

(b) `k = i*i+j` if `(end - start)` is odd

(c) `j = k-i` if `(end - start)` is even

(d) `i = k-j` if `(end - start)` is odd

18. Given the following definition of the function `foo`, what does `foo(1037,2)` return? Note that `a//b` denotes the quotient (integer part) of $a \div b$, for integers `a` and `b`. For instance `7//3` is `2`.

```
function foo(n,d){
  x := 0;
  while (n >= 1) {
    x := x+1;
    n := n//d;
  }
  return(x);
}
```

19. A very special island is inhabited only by knights and knaves. Knights always tell the truth, and knaves always lie. You meet three inhabitants: Ramesh, Bharat and Menaka. Ramesh claims, "Bharat is a knave." Bharat says, "Menaka and I are both knights or both knaves." Menaka claims that Bharat is a knave.

    Which of the following are correct?

    (a) Bharat is a knave
    (b) Ramesh is a knight
    (c) Exactly two of the three are knaves
    (d) Exactly two of the three are knights

20. There are 7 switches on a switchboard, some of which are *on* and some of which are *off*. In one move, you pick any 2 switches and toggle each of them—if the switch you pick is currently off, you turn it on, if it is on, you turn it off. Your aim is to execute a sequence of moves and turn all 7 switches on. For which of the following initial configurations is this possible? Each configuration lists the initial positions of the 7 switches in sequence, from switch 1 to switch 7.

    (a) (off,on,on,on,on,off,on)
    (b) (off,on,on,on,off,on,off)
    (c) (off,on,off,off,on,off,on)
    (d) (off,on,off,off,off,on,off)

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.: $\binom{n}{r}$, ^nP_r, $n!$ etc.

1. Let $N = \{1, 2, 3, \ldots\}$ be the set of natural integers and let

$$f : N \times N \mapsto N$$

   be defined by $f(m, n) = (2m - 1) * 2^n$. Is f injective? Is f surjective? Give reasons.

2. Suppose A, B and C are $m \times m$ matrices. What does the following algorithm compute? (Here $A(i, j)$ denotes the $(i.j)^{th}$ entry of matrix A.)

```
for i = 1 to m
    for j = 1 to m
        for k = 1 to m
            C(i, j) = A(i, k) * B(k, j) + C(i, j)
        end
    end
end
```

3. Find A^{10} where A is the matrix $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Justify your answer.

4. In computing, a floating point operation (flop) is any one of the following operations performed by a computer: addition, subtraction, multiplication, division. For example, the dot product of two vectors $(u, v, w) \cdot (x, y, z) = ux + vy + wz$ involves 3 multiplications and 2 additions, a total of 5 flops.
   Calculate the exact number of flops required computing $C = AB$ for two 5×5 matrices A and B using a direct implementation of $c_{ij} = \sum_{k=1}^{5} a_{ik} b_{kj}$. How does this number change if both the matrices are upper triangular?

5. A function f from the set A to itself is said to have a *fixed point* if $f(i) = i$ for some i in A. Suppose A is the set $\{a, b, c, d\}$. Find the number of bijective functions from the set A to itself having no fixed point.

6. A 4-digit number is represented as $abcd$ i.e. $a \times 10^3 + b \times 10^2 + c \times 10 + d$, where $a \neq 0$. Suppose the number $dcba$, obtained by reversing the digits of $abcd$, is 9 times $abcd$. Find the number $abcd$.

7. A computer password requires you to use exactly 1 uppercase letter, 3 lowercase letters, 3 digits and 2 special characters (there are 33 special characters that can be used). In how many ways can you create such a password?

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

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

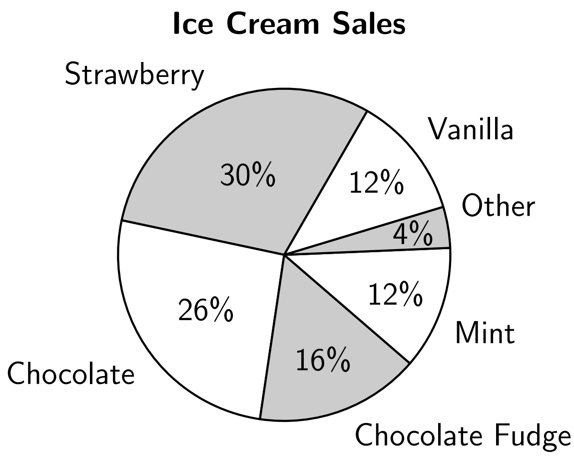
The expected number of success is $\mathbb{E}(X) = np$.

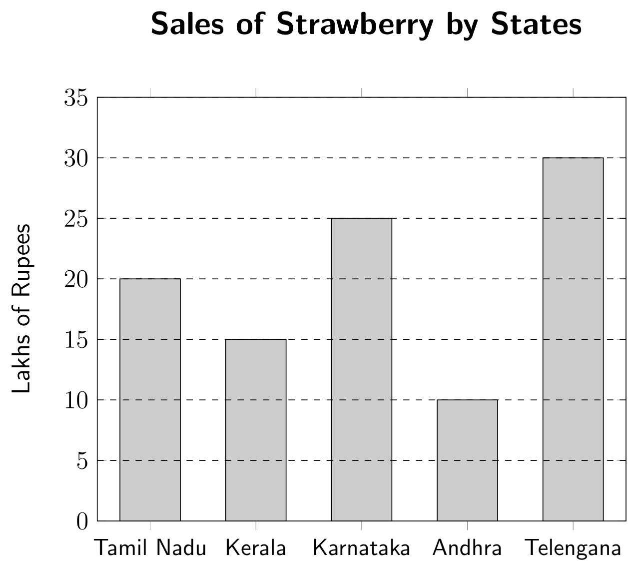
8. The probability that an individual will default on his/her credit is $\frac{1}{100}$. What is the probability that out of 200 debtors of the bank, there will be at least one credit default in a year. You can assume that whether a given debtor will default or not is independent of the behavior of other debtors.

9. For the situation in the previous problem, what is the expected number defaults?

10. Suppose the bank earns 10% of the loan amount for every loan that is repaid. If the person defaults on his/her credit then the bank loses the entire loan amount. What is the expected revenue of the bank from a loan of Rs. 100,000?

Description for the following four questions:
An ice-cream company mainly operates in the five southern states of India. The pie chart shows the breakdown of revenues (in percentages) for the ice cream company over the last summer. The bar chart shows the detail of breakdown for strawberry flavor by states in lakhs of rupees.





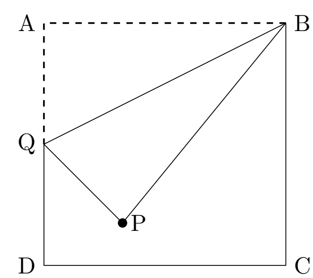
11. What are the total sales of the strawberry flavor?

12. What is the total revenue of the ice cream company?

13. What are the total sales of the chocolate flavor?

14. If you assume that the chocolate flavor and the strawberry flavor are sold in the same proportion across the five states, then what are the sales of chocolate in Tamil Nadu, in lakhs of rupees?

15. A square piece of paper $ABCD$ of side length 1 is folded along the segment that connects the upper right corner B and the midpoint Q of the left edge AD, as shown. What is the vertical distance between the base edge (segment DC) and the point P (which was originally point A)?



16. A boolean value is a value from the set {`True`,`False`}. A 3-ary boolean function is a function that takes three boolean values as input and produces a boolean value as output. Let f and g be 3-ary boolean functions. We say that f and g are *neighbours* if f and g agree on at least one input triple and disagree on at least one input triple: that is, there exists a triple (x, y, z) such that $f(x, y, z) = g(x, y, z)$ and a triple (x', y', z') such that $f(x', y', z') \neq g(x', y', z')$.
Suppose we fix a 3-ary boolean function h. How many neighbours does h have?

Description for the following four questions:

A golf club has m members with serial numbers $1, 2 \ldots, m$. If members with serial numbers i and j are friends, then $A(i, j) = A(j, i) = 1$, otherwise $A(i, j) = A(j, i) = 0$. By convention, $A(i, i) = 0$, *i.e.* a person is not considered a friend of himself or herself. Let $A^k(i, j)$ refer to the $(i, j)^{th}$ entry in the k^{th} power of the matrix A.
Suppose it is given that $A^9(i, j) > 0$ for all pairs i, j where $1 \leq i, j \leq m$, $A^2(1, 2) > 0$ and $A^4(1, 3) = 0$. Then which of the following are necessarily true? Give reasons.

17. Member 1 and member 2 have at least one friend in common.

18. $A^2(i, i) > 0$ for all i, $1 \leq i \leq m$.

19. $m \leq 9$.

20. $m \geq 6$.