GATE WALLAH

DATA SCIENCE AND ARTIFICIAL INTELLIGENCE

EXAM HELD ON

3rd FEBRUARY 2024

MORNING SESSION

DETAILED SOLUTION BY TEAM

DATA SCIENCE & ARTIFICIAL INTELLIGENCE

[MCQ]

Q.1. "Which algorithm sort the array in exactly using 2 swap [4, 3, 2, 1, 5]

Selection sort

Insertion sort

Bubble sort"

(a) Selection sort (b) Insertion sort

(c) Buffel sort (d) None of these

Sol. **(a)**

$$\begin{array}{ccccl} 4 & 3 & 2 & 1 & 5 \\ 1 & 3 & 2 & 4 & 5 \leftarrow \text{end of pass 1} \\ & \uparrow & & & \\ 1 & 2 & 3 & 4 & 5 \leftarrow \text{end of pass 2} \end{array}$$

$\therefore$ 'a' is correct answer.

[NAT]

Q.2. How many 4 digit number can be formed using digits 1, 3, 4, 6, 7 divisible by 3.

Sol. On deleting 3 from digits the remaining digits would sum up to 18 which is multiple of 3. So the 4 digit number can be formed using digit 1, 4, 6, 7 are 24.

On deleting 6 from digits the remaining digits would sum up to 15 which is multiple of 3. So the 4 digit number can be formed using digit 1, 3, 4, 7 are 24.

Therefore, total 4 digits no. will be 24 + 24 = 48.

[NAT]

Q.3. Let $\operatorname{Lim}_{x\to 0}\frac{\ln\left((x^2+1)\cos x\right)}{x^2}$

Sol. **(1/2)**

$$\lim_{x\to 0}\frac{ln\left[\left(x^2+1\right)\cos x\right]}{x^2}$$

Putting x $\to$ 0 in the above part we get

$$\frac{ln\{(1)\times\cos x\}}{0}=\frac{0}{0}\ from$$

Thus, using L-hospital Rule there

$$\lim_{x\to a}\frac{f(n)}{g(n)}=\lim_{x\to a}\frac{f'(n)}{g'(n)}$$

So, differentiating it we get-

$$\Rightarrow \frac{\frac{1}{\left(x^2+1\right).\cos x}\left[\cos x\cdot 2x+\left(x^2+1\right)-\sin x\right]}{2x}$$

$$\Rightarrow \frac{\frac{2x}{x^2+1} - \frac{\sin x}{\cos x}}{2x}$$

$$\Rightarrow \frac{1}{x^2+1} - \frac{\sin x}{2x \cdot \cos x}$$

Thus, $\lim_{x\to 0} \frac{ln\{(x^2+1)\cos x\}}{x^2} = \lim_{x\to 0}\left\{\frac{1}{x^2+1} - \frac{\sin x}{2x\cos}\right\}$

$$1 - \frac{1}{2}\lim_{x\to 0}\frac{\sin x}{x} \times \frac{1}{\cos x}$$

$$1 - \frac{1}{2} \times 1 \times 1$$

$$1 - \frac{1}{2} = \frac{1}{2}$$

[NAT]

Q.4. A dice is thrown repeatedly. Expected number of throws required until two consecutive throws shows even number _____.

Sol. **(6)**

PI = ½

$$N = \frac{1}{p} + \frac{1}{p^2} = \frac{1}{1/2} + \frac{1}{1/4} = 6$$

[NAT]

Q.5. $M = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 3 \\ 4 & 3 & 6 \end{bmatrix}$ det $[M^2 + 12M]$ _______.

Sol. **(0)**

= det $[M^2 + 12M]$

$|M^2 + 12M|$

$|M(M + 12I|)|$

$|M|\ \ |M + 12I|$

Now, $|M| = 1\begin{vmatrix} 1 & 3 \\ 3 & 6 \end{vmatrix} - 2\begin{vmatrix} 3 & 3 \\ 4 & 6 \end{vmatrix} + 3\begin{vmatrix} 3 & 1 \\ 4 & 3 \end{vmatrix}$

$= 1\ (6 - 9) - 2\ (18 - 12) + 3\ (9 - 4)$

$= -3 - 12 + 15$

$= -15 + 15 = 0$

Det $[M^2 + 12M]$

$|M|\ \ |M + 12I| = 0 \times (|M + 12I| = 0)$

[NAT]

Q.6. Data set with k -binary value attribute (where k>2) for a two class classification task the number of parameters to be estimated for learning a view … b…. …….

(a) $2k + 1$ (b) $k^2 +1$

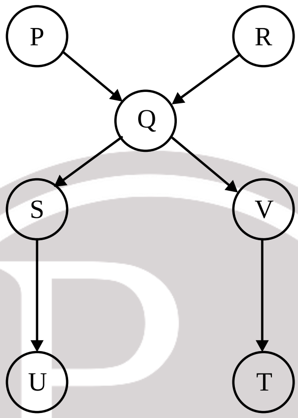
(c) $2^{k+1}+1$ (d) 2^k+1

Sol. **(b)**

The total number of parameter is $3k + 2$ & when you subsituted $3k + 2$ for $k^2 + 1$.

[MSQ]

Q.7. Consider the DAG below:



Which of the following is/are valid vertex ordering that can be obtained from topological sort of the DAG?

(a) PRQVSUT (b) PQRSVUT

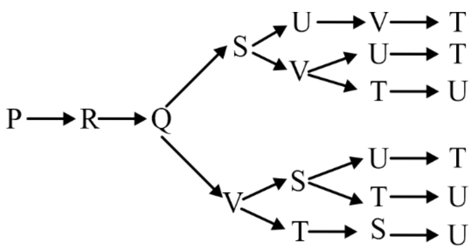
(c) PQRSTUV (d) PRQSVTU

Sol. **(a, d)**

Sol.(1). G = (V,E) directed

(1) Visit vertex with in-degree '0'

(2) Delete that vertex from the graph and repeat step-1 until graph becomes empty.



Sol.(2). By observing graph we can say that, Q cannot be selected as there are incoming edge, so the option c and b are eliminated.

If we select R, there is only one node to be selected that is Q, after selecting Q the remaining nodes. Left deleted us. Hence, we are left with the modes (S, U) and (V, T), any of these modes can be selected and deleted in any order.

∴ Correct option are 'a' and 'd'.

[MCQ]

Q.8. Consider the function computes (X) whose pseudo code is given below:
Which one of the following
Values is returned by the computes (X) for X= [6, 3, 5, 4, 10]

```
computes (X)
S[1] ← 1
for i ← 2 to length (X)
S[i] ← 1
    if X[i – 1] ≤ X[i]
    S[i] ← S[i] + S [i –1]
end if
end for
return S
```

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

Sol. **(c)**

For i = 2

S[2] ← 1

If 6 < 3 // condition false

For i = 3

S[3] ← 1

If 3 < 5 // condition true

S[3] ← S[3] + S [2] = S[3] ← 1 + 1 = 2

Therefore, S[1] = 1, S[2] = 1, S[3] = 2, S[4] = 1, S[5] = 2.

Hence, the option (c) is correct.

[MSQ]

Q.9. Given the relational schema R = (U, V, W, X, Y, Z) and the set of functional dependencies:

$$F = \begin{Bmatrix} U \to V \\ U \to W \\ WX \to Y \\ WX \to Z \\ V \to X \end{Bmatrix}$$

Which of the following FD can be derived from the above set?

(a) $VW \to Y$
(b) $VW \to YZ$
(c) $VW \to U$
(d) $WX \to YZ$

Sol. **(a, b, d)**

Given R = (U, V, W, X, Y, Z)

FD set: $\{U\to V, U\to W, WX\to Y, WX\to Z, V\to X\}$

$[VW]^+ = \{V, X, W, Y, Z\}$
Therefore VW→Y true, also VW→YZ also true and VW→U is false
$[WX]^+ = \{W, X, Z, Y\}$
Therefore WX → YZ true.
Option a, b and d are correct.

[MCQ]

Q.10. Let $f(n)$ denotes the maximum number of comparison made while searching for an entry in a sorted array of size n using binary search.
Which one of the following options is TRUE?

(a) $F(n) = F(\lfloor n/2 \rfloor)$
(b) $F(n) = F(\lfloor n/2 \rfloor) + F(\lfloor n/2 \rfloor)$
(c) $F(n) = F(n-1) + 1$
(d) $F(n) = F(\lfloor n/2 \rfloor) + 1$

Sol. **(d)**

In bianry search in one comparision entire array is divided into two halves. If we have input array size of n, then in one comparison we are making the size to $n/2$.

So, $F(n) = F(\lfloor n/2 \rfloor) + 1$

Hence, option (d) is correct.

[NAT]

Q.11. Dataset with data point $\{(x_1y_1)(x_2y_2).....(x_6y_6)\}$

$x_1 = \begin{bmatrix}1\\0\end{bmatrix}, x_2 = \begin{bmatrix}6\\1\end{bmatrix}, x_3 = \begin{bmatrix}0\\-1\end{bmatrix}, x_4 = \begin{bmatrix}-1\\0\end{bmatrix}, x_5 = \begin{bmatrix}2\\2\end{bmatrix}, x_6 = \begin{bmatrix}-2\\-2\end{bmatrix}$ labels $y_1 = y_2 = y_5 = 1$,
$y_3 = y_4 = y_6 = -1$.

A hard margin linear support vertex machine is trained on the above data set support vector.?

x_4 x_3 | x_3 x_4 x_5 | x_1 x_2 x_3 x_4 | x_1 x_2 x_5 joint probabilities.

Eg: $f_{(x,y)}(x, y)\ \{0, 2xy\ \ 0 < x < 2\ \ 0 < y < x$

$E[y \mid x = 1.5]$ is

Sol. **(6)**

The support vectors for the hard margin linear SVM trained on the given dataset are:
[[0. –1.]]
[[–1. 0.]]
[[1. 0.]]
2. Joint Probabilities
The joint probabilities for the specified combinations of data points are:
X4 X3 0.333333

X3 X4 X5: 0.166667

X1 X2 X3 X4: 0.166667

X1 X2 X5: 0.333333

3. E[y | x = 1.5]:

The conditional expectation E[y |x = 1.5] is calculated to be –-1.5

Explanation: The support vectors represent the data points that lie closest to the decision boundary of the SVM They are crucial for defining the margin and the orientation of the separating hyperplane Joint probabilities describe the likelihood of two events occurring together. In this case: they indicate the probability of specific combinations of data points.

E[y | x = 1.5] represents the expected value of y given that x = 1.5. It's a conditional expectation that measures the average value of y we expect to observe when x takes the value of 1.5.

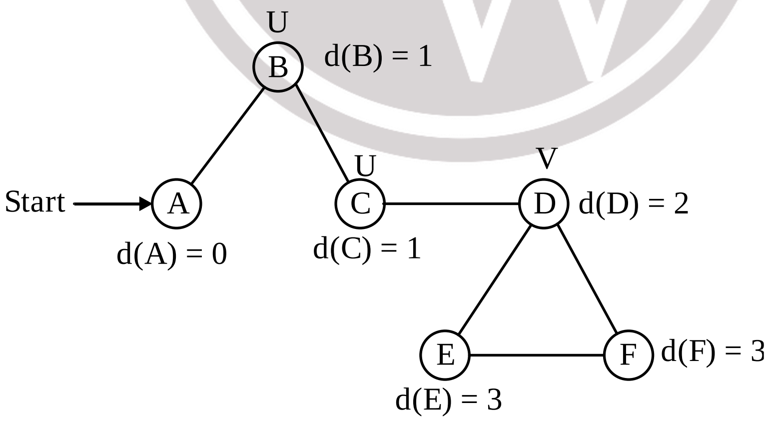
[MSQ]

Q.12. Consider performing DFS on an undirected unweighted graph G, starting at vertex S. For any vertex U in G, d[U] is the length of the shortest path from S to U. Let (U, V) be an edge in G, Such that d(V) < d(V) . If the edge (U, V) is explored first in the direction from U to V, during the above DFS, then (U, V) becomes _____ edge.

(a) Back (b) Gray

(c) Cross (d) Tree

Sol. **(d)**



Taking possible DFS starting from ‘S’.

$A \rightarrow B \rightarrow C \rightarrow D \rightarrow F \rightarrow E$

From above illustration, edges C, D are past of the tree.

Therefore, Option (d) is correct.

[MCQ]

Q.13. Match the items in column 1 with the item in column 2 in the following table.

Column 1		Column 2	
(p)	First In First out	(i)	Stacks
(q)	Lookup Operation	(ii)	Queues
(r)	Last in First out	(iii)	Hash table

(a) p-(i), q-(ii), r-(iii)
(b) p-(iii), q-(ii), r-(i)
(c) p-(ii), q-(iii), r-(i)
(d) p-(iii), q-(i), r-(ii)

Sol. **(c)**

(i) Stacks : Stacks follow last in first out.
(ii) Queues : Queues follow first in first out
(iii) Hash table : Lookup Operation performed on hash table.

[MSQ]

Q.14. Consider the following tree traversal on FBT
(i) Pre order
(ii) In order
(iii) Post order
Which of the following traversal options is/are sufficient to uniquely reconstruct the FBT?

(a) (i) and (ii) only
(b) (ii) and (iii) only
(c) (iii) and (i) only
(d) (i) only

Sol. **(a, b)**

Full binary tree can be constructed uniquely with given data.
(i) In order and Pre order
(ii) In order and post order

[NAT]

Q.15. Consider sorting the following array of integer in ascending order using an inplace Quick sort that uses last element as the pivot table

60	70	80	90	100

The minimum number of swaps performed during the Quick sort?

Sol. **(0)**

Given array, 60, 70, 80, 90, 100. During the partitioning step, all elements are compared with the pivot (which is 100 in this case). Since all other elements are smaller than 100, there won't be any swaps in the partitioning step. Thus, the minimum number of swaps performed during the Quick Sort is 0.

[NAT]

Q.16. Let $f: R \rightarrow R$ be the function $f(x) = \frac{1}{1+e^{-x}}$. The value of the derivative of at x where $f(x) = 0.4$ is.

Sol. **(0.24)**

$$f(x) = 0.4$$
$$\frac{1}{1+e^{-x}} = 0.4$$
$$1+e^{-x} = \frac{1}{0.4}$$
$$e^{-x} = \frac{1}{0.4} - 1 = \frac{1-0.4}{0.4}$$
$$e^{-x} = \frac{0.6}{0.4}$$
$$e^{-x} = \frac{6}{4} = \frac{3}{2}$$
$$\frac{1}{e^{x}} = \frac{3}{2}$$
$$e^{x} = \frac{2}{3}$$
$$ln\, e^{x} = ln\frac{2}{3}$$
$$x = ln\frac{2}{3}$$
$$f'(x) = \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right)$$
$$= \frac{-1}{\left(1+e^{-x}\right)^{2}} \times e^{-x} \times (-1) = \frac{e^{-x}}{\left(1+e^{-x}\right)^{2}}$$

Now, $x = ln\frac{2}{3}$

$$e^{-x} = e^{-ln\frac{2}{3}}$$
$$= e^{ln\frac{2}{3}} = \frac{3}{2}$$
$$= \frac{\frac{3}{2}}{\left(1+\left(\frac{3}{2}\right)\right)^{2}} = f'\left(ln\frac{2}{3}\right) = 0.24$$

[MCQ]

Q.17. $M = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$

(a) Eigen values of M are +ve and real

(b) Eigen value of M non-negative and zero

(c) Eigen value of M are complex conjugates pair

(d) Eigen value of M are negative and real.

Sol. **(c)**

$M = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$

Calculating the eigen value of M we get

$$M - \lambda I| = \left|\begin{bmatrix} 2-\lambda & -1 \\ 3 & 1-\lambda \end{bmatrix}\right| = 0$$

$$\begin{vmatrix} 2-\lambda & -1 \\ 3 & 1-\lambda \end{vmatrix} = 0$$

$(2-\lambda)(1-\lambda) + 3 = 0$

$2 - 2\lambda - \lambda + \lambda^2 + 3 = 0$

$\lambda^2 - 3\lambda + 5 = 0$

$D = 9 - 20$

$D < 0$

Thus eigen values will be complex conjugates.

[NAT]

Q.18. $f : R \to R$ is a function from R to R such that at some $x^* \in R$, $f'(x^*) = 0$

$f''(x^*) > 0$, thus of necessarily has a

Sol. **(Local Minima)**

f will have a local minima a at x^* as $f'(x^*) = 0$ denotes that x^* is a critical point. Now as the double derivative of at (x^*) is greater than 0, thus the function will have a local minima at $x = x^* \in R$

[NAT]

Q.19. Match the following:

(p) Discrimination	(i)PCA.
(q) Dimentionality Reduction	(ii)Naive Baye.
(r) Generative Model	(iii) Logistic Regression.

Sol.

(i)	PCA	(q)	Dimentionality Reduction
(ii)	Naive Baye.	(p)	Discrimination
(iii)	Logistic Regression.	(r)	Generative Model

[MCQ]

Q.20. Three fair coins are tossed independently. T is the event that two or more tosses result in heads. S is the event that two or more tosses result in tails. The probability of the event $T \cap S$ is –

(a) 0.5 (b) 0

(c) 0.25 (d) 1

Sol. **(b)**

Example space for tossing three fair coins independently is

HHH, HHT, HTT, TTT, TTH, THT, THH,

Now, T = {HHT, HHH, HTH, THH}

S = {HTT, TTT, TTH, THT}

$$P(T \cap S) = \frac{n(T \cap S)}{n(S)}$$

$$T \cap S = \phi$$

$$N\ (T \cap S) = 0$$

$$P\ (T \cap S) = \frac{0}{8} = 0$$

[MCQ]

Q.21. Consider two events T & S. Let $\overline{T}$ denotes compliment of event T. The probability associated with different event are given as follows:

$P(\overline{T}) = 0.6$

$P(S/T) = 0.3$

$P(S/\overline{T}) = 0.6$

$P(T/S) = ?$

(a) 0.25 (b) 1

(c) 0.75 (d) 0.5

Sol. **(a)**

$P(\overline{T}) = 0.6,$

$1 - P(T) = 0.6$

$P(T) = 0.4$

$P(S/T) + \dfrac{P(S \cap T)}{P(T)}$

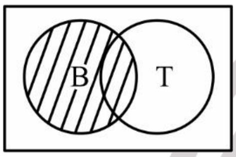
$0.3 \times 0.4 = P(S \cap T)$

$P(S \cap T) = 0.12$

$P\left(\dfrac{S}{\overline{T}}\right) = \dfrac{P(S \cap \overline{T})}{P(\overline{T})}$

$P(S \cap \overline{T}) = 0.6 \times 0.6 = 0.36$

Now



$(S \cap T)(S \cap \overline{T}) = S$

$P\{(S \cap T)(S \cap \overline{T})\} = P[S]$

$P(A + B) = P(A) + P(B) - P(A \cap B)$

Now

$P(S \cap T) + (S \cap \overline{T}) = P(S \cap T) + (S \cap \overline{T}) = F(S)$

$0.12 + 0.36 - P(S \cap T \cap \overline{T} \cap C) = F(S)$

$0.48 - P(\cap(\phi \cap S)) = P(S)$

$0.48 - P(S \cap \phi) = P(S)$

$0.48 - P(\phi) = P(\overline{S})$

$P(S) = 0.48$

$P\left(\dfrac{T}{S}\right) = \dfrac{P(S \cap T)}{P(S)} = \dfrac{0.12}{0.48} = t_1 = 0.25$

[MCQ]

Q.22. $f_x(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{other} \end{cases}$

If 5E(x) = Var (x), E(x) & Var (x) indicate the expectation and variance of x the value of λ is ______.

(a) 1/5 (b) 1/2

(c) 2/5 (d) 2

Sol. **(a)**

$$f_x(x) = \begin{Bmatrix} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{Bmatrix}$$

If 5E(X) = var (x) have E(x) denotes the expectation of x and var(x) denotes the variance of x, then the values of λ is ___

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_x(x) \cdot dx$$

$$\text{Var}(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f_x(x) \cdot dx$$

$$E(x) = \int_{-\infty}^{\infty} x f_x(x) dx$$

$$\lambda \int_0^{\infty} x e^{-\lambda x} dx$$

$$= \lambda \left[\frac{x e^{-\lambda x}}{-\lambda} - \int \frac{e^{-\lambda x}}{-\lambda} dx \right]_0^{\infty}$$

$$= \lambda \left[\frac{-1}{\lambda} e^{-\lambda x} x - \frac{e^{-\lambda x}}{\lambda^2} \right]_0^{\infty}$$

$$= \frac{1}{\lambda}$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f_x(x) dx$$

$$= \lambda \left[x^2 \frac{e^{-\lambda x}}{-\lambda} - \int 2x \frac{e^{-\lambda x}}{-\lambda} dx \right]_0^{\infty}$$

$$= \lambda \left(\frac{2}{\lambda^2} \times \frac{1}{\lambda} \right) = \frac{2}{\lambda^2}$$

$$\text{Var}(x) = E(x^2) - (E(x))^2$$

$$= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2}$$

$$\text{Var}(x) = 5E(x)$$

$$\frac{1}{\lambda^2} = \frac{5}{\lambda}$$

$$\lambda(5\lambda - 1)$$

$$\lambda = \frac{1}{5}$$

[MCQ]

Q.23. Probability to get a boy & girl is 1/2 if A man has three children, the probability to get 2 girls and 1 boy.

(a) $\frac{1}{8}$ (b) $\frac{2}{8}$

(c) $\frac{3}{8}$ (d) $\frac{1}{2}$

Sol. **(b)**

The probability to get a boy and girl is $\frac{1}{2}$. A man has three children, the probability to get 2 girl and 1 boy is___.

S = {BBB, GGG, BBG, GGB}

n(S) = 4

$$P(E) = \frac{n(E)}{n(S)} = \frac{1}{4} = \frac{2}{8}$$

[NAT]

Q.24. Consider the following relations Raider and Team managed by Kabaddi as follows:

Raider

ID	Name	Raids	Raid Point
1	-	200	250
2	-	190	219
3	-	150	200
4	-	150	190
5	-	175	220
6	-	193	215

Team

City	ID	Bid point
Jaipur	2	200
-	3	195
-	5	175
Jaipur	1	250
-	4	200
Jaipur	6	200

Select * FROM Raider Team WHERE Raider .ID = Team. ID and City = “Jaipur” and Raid point > 200.

Sol. **(3)**

Raider × Team [Cross product]

Id	Name	Raids	Raids Point	City	Bid point
1		200	250	Jaipur	250
2		190	219	Jaipur	200
3		150	200	-	195
4		150	190	-	200
5		175	220	-	175
6		193	215	Jaipur	200

The output will contain tuples where 'Raider ID' = Team · ID and City = "Jaipur"

The final output relation will consist of 3 tuples.

[MCQ]

Q.25. $f_{X,Y}(x,y)=\begin{cases}2xy & 0<x<2 \\ & 0<y<x \\ 0 & \text{otherwise}\end{cases}$ find $E[Y|x=1.5]=$

(a) 1
(b) 2
(c) 3
(d) 4

Sol. **(a)**

$$E[Y|X=x]=\int_{-\infty}^{\infty} y f_{Y/X}(y/x)dy$$

$$f_{Y/X}(y/x)=\frac{f_{X,Y}(x,y)}{f_X(x)}$$

$$f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$f_X(x)=0$ for $x\le 0,\ \ x\ge 2.$

$$f_X(x)=\int_0^x f_{X,Y}(x,y)\,dy \qquad 0<x<2$$

$$f_X(x)=\int_0^x 2xy\,dy=2x\left[\frac{y^2}{2}\right]_0^x=x^3$$

Now, $f_{Y/X}\left(\frac{y}{x}\right)=\frac{f_{X,Y}(x,y)}{f_X(x)}$

$f_{Y/X}\left(\frac{y}{x}\right)=0$ for $\begin{matrix} x\le 0, & x\ge 2 \\ y\ge x & y<0\end{matrix}$

$$f_{Y/X}\left(\frac{y}{x}\right)=\frac{2xy}{x^3} \quad \text{for} \quad \begin{array}{l} 0<x<2 \\ 0<y<x \end{array}$$

$$f_{Y/X}\left(\frac{y}{x}\right)=\frac{2y}{x^2}$$

Now, $E\left[\frac{Y}{X=x}\right]=\int_{-\infty}^{\infty} y f_{Y/X}\left(\frac{y}{x}\right)dy$

$$E\left[\frac{Y}{X=x}\right]=\int_0^x y\,\frac{2y}{x^2}dy \quad \text{for} \quad 0<x<2 \text{ only}$$

$$E\left[\frac{Y}{X=x}\right]=\frac{2}{x^2}\left[\frac{y^3}{3}\right]_0^x=\frac{2}{3}x \quad \text{for} \quad 0<x<2$$

Now, $E\left[\frac{Y}{X}=1.5\right]=\frac{2}{3}\times 1.5=1$

GATE WALLAH SOLDIERS

THANK YOU