CHENNAI MATHEMATICAL INSTITUTE
M.Sc. Data Science Entrance Examination 2021 - Solution

Part A - Answers

A1. (c)	A11. (a),(b)
A2. (b)	A12. (a),(d)
A3. (a)	A13. (b)
A4. (d)	A14. (b),(d)
A5. (c)	A15. (a),(b),(c)
A6. (d)	A16. (b),(c)
A7. (a),(b)	A17. (a),(c),(d)
A8. (b),(c),(d)	A18. (d)
A9. (d)	A19. (c)
A10. (a),(c)	A20. (b)

Multiple-choice questions

1. There are two longest subsequences, not necessarily contiguous, common to the strings "ARTIFICIAL" and "INTELLIGENCE". They are "IIC" and "TIC" which are of length three.

   Consider two strings S1 = "CORONAVIRUS" and S2 = "SARSCOVID". Let x be the length of a longest common subsequence between S1 and S2 and let y be the number of such longest common subsequences of length x between S1 and S2. What is $x + 5y$?

   (a) 13
   (b) 15
   (c) 14
   (d) 16

   Solution: *(b) is correct. The two ($y = 2$) lcs sequences are "covi" and "rovi". The length, $x = 4$. Therefore,* $x + 5y = 4 + 5 * 2 = 14$.

2. The roots of the polynomial $p(x) = x^4 - 2x^3 - 2x^2 + 8x - 8$ are:

   (a) $1, \ -1, \ 2, \ 2+3i$
   (b) $1+i, \ 1-i, \ 2, \ -2$
   (c) $1, \ -1+i, \ 2, \ 2+3i$
   (d) $1+i, \ -1+i, \ 2, \ -2$

   Solution: *(b) is correct. (a), (c) and (d) are wrong because complex roots can appear only in conjugate pairs.*

3. Consider the following code, in which `A` is an array indexed from `0`.

```
function foo(A,n) {
     m = A[0];
     x = 0;
     for i = 0 to n-1 {
          x = x + A[i];
          if (m < x) {
               m = x;
          }
          if (x < 0) {
               x = 0;
          }
     }
     return(m);
}
```

   If `A = [-12,-3,5,10,8,-16,-23,12,-5,7]`, what will `foo(A,10)` return?

   (a) 23
   (b) 17
   (c) -17
   (d) 35

   Solution: *(a) is correct. This code computes the maximum subarray sum which is 5 + 10 + 8 = 23.*

4. A student has an average score of 80 from her first four Mathematics tests, and 88 from her first five Physics tests. How much must she score in her upcoming tests to raise her average score in Mathematics and Physics to 82 and 89, respectively?

   (a) She should score 92 in Mathematics and 100 in Physics in the next tests.
   (b) She should score 86 in Mathematics and 90 in Physics in the next tests.
   (c) She should score 88 in Mathematics and 92 in Physics in the next tests.
   (d) She should score 90 in Mathematics and 94 in Physics in the next tests.

   Solution: *(d) is correct. The average of her first four scores in Mathematics is sum of the scores on the tests divided by* 4. *Since the average of first four test is* 80, *the sum of the four score is* $80 \times 4 = 320$. *Suppose the score of her 5th Math test is* x, *the sum of the 5 Math scores is* $320 + x$, *hence*

$$\frac{320 + x}{5} = 82.$$

   Solving for x, *we find* $x = 90$, *so she needs a* 90 *on her fifth test to raise her average of Mathematics to* 82. *Similarly, the sum of five Physics score* $88 \times 5 = 440$. *Suppose the score of her 6th Physics test is* y, *the sum of 6 Physics score would be* $440 + y$, *hence*

$$\frac{440 + y}{6} = 89.$$

   Solving for y, *we find* $y = 94$, *so she needs a* 94 *on her sixth Physics test to raise her average to* 89.

5. A binary tree starts with a single root node at the top of the tree. Each node can have either a left child or a right child, or both, or neither. The children of a node are drawn below it, connected by edges. Here are the five possible binary trees with three nodes.

```
      0             0     0              0      0
     / \           /       \            /        \
    0   0         0         0          0          0
                 /           \          \        /
                0             0          0      0
```

   Note that the directions left and right of the children matter. In the second tree, the root has a left child that has a left child, while, in the fourth tree, the root has a left child that has a right child, and so on.

   How many different binary trees can be constructed with four nodes?

   (a) 3
   (b) 5
   (c) 14
   (d) 30

   Solution: *(c) is correct.* $T(n) = (2n)!/(n+1)!n!$ *binary trees can be constructed with* n *unlabeled nodes. So,* $T(4) = 8!/5!4! = 14$.

6. Consider the following code.

```
function foo(A,n) {
    x = 0;
    for i = 0 to n-1 {
        x = A[i]^x;
    }
    return(x);
}
```

Here, `a^b` represents the bitwise *Exclusive OR* function over variables `a` and `b`. For example, `3^4 = 7` and `9^5 = 12`. The truth table for the *Exclusive OR* function is provided for your reference.

a	b	`a^b`
0	0	0
0	1	1
1	0	1
1	1	0

The truth table for the *Exclusive OR* function.

Let `A = [2,3,7,2,3]` be an array indexed from `0`. What will `foo(A,5)` return?

(a) 0
(b) 2
(c) 3
(d) 7

Solution: *(d) is correct. This code finds the only unique value (if any) among duplicates in this array. 7 will be returned.*

7. Fifteen telephones are received at a service center. Of these, 5 are mobile, 6 are cordless, and 4 are wired. These 15 phones are randomly numbered from 1 to 15 to establish the order in which they are serviced. Which of the following statement(s) is/are correct?

(a) The probability that among the first 3 serviced, the first and third are mobile and the second is not, is $\dfrac{5 \times 10 \times 4}{15 \times 14 \times 13}$.

(b) The probability that the first four serviced are all the wired phones, is $\dfrac{1}{\binom{15}{4}}$.

(c) The probability that after servicing ten of these phones, only one of the three types remain to be serviced, is $\dfrac{\binom{6}{5}}{\binom{15}{5}}$.

(d) The probability that two phones of each type are among the first six serviced, is $\dfrac{\binom{5}{2} + \binom{6}{2} + \binom{4}{2}}{\binom{15}{6}}$.

Solution:
(a) is correct.
(b) is correct.
(c) is wrong. The probability that after servicing ten of these phones, only one of the three types, remain to be serviced is following: The 5 to be serviced are either mobile or cordless. So the probability is $\dfrac{\binom{6}{5} + \binom{5}{5}}{\binom{15}{5}}$
(d) is wrong. The probability that two phones of each type are among the first six serviced is $\dfrac{\binom{5}{2}\binom{6}{2}\binom{4}{2}}{\binom{15}{6}}$

8. Which of the following statements is/are true?

(a) For $k = 69597$, $x_1, x_2, \ldots, x_k \in (0, \infty)$

$$\frac{x_1 + x_2 + \cdots + x_{2^k}}{2^k} \leq (x_1 x_2 \cdots x_{2^k})^{\frac{1}{2^k}}$$

(b) For any three real numbers x, y, z,

$$|x - z| \leq |x - y| + |y - z|$$

(c) For $|r| < 1$,

$$\sum_{n=0}^{\infty} 37 \cdot r^n = \frac{37}{1-r}.$$

(d) $(n^3 - n)$ is divisible by 3 for $n = 1, 2, 3, \ldots$

Solution: (b), (c) and (d) are correct.

(a) is wrong. It is $AM \geq GM$ inequality.

(b) is correct. It is triangle inequality.

(c) is correct. Sum of GP series.

(d) is correct. Suppose $n = 1$, then $1^3 - 1 = 0$, i.e., 0 is divisible by 3. Therefore, the statement is true for $n = 1$. Let us assume the statement is true for $n = k$, i.e., $k^3 - k = 3x$. Now, we need to prove that if the statement is true for $n = k$ then it is also true for $n = k + 1$,

The expression is the product of the three consecutive integers $(n - 1), n, (n + 1)$. One of these three must be divisible by 3.

9. Which of the following statements is/are true for an arbitrary $n \times n$ matrix A?

   (a) exchanging two rows of A does not change its determinant
   (b) exchanging two rows of A does not change its trace
   (c) replacing each diagonal element of A with a 1 does not change its determinant
   (d) exchanging two columns of A negates its determinant.

   Solution: *Only (d) is correct.*

10. Let A, B be $n \times n$ matrices. Which of the following properties of A and B are preserved under matrix multiplication?

    (a) Being upper triangular
    (b) All diagonal elements being zero
    (c) Being diagonal
    (d) Being symmetric

    Solution: (a) and (c) are preserved under matrix multiplication, (b) and (d) are not.

11. The proportion of visitors to an e-commerce website in a week who would purchase a product follows the beta distribution with the probability density function

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1$$

where Γ is the gamma function, $\alpha > 0$, $\beta > 0$ and $(\alpha+\beta) > 2$. The mean and mode of the beta distribution are $\mathbb{E}(X) = \frac{\alpha}{\alpha+\beta}$ and $\mathbb{M}ode(X) = \frac{\alpha-1}{\alpha+\beta-2}$. From historical data the mean and mode of the proportion of buyers are estimated as $\frac{1}{4}$ and $\frac{1}{5}$ respectively. Which of the following statements is/are correct?

(a) $\alpha = 3$ and $\beta = 9$

(b) The probability that in a given week the proportion of visitors who buy the product is between 10% and 20% is given by

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left[\int_0^{0.2} x^{\alpha-1}(1-x)^{\beta-1}\, dx - \int_0^{0.1} x^{\alpha-1}(1-x)^{\beta-1}\, dx \right]$$

(c) The mean proportion of buyers can be calculated as follows

$$\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx$$

(d) The probability that in a given week the proportion of visitors who buy the product is less than 20% is given by

$$1 - \int_{0.2}^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha}(1-x)^{\beta}\, dx$$

Solution: (a) is correct

$$\begin{aligned} \frac{\alpha}{\alpha+\beta} &= \frac{1}{4} \\ 4\alpha &= \alpha+\beta \quad (1) \end{aligned}$$

$$\begin{aligned} \frac{\alpha-1}{\alpha+\beta-2} &= \frac{1}{5} \\ \text{From (1)} \quad \frac{\alpha-1}{4\alpha-2} &= \frac{1}{5} \\ 5\alpha-5 &= 4\alpha-2 \\ \alpha &= 3 \end{aligned}$$

Solving for β we have $\beta = 9$

(b) is correct. As

$$\begin{aligned} \mathbb{P}(0.1 \le x \le 0.2) &= \mathbb{P}(X \le 0.2) - \mathbb{P}(X \le 0.1) \\ &= \int_0^{0.2} f(x)dx - \int_0^{0.1} f(x)dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \Big[\int_0^{0.2} x^{\alpha-1}(1-x)^{\beta-1}dx - \int_0^{0.1} x^{\alpha-1}(1-x)^{\beta-1}dx \Big] \end{aligned}$$

(c) is wrong. The average should be calculated as

$$\int_0^1 x f(x)dx = \frac{1}{4}$$

(d) is wrong. The chance that on a given week the proportion of visitor would be less than 20% is

$$1 - \int_{0.2}^{1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx$$

12. There are 10^{10} functions from $\{1, 2, \ldots, 10\}$ to $\{1, 2, \ldots, 10\}$. Pick one such function f at random. Let p_1 be the probability that $f^{-1}(3)$ has exactly four elements, and p_2 the probability that $f^{-1}(2)$ has exactly six elements. Which of the following is/are true?

(a) $p_1 > p_2$

(b) $p_1 < p_2$

(c) $p_1 = p_2$

(d) $p_1 = 3^4 p_2$

Solution: *(a) and (d) are correct.* $p_1 = \binom{10}{4}9^6/c$ *and* $p_2 = \binom{10}{6}9^4/c$ *so* $p_1 = 81p_2$.

13. For a non-zero real number a, the inverse of $J = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}$ is

(a) $\begin{pmatrix} a^{-1} & a^{-2} & a^{-3} \\ 0 & a^{-1} & a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

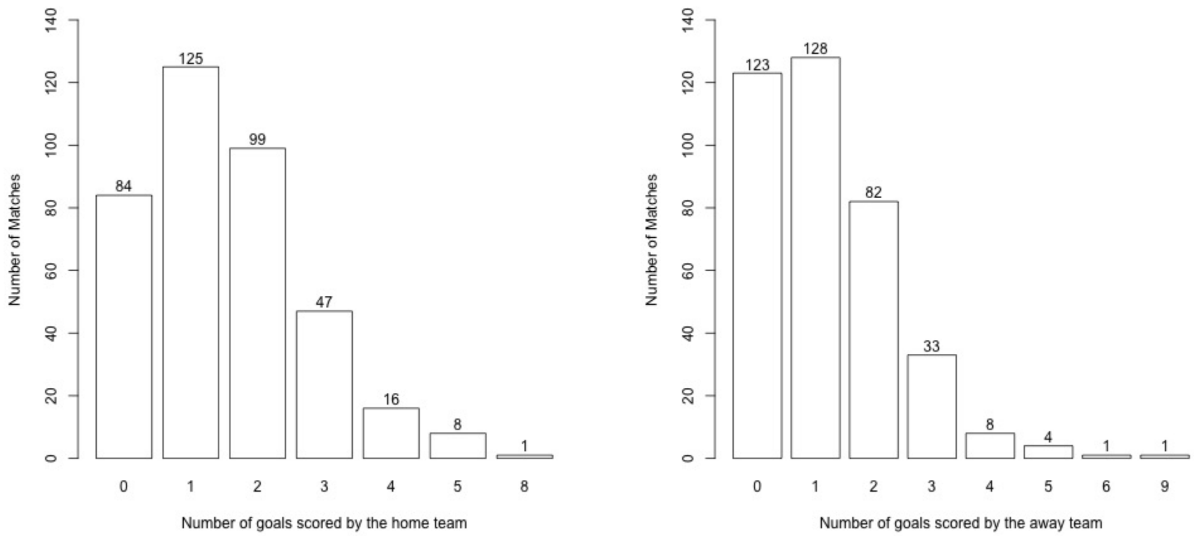
(b) $\begin{pmatrix} a^{-1} & -a^{-2} & a^{-3} \\ 0 & a^{-1} & -a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

(c) $\begin{pmatrix} a^{-1} & 1 & 0 \\ 0 & a^{-1} & 1 \\ 0 & 0 & a^{-1} \end{pmatrix}$

(d) $\begin{pmatrix} a^{-1} & a^{-2} & 0 \\ 0 & a^{-1} & a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

Solution: *(b) is correct. One way to get the answer is row/column elimination on the augmented matrix* $(J|I_{3\times3})$, *requires 4 simple steps this way.*

Description for following two questions: In the 2019-2020 season of the English Premier League (EPL), 380 matches were played in a home and away format. The figure below describes the number of goals scored by the home team and the away team against the number of matches played. For example, the home team scored one goal in 125 matches.



(a) (b)

14. Which of the following statements is/are correct?

    (a) The average number of goals scored by the home team over 380 matches is less than 1.5, and for away teams are less than 1.

    (b) In the whole season of the EPL, a total of 1034 goals were scored.

    (c) The probability that away team failed to score any goal is 22.11%.

    (d) The probability that home team scored more than 3 goals is 6.6%.

Solution: *(a) is wrong. The average number of goals scored by home team over 380 matches are* $576/380 \approx 1.515 > 1.5$ *Similarly, average number of goals scored by away team is* $458/380 \approx 1.21$
(b) is correct. The total number of goal scored by home teams are $0+1\times125+2\times99+3\times47+4\times16+8+8\times1 = 576$ *goals. Similarly, the total number of goal scored by away teams are* $0+1\times128+2\times82+3\times33+4\times88+5\times4+6\times1+9\times1 = 458$ *goals. So total number of goals scored by home and away team is* $576+458 = 1034$.
(c) is wrong. The probability that away team will fail to score any goal is $\frac{123}{380} \times 100 \approx 32.37\%$.
(d) is correct. From Figure (part (a)) in 16 matches the home team score 4 goals, in 8 matches home team scored 5 goals and in one match the home team scored 8 goals. So total 25 matches home team scored more than 3 goals out of 380 matchs, i.e., $\frac{25}{380} \times 100 \approx 6.6\%$.

15. Which of the following statements is/are correct?

    (a) In more than 50% matches, the home team scored at most one goal.

    (b) In more than 10% matches, the away team scored more than 2 goals

    (c) In more than 15% matches, the home team scored three or more goals

    (d) In more than 90% matches, the away team scored less than 3 goals

Solution: (a) is correct. The home team filed to score any goal in 84 matches and the home team scored 1 goal in 125 matches. So in 209 matched the home team scored one or less goal out of 380 matched played during the season. Hence the statment is correct.

(b) is correct. There are 47 matches where the away team scored more than 2 goals (33+8+4+1+1=47), which is $47/380 \times 100 = 12.37\%$

(c) is correct. There are 72 (=47+16+8+1) matches where the home team score 3 or more goals out of 380 matches, that is $(72/380) \times 100 \approx 18.95\%$

(d) is wrong. There are 333 (=123+128+82) matched where the away team scored less than 3 goals, that is $(333/380) \times 100 = 87.63\%$

16. You are given a sequence of N digits. From this you can generate N-digit numbers as follows: Pick up each digit in the given order, starting from the first one. Place it either to the left or to the right of the number you have generated so far. For instance, if you start with the sequence $[4, 3, 2, 5, 2]$, you can generate 53422 through the steps $4 \to 34 \to 342 \to 5342 \to 53422$, and 22345 through the steps $4 \to 34 \to 234 \to 2345 \to 22345$.

    Which of the following sequences could have generated 51324?

    (a) $[2, 1, 3, 4, 5]$
    (b) $[2, 3, 1, 5, 4]$
    (c) $[2, 3, 4, 1, 5]$
    (d) $[2, 4, 1, 3, 5]$

    Solution (b) $[2, 3, 1, 5, 4]$ and (c) $[2, 3, 4, 1, 5]$

    Explanation

    - (a) From input $[2, 1, 3, 4, 5]$, we cannot generate 132 from $[2, 1, 3]$.
    - (b) Input $[2, 3, 1, 5, 4] : 2 \to 32 \to 132 \to 5132 \to 51324$
    - (c) Input $[2, 3, 4, 1, 5] : 2 \to 32 \to 324 \to 1324 \to 51324$
    - (d) From input $[2, 4, 1, 3, 5]$, we cannot insert 3 into the sequence 124 generated from $[2, 4, 1]$.

17. Which of the following are true?

    (a) $2^{\frac{1}{2}} < 3^{\frac{1}{3}}$ and $3^{\frac{1}{3}} > 4^{\frac{1}{4}}$
    (b) $2^{r-1}r! \geq (2r-2)!$ for all natural numbers r
    (c) $\sin(2x) \geq \cos(x)$ for all values of x in $\left[\frac{\pi}{6}, \frac{\pi}{2}\right]$
    (d) $|x-1| - 1 \leq ||x-1| - 1|$ for all real numbers x

    Solution:

    (a) True: checked easily by raising to appropriate powers; raise to 6 for the first part and raise to 12 for the second part.

    (b) False: $2^{r-1}r! \geq (2r-2)!$ is true only for $r = 1, 2, 3$; for $r \geq 4$, the reverse inequality holds.

    (c) True: Check it at $k = 1$, and observe that the phenomenon repeats with period 2π; there are other intervals also where this is true, but we aren't mentioning those.

    (d) True: observe their graphs. On $(0, 2)$, we have $|x-1| - 1 < ||x-1| - 1|$, while both functions agree on values of x outside this interval.

18. Let $A_{11}, A_{12}, A_{21}, A_{22}, B_{11}, B_{12}, B_{21}, B_{22}$ be $n \times n$ matrices. Consider the $2n \times 2n$ *block* matrices $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$. The product AB is given by:

   (a) $\begin{bmatrix} A_{11}B_{11} & A_{12}B_{12} \\ A_{21}B_{21} & A_{22}B_{22} \end{bmatrix}$

   (b) $\begin{bmatrix} A_{11}B_{21} + A_{12}B_{11} & A_{11}B_{22} + A_{12}B_{21} \\ A_{21}B_{21} + A_{22}B_{11} & A_{21}B_{22} + A_{22}B_{21} \end{bmatrix}$

   (c) $\begin{bmatrix} A_{11}(B_{11} + B_{21}) & A_{12}(B_{12} + B_{22}) \\ A_{21}(B_{11} + B_{21}) & A_{22}(B_{12} + B_{22}) \end{bmatrix}$

   (d) $\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$

   Solution: *(d) is correct; block matrix multiplication works like regular matrix multiplication in this case.*

19. Consider right angled triangles with integer side lengths a, b, c where c is the hypotenuse. Call a right angled triangle *special* if a and b are odd integers. How many special right angled triangles exist?

   (a) 23
   (b) infinitely many
   (c) 0
   (d) 12

   Solution: *(c) is correct. The question is to identify all pairs of odd integers (a, b) such that $a^2 + b^2$ is a square. Suppose $a = 2k+1$ and $b = 2l+1$ for some non-zero integers k, l. Then $a^2+b^2 = 4(k^2+l^2)+4(k+l)2$, so that 2 is the remainder when $a^2 + b^2$ is divided by 4. On the other hand, any square when divided by 4 will leave remainder 0 or 1 (this can be easily checked by analyzing the even and odd integer cases; in other words, there are no squares that are 2 modulo 4. Thus $a^2 + b^2$ cannot be a square for odd numbers a and b.*

   Thus there are no special right angled triangles.

20. Given a sequence of numbers $[x_1, x_2, \ldots, x_n]$, we say that x_i is left-visible if each number in $[x_1, x_2, \ldots, x_{i-1}]$ is strictly smaller than x_i. Similarly x_i is right-visible if each number in $[x_{i+1}, x_{i+2}, \ldots, x_n]$ is strictly smaller than x_i. For instance, in the sequence $[13, 37, 24, 55, 46]$, the numbers 13, 37, 55 are left-visible and 55, 46 are right-visible.

   Suppose we construct a sequence from the set $\{17, 22, 38, 49, 53\}$ such that 4 numbers are left-visible and 2 numbers are right-visible. In addition, we are told that 22 is at the second position from the left. What are the possible numbers at the third and fifth positions?

   (a) $\{38, 53\}$
   (b) $\{38, 49\}$
   (c) $\{17, 49\}$
   (d) $\{17, 38\}$

   Solution (b) $\{38, 49\}$

   Explanation

   Since 4 numbers are left-visible and 2 are right-visible, we must have 53 at the fourth or fifth position. If it were at the fifth position, only 1 number would be right-visible, so it must be in the fourth position. This gives us the partial sequence $[-, 22, -, 53, -]$. Since 4 numbers are left-visible, the first four numbers should be in ascending order. This gives us two options, $[17, 22, 38, 53, 49]$ and $[17, 22, 49, 53, 38]$. Hence, the third and fifth positions are occupied by $\{38, 49\}$ in either order.

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. Consider the following code.

```
function foo(n) {
    answer = 0;
    x = n;
    while (x > 0) {
        y = x % 10;
        answer = (answer * 10) + y;
        x = x // 10;
    }

    x = answer;
    answer = (answer * 10) + 1;

    while(x > 0) {
        y = x % 10;
        answer = (answer * 10) + y;
        x = x // 10;
    }

    return(answer);
}
```

Here,

- `a % b` represents the remainder when a is divided by b. For example, `23 % 10 = 3`.
- `a // b` represents integer division. For example, `23 // 10 = 2`.

What will `foo(2021)` return?

Solution: *Both while loops in the code reverse the digits in the input. So, after the first loop, 2021 becomes 1202. The we concatenate 1 to it. In the next loop we reverse again to add 2021 to the result. So, the final answer is 120212021.*

2. Food delivery agents Aman, Boni, Chan, Dong and Eman were assessed on five parameters P1 to P5. They received an integer rating between 1 and 5 for each parameter. None of them received the same rating in four or more parameters. Everyone received a score of 1 in P2 or P5. All of them received the same rating in at least two parameters. The partial information related to the ratings are given in the table below.

Name	P1	P2	P3	P4	P5
Aman				3	
Boni	4				
Chan			5		
Dong	2				
Eman				1	

Considering the above information, answer the following questions:

(a) What is the maximum average rating Boni could have achieved across all the five parameters?

(b) What is the minimum average rating Chan could have received across all the five parameters?

Solution: *Boni would receive the best average if he scores either* `<4,1,5,5,5>` *or* `<4,5,5,5,1>`. *In either case, the average is 4. To receive minimum average, Chan must have received three ones, one five and a two. So Chan could have received either* `<1,1,5,1,2>`, `<1,1,5,2,1>`, `<1,2,5,1,1>` *or* `<2,1,5,1,1>`. *In each case the average is 2.*

3. Consider the following code where `A` is an array indexed from `0`.

```
function foo(A, year, n) {
    l = 0, r = n - 1, c = 0;
    while (l <= r) {
        c = c + 1;
        m = l + (r - l) // 2;
        if (A[m] == year) {
            return(c * m);
        }
        if (A[m] < year) {
            l = m + 1;
        } else {
            r = m - 1;
        }
    }
    return(-1);
}

function bar() {
    A = [2016, 2017, 2018, 2019, 2020, 2021, 2022];
    result = foo(A, 2021, 7);
    print(result);
}
```

Here, `a // b` represents integer division. For example, `23 // 10 = 2`. What will be printed when `bar()` is executed?

Solution: *This code performs binary iterative search for* 2021 *in the array* `A` *whose length is* 7. *This code locates* 2021 *during the second iteration. Therefore* $c = 2$. *Since 2021 appears at index* 5, $m = 5$. *The output will be* $5 \times 2 = 10$.

4. The CMI MSc DS 2021 batch of 60 students is holding an online event to celebrate their joining CMI. Student Aruni is in charge of organizing the musical section, and she sends out an online form where each student has to mark whether they agree or decline to performing two activities during the event: singing, and playing a musical instrument. Each student can mark one of four sets of choices: (i) agree to both singing and playing an instrument, (ii) agree to sing and decline to play an instrument, (iii) agree to play an instrument and decline to sing, or (iv) decline to do either activity.

All the students respond within the deadline, and Aruni sits down to tabulate the results so that she can plan the musical events. She finds that thirty five students agreed to sing or play an instrument, of whom twenty students agreed to do both.

How many students agreed to do exactly one activity, and how many declined to participate?

Solution: *Twenty students agreed to do both activities, and fifteen declined to do exactly one activity. Can be found by drawing a Venn diagram.*

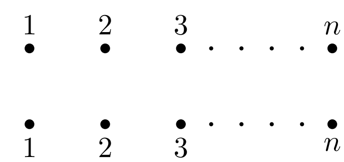
5. Show that every selection of 503 numbers from $\{1, 2, 3, \ldots, 987\}$ has two numbers with g.c.d. 1. Recall that the g.c.d. of two positive integers x, y is the largest positive integer smaller than x, y which divides both x and y.

   Solution: *Use pigeonhole principle. Pair numbers as* $(1, 2),\ (3, 4), \ldots, (2n-1, 2n)$. *When we select* $n+1$ *numbers we would certainly have to select a pair, thus consecutive numbers. In this case, we will have* 494 *'holes'* $(1, 2),\ \ldots (985, 986), (987)$, *and since we need to choose* 503 *numbers, at least 2 must come from the same hole, so at least 2 numbers must be consecutive, thus having g.c.d.* 1.

6. Let S be the sphere given by the equation $x^2 + y^2 + z^2 - 6x + 4y - 4z - 272 = 0$. What are the centre and radius of S? What is the minimum distance between S and the plane $2x + 2y + z = 214$?

   Solution: $S = (x-3)^2 + (y+2)^2 + (z-2)^2 = 272 + 17 = 289$. *The centre of the sphere is (3,-2,2) and its radius is 17. Distance of the centre from* $2x + 2y + z = 214$ *is* $|2*3 - 2*2 + 2*1 - 214/\sqrt{4+4+1}|$ *is 70. So the closest point is at a distance of 53 from the sphere.*

7. We draw two rows of n points with labels $\{1, 2, \ldots, n\}$, as shown in the figure below.



   We connect each point on the top with a point on the bottom at random, making sure that no two points on the top are connected to the same point on the bottom. We say point j on the top is *special* if it gets connected to point j below.

   Let X be the number of special points. What is $\mathbb{E}(X)$, the expected value of X?

   Solution: *These are random permutations of* $[n]$. *Probability that* i *is mapped to* i *is* $1/n$. *So if* X_i *denotes the event that* i *connects to* i, $E(X_i) = 1/n$. *So* $E(X) = n * 1/n = 1$.

8. Let T_n denote the set of all sequences over $\{0, 1, 2\}$ of length n. If we pick an element x in T_n uniformly at random what is the expected number of 1's in x?

   Solution *Let* X *denote the number of* 1*'s in* x. $X = X_1 + X_2 + \ldots + X_n$, *where* X_i *is 1 if* 1 *the i-th position, else* X_i *is zero. Since* $\mathbb{E}(X_i) = 1\left(\frac{1}{3}\right) + 0\left(\frac{2}{3}\right) = \frac{1}{3}$, *so* $\mathbb{E}(X) = \sum \mathbb{E}(X_i) = \frac{n}{3}$.

9. What is the smallest possible value of $\sum_{i=1}^{n} \frac{1}{d_i + 1}$, if the d_i's are constrained to be nonnegative real numbers and satisfy $\sum_{i=1}^{n} d_i = N$.

   Solution : *If* $d_i > d_j$, *then set* $d_i' = d_j' = (d_i + d_j)/2$. *The constraint* $\sum_{i=1}^{n} d_i = N$ *continues to hold with the new values of* d_i', d_j' . *However* $\frac{1}{d_i'+1} + \frac{1}{d_j'+1} < \frac{1}{d_i+1} + \frac{1}{d_j+1}$. *So this reduces the objective function. The minimum is therefore when all are equal, so each* $d_i = N/n$. *Harmonic mean. The objective function has value* $\sum_{i=1}^{n} \frac{n}{(N/n)+1} = \frac{n^2}{(N+n)}$

10. Let A be the 3×3 real matrix $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Suppose $x^T Ax \geq 0$ for every $x \in \mathbb{R}^3$. Then show that all of a, e and i are non-negative.

Solution: *Just take x to be the 3 standard basis vectors of $\mathbb{R}^3$; e.g. if $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, then $x^T Ax = a$, etc. Understand what $x^T Ax$ does, and deduce that these choices of x will give the answer.*

11. In the finale of the Indian Idol programme, Ladies Special, there are four women contestants — Arunima, Priyamani, Razi and Shriya. The organisers send all participants a list of 7 songs, asking each one to pick one song for the finale. All the four girls pick the same song, *Pyar Hua Chupkese* from the film *1942 A Love Story*. To resolve the tie, two days before the finale, the organisers prepare 7 sheets of paper and fold them. Exactly one of the 7 songs from the list is written on each folded sheet and the folded sheets are identical in all other aspects. First Arunima is asked to pick one sheet out of 7, then Priyamani is asked to pick one sheet out of the remaining 6, then Razi is asked to pick one out of the remaining 5 and finally Shriya is asked to pick one out of the remaining 4.

(a) What is the probability that the sheet chosen by Razi has the song *Pyar Hua Chupkese*?

(b) When Shriya is asked to choose a sheet, she argues that the three sheets chosen by the others should be opened, and if *Pyar Hua Chupkese* is not chosen by anyone, she should be allowed to pick that song. This is agreed to. What is the probability that Shriya gets to sing the song *Pyar Hua Chupkese*?

Solution:

(a) The number of ways in which Razi can choose the said song is same as counting the number of ways a song can be chosen if only 6 options are made available to the others. So Arunima can choose a song in $\binom{6}{1}$ ways, Priyamani can choose a song in $\binom{5}{1}$ ways. So the total number of ways in which Razi chooses Pyar Hua Chupkese is 30. The total number of choices is $7 \times 6 \times 5 = 210$. Thus the required probability is $\frac{30}{210} = \frac{1}{7}$.

(b) This is the probability that none of the earlier contestants chooses the song Pyar Hua Chupkese, which can happen in $6 \times 5 \times 4$ ways. So the required probability is $\frac{6 \times 5 \times 4}{7 \times 6 \times 5} = \frac{4}{7}$.

Description for following two questions: A Non-Banking Finance Corporation (NBFC) declares fixed annual rates of simple interest on their auto and housing loans each year. The rates of interest offered by the company differ from year to year depending on the variation in macro economic indicators like inflation, RBI's repo rate etc. The annual rates of interest offered by the company for the Auto and Housing sectors over the years are shown in the figure.

12. In 2013, a customer took a housing loan and a car loan. The total loan amount was ₹60 lakhs. The interest paid by the customer after one year was ₹7.36 lakhs. What was the housing loan amount?

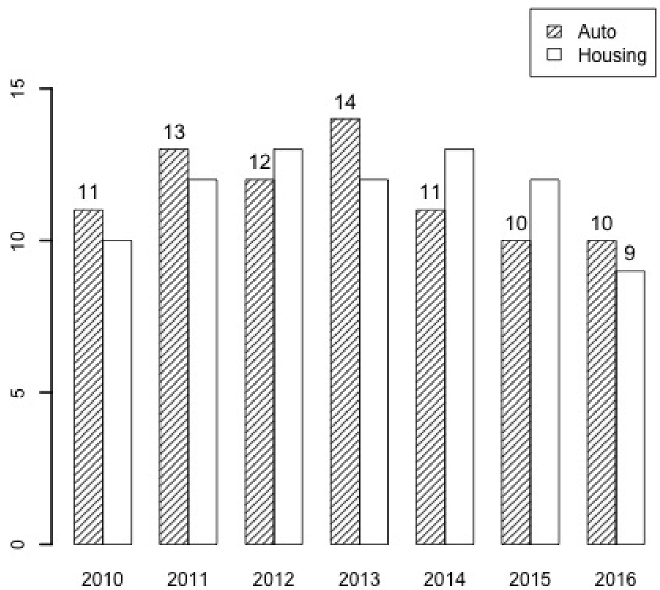
Solution*: Suppose ₹x lakhs was the housing loan amount in 2014, the amount of car loan in 2014 =₹(60 - x) lakhs. Since the loan was only one year old so it must have been agreed upon at the rate of interest prevalent in 2013, which (from the figure) was 14%. The rate of interest on the housing loan was 12%.*

Total interest received from the loans after 1 year is:

= ₹[(12% of x) + 14% of (60 - x)] lakhs

*= ₹[0.12*x + 0.14*60 -0.14*x] lakhs*

= ₹[8.4 - 0.02x] lakhs



= ₹*[8.4 -* $\frac{2}{100}x$*] lakhs*

Therefore

8.4 - $\frac{2}{100}x = 7.36 \implies x = 52$

The amount of housing loan is ₹52 lakh.

13. In 2012, Customer A took a fixed interest auto loan of ₹5 lakh for 4 years, which meant he paid interest according to the 2012 rate each year. Also in 2012, Customer B took a variable interest auto loan for ₹5 lakh for 4 years which meant his interest each year was calculated based on the prevailing interest rate for that year. Which of them paid more interest over 4 years? How much more?

    Solution*: Customer A pays* $4 \times 12 = 48\%$ *interest over 4 years. Customer B pays* $12 + 14 + 11 + 10 = 47\%$ *interest over 4 years. So Customer A pays* 1%*, or ₹5000, more.*

14. Let Y be a continuous random variable that takes values in $[0, 1]$.

    (a) $g(y)$ is a probability density function that can be used to model Y,

    $$g(y) = c.y^2(1 - y), \quad 0 \leq y \leq 1$$

    Determine c.

    (b) Using $g(y)$, find the probability that Y is greater than 0.8.

    Solution: *(a)*

    $$\int_0^1 y^2(1-y)dy = \int_0^1 y^2 dy - \int_0^1 y^3 dy = \left[\frac{y^3}{3}\right]_0^1 - \left[\frac{y^4}{4}\right]_0^1 = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$$

So the probability function is

$$g(y) = 12y^2(1-y), \quad 0 \le y \le 1$$

(b) The probability is

$$\begin{aligned}\mathbb{P}(Y > 0.8) = \int_{0.8}^{1} 12y^2(1-y)dy &= 12\Big\{\Big[\frac{y^3}{3}\Big]_{0.8}^{1} - \Big[\frac{y^4}{4}\Big]_{0.8}^{1}\Big\} \\ &= 12\Big\{\Big(\frac{1}{3} - \frac{0.8^3}{3}\Big) - \Big(\frac{1}{4} - \frac{0.8^4}{4}\Big)\Big\} \\ &\approx 12(0.16267 - 0.1476) \\ &= 0.1808\end{aligned}$$

15. The state government has decided to regulate the sale and distribution of solar power. There are multiple suppliers and consumers, as described in the tables below. The supplier table lists the amount of electricity generated by each supplier, in megawatts (MW), and the supplier's selling price per MW. The consumer table lists the amount of electricity required by each consumer, in MW, and the price the consumer will pay per MW.

Supplier ID	Quantity	Price
1	400	150
2	180	225
3	250	170
4	120	300
5	240	200

Consumer ID	Quantity	Price
1	400	180
2	300	210
3	560	200
4	240	160

The government wishes to fix a uniform market price per MW for solar power. If this price is p, any seller whose selling price is at most p will be willing to sell some or all of the electricity they generate at price p. Similarly, any consumer who is willing to pay at least p can buy as much electricity as they need at price p, upto their required quantity.

Find the value of p at which the maximum amount of solar power can be bought and sold on the market.

Solution: *As the market price increases, the supply available for sale increases and the consumer demand drops. Here are tables describing how supply and demand vary with market price.*

Market price	*0–149*	*150–169*	*170–199*	*200–224*	*225–299*	*300–*
Supply available	*0*	*400*	*650*	*890*	*1070*	*1190*

Market price	*0–160*	*161–180*	*181–200*	*201–210*	*211–*
Consumer demand	*1500*	*1260*	*860*	*300*	*0*

From this, it is clear the optimum market price is 200. At this price, the supply available is 890 MW and the demand is 860 MW so the volume of power bought and sold will be 860 MW. If the market price drops to 199, the supply drops to 650 and if the market price increases to 201, the demand drops to 300.

16. (a) If a polynomial $f(x)$ has roots $r_1, r_2, \ldots r_n$, what are the roots of the polynomial $f(x+1)$?

    (b) Let f be a polynomial of degree three. Set $g(x) = f(x-1)$. Assume none of the roots of f are -1. What are the roots of the polynomial $x^3 g(1/x)$?

Solution: *$h(x) = f(x+1)$. Then $h(r-1) = f(r-1+1) = f(r) = 0$ for every root of f. So the roots of h are $r-1$, r a root of f.*

The roots of $g(x)$ are $r+1$, where r is a root of f. Since -1 is not a root of f, none of the roots of g are zero. Let $h(x) = x^3 * g(1/x)$. If r is a root of g which is non zero, $h(1/r) = (1/r)^3 g(r)$ is zero. So the roots of $h(x)$ are $1/(r+1)$ where r is a root of f.

17. Is the following statement true?

If $f(x) \geq 0$ for all x, and $\int_{-\infty}^{\infty} f(x)\, dx < \infty$ then $\int_{-\infty}^{\infty} x^2 f(x)\, dx \geq \epsilon^2 \int_{\epsilon}^{\infty} f(x)\, dx$, for all $\epsilon > 0$.

Solution: *The statement is correct, because*

$$\begin{aligned}\int_{-\infty}^{\infty} x^2 f(x)\, dx &= \int_{-\infty}^{\epsilon} x^2 f(x)\, dx + \int_{\epsilon}^{\infty} x^2 f(x)\, dx \\ &\geq \int_{\epsilon}^{\infty} x^2 f(x)\, dx \\ &\geq \epsilon^2 \int_{\epsilon}^{\infty} f(x)\, dx\end{aligned}$$

18. (a) Consider the function $f(x) = \frac{\ln x}{x}$. Find the critical point(s) of f and say whether f has a maximum or minimum at that point.
    (b) Prove that for integers $a > b \geq 3$, we must have $a^b < b^a$.

Solution: *Using the first and second derivatives it can be shown that $f(x)$ has one critical point, $x = e$, at which it attains its maximum ($f'(x) = \frac{1-\ln x}{x^2}$ and $f''(x) = \frac{2\ln x - 3}{x^3}$). So for $x < e$, the function is increasing and for $x > e$ the function is decreasing. This implies the conclusion.*

19. Recall that if h is a function from X to Y and g is a function from Y to Z then, $g \circ h$ is the function from X to Z such that $(g \circ h)(x) = g(h(x))$, for all $x \in X$.

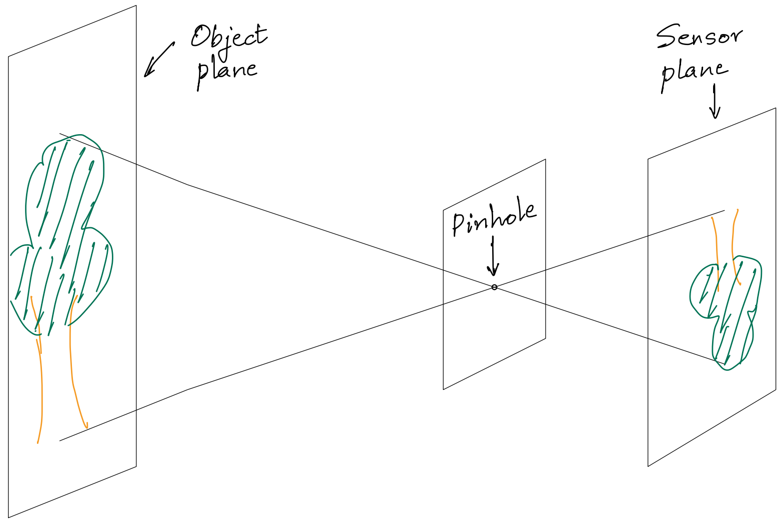
Let S be the set of all functions f from $\{1,2,3,4,5,6\}$ to $\{1,2,3,4,5,6\}$ such that $f \circ f = f$.

(a) Compute the number of functions $f \in S$ whose range has three elements.
(b) What is the cardinality of S?

Solution: *If y is an element in the range of f then $f(y) = y$ since f is idempotent. Because if z maps to y, $y = f(z) = f(f(z)) = f(y)$.*

*Can pick 3 elements in the range of f in $\binom{6}{3}$ ways. Now these elements have to go to themselves under f since $f \circ f = f$. . And the remaining 3 elements have to map to one of them. So the number of such functions is 3 * 3 * 3 =27. So the total number of functions with range of size 3 is $\binom{6}{3} * 27$.*

20. A pinhole camera is placed between an object and a sensor. Recall that images from a pinhole camera are inverted (see figure below.) Assume that the object lies on a plane and that the planes containing the object, the plane containing the pinhole and the sensor/image plane are all parallel to each other.



(a) Show that straight lines are projected to straight lines. In particular, show that if A, B, C are three collinear points on an object, then their corresponding images a, b, c are also collinear.

(b) Suppose two identical objects M and N having height 60cm are placed before the pinhole camera at distances 90cm and 120cm respectively. Suppose also that the distance between the pinhole and the image plane is 30cm. What are the heights of their images after projection through the pinhole camera?

Solution: *Let O denote the pinhole.*

(a) Notice the pairs of similar triangles: $\triangle AOB \backsim \triangle aOb$ and $\triangle BOC \backsim \triangle bOc$ implies that $\angle ABO = \angle abO$ and $\angle CBO = \angle cbO$. SO

$$\angle abO + \angle cbO = \angle ABO + \angle ABO = 180^\circ.$$

Thus, a, b, c are collinear.

(b) Let D_1 denote the distance between O and first object plane, and D_2 denote the distance between O and second object plane. Let d denote the distance between O and the image plane.

In general, if H is the height of the object and h is the height of its image, then similarity gives $\frac{H}{h} = \frac{D}{d}$, where D is the distance of the object plan from O.

From the given information we have:

$$\frac{H}{h_1} = \frac{D_1}{d} = \frac{90}{30} = 3, \text{ and } \frac{H}{h_2} = \frac{D_2}{d} = \frac{120}{30} = 4.$$

$$\text{So } \; h_1 = \frac{H_1}{3} = \frac{60}{3} = 20\text{cm}, \text{ and } h_2 = \frac{H_2}{4} = \frac{60}{4} = 15\text{cm}.$$

Thus, image of M has height 20cm and image of N has height 15cm.

Objects placed farther from the pinhole appear smaller in size.