Question 1:- Graph Theory: Consider the following incidence matrix of a graph $G = (V, E)$ with $V = \{a, b, c, d\}$ and $\{e_1, e_2, e_3, e_4, e_5, e_6\}$

$$M = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & 1 & -1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{pmatrix} \end{array}$$

Based on the information you obtained from the incidence matrix M, answer these questions:

a) What type of graph does M represent?

ans.) Degree (a) = 2

Degree (b) = 2

Degree (c) = 2

Degree (d) = 2

All the vertices have a degree of 2. Thus, ~~B~~ M is a undirected regular graph. ~~w~~ with (a, b, c, d) all have a degree of 2.

b) Find the adjacency matrix A for this graph.

ans.)

$$M = \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{bmatrix}$$

Creating empty 4 x 4 matrix:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Finding non zero entries in M =

- Column 1 = Rows 1 & 2. (a & b)
- Column 2 = Rows 1 & 2 (a & b)
- Column 3 = No zero entry
- Column 4 = Rows 1 & 3 (a & c)
- Column 5 = Rows 1 & 3 (a & c)
- Column 6 = Rows 3 & 4 (c & d).

∴ A [1] [2] = A [2] [1] = 1.

A [1] [3] = A [3] [1] = 1

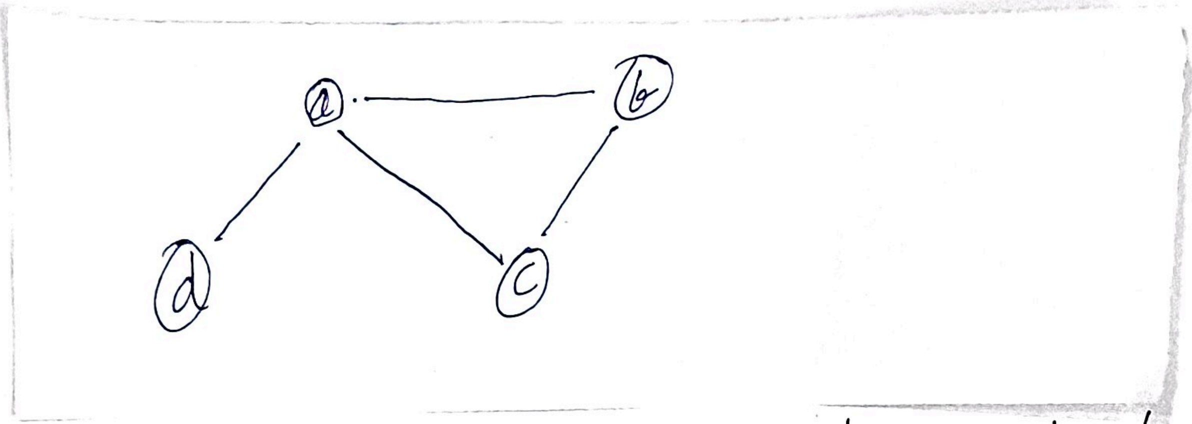
A [3] [4] = A [4] [3] = 1.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$A[i][j] = 1.$

c) Draw the graph.

ans.)



d) How many paths of 2 are there between nodes b & c

ans.) ~~Zero~~ 2.

e) In terms of connectivity of the graph, what is your interpretation of $\mathrm{tr}(A^2)$? [The question is not about the val

A^2 means there are 2 paths of length 2.

Question 2:- Probability Baye's Theorem. A data scientist is interested in studying the relationship between having a social media account & obesity. He collected some data represented in the following table.

Individual	Having social media account	Obesity
1	Yes	Yes
2	Yes	Yes
3	Yes	Yes
4	No	No
5	No.	Yes.

a) Find the probability distribution of having a social media account & being obese.

ans)

		Obese Yes	No	Margin
SM	Yes	3	0	3
	No	1	1	2
		4	1	5

Probability distribution of social media account:

$P(SM\ Yes) = \frac{3}{5} = 0.6$, $P(SM\ No) = 1 - 0.6 = 0.4$.

Probability distribution of obese:

$P(obese) = \frac{4}{5} = 0.8$, $P(not) = 1 - 0.8 = 0.2$

d) The data scientist is interested in whether having a social media account helps gain information about obesity. The concept of mutual info. quantifies how one feature is related to another. For two random X & Y, mutual info. is defined as.

$$\text{mutual information} = \sum_x \sum_y P_{X,Y}(x,y) \ln\left(\frac{P_{X,Y}(x,y)}{P_X(x) P_Y(y)}\right)$$

Suppose the joint probability distribution of having SM account & obesity is shown by $P_{X,Y}(x,y)$ & the marginal distribution of having social media & obesity are $P_X(x)$ & $P_Y(y)$.

Compute mutual information.

$$P_1 = P(SM=Y, O=Y) \ln \frac{P(SM=Y, O=Y)}{P(SM=Y) \times P(O=Y)}$$

$$= \frac{3}{5} \ln \frac{3/5}{3/5 \times 4/5}$$

$$= \frac{3}{5} \ln \frac{5}{4}.$$

$$P_2 = P(SM=Y, O=N) \ln \frac{P(SM=Y, O=N)}{P(SM=Y) \times P(O=N)}.$$

$$=$$

$$P_3 = P(SM=N, O=Y) \ln \frac{P(SM=N, O=Y)}{P(SM=N) \times P(O=Y)}$$

$$= \frac{1}{5} \ln \frac{1/5}{2/5 \times 1/5}$$

$$= \frac{1}{5} \ln \frac{5}{8}.$$

$$P_4 = P(SM=N, O=N) \cdot \ln \frac{P(SM=N, O=N)}{P(SM=N) \times P(O=N)}$$

$$= \frac{1}{5} \ln \frac{1/5}{2/5 \times 1/5}$$

$$= \frac{1}{5} \ln \frac{5}{2}.$$

Mutual information $= P_1 + P_2 + P_3 + P_4$

$$= \frac{3}{5} \ln \frac{5}{4} + \frac{1}{5} \ln \frac{5}{8} + \frac{1}{5} \ln \frac{5}{2}$$

$$= 0.22.$$

e) What is the mutual information of two independent random variable theoretically?

ans.) Mutual information is the mesure of amount of information that random variables share. When two are independent it is 0.

Question 3:- Probability Distribution: Let x be a discrete random variable that takes values in $\{-2,-1,0,1,2\}$. with equal probability, Also, Y is another discrete random variable defined as $Y = |X|$.

a) Construct the joint probability distribution table.

ans.)

$X = x$	-2	-1	0	1	2
$P(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

$$Y = |X|$$

$$\therefore Y = \{2, 1, 01, 2.$$

Hence,

$Y = y$	0	1	2
$P(y)$	$1/5$	$2/5$	$2/5$

Joint Probability Distribution :-

$Y \backslash X$	-2	-1	0	1	2	
0	0	0	$1/5$	0	0	$1/5$
1	0	$1/5$	0	$1/5$	0	$2/5$
2	$1/5$	0	0	0	$1/5$	$2/5$
	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$	1

b) Are X & Y independent? Justify:

$P(X=-2) = 1/5$

$P(Y=2 = 2/5$.

$P(X=-2 \ \& \ Y=2) = 1/5$.

We can see that when X changes Y also changes. Y ~~either~~ takes a positive value of every value that X has hence they both are dependent.

c) Find Corr (X, Y).

ans). Covariance X, Y =

$$Cov(X,Y) = \frac{1}{n} \Sigma (x_i - \bar{x}) \times (y_i - \bar{y}).$$

$$\bar{x} = \frac{\Sigma x_i}{n} \quad \& \quad \bar{y} = \frac{\Sigma y_i}{n}$$

$\therefore$

$$\frac{1}{n} \Sigma \left(x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x}\bar{y} \right).$$

$$= \frac{1}{n} \Sigma x_i y_i - \bar{y} \frac{\Sigma x_i}{n} - \bar{x} \frac{\Sigma y_i}{n} + \bar{x}\bar{y}.$$

$$= \frac{1}{n} \Sigma x_i y_i - \bar{y}\bar{x} - \bar{x}\bar{y} + \bar{x}\bar{y}$$

$$= \frac{1}{n} \Sigma x_i y_i - \bar{x}\bar{y}$$

X	Y	XY
-2	2	-4
-1	1	-1
0	0	0
1	1	1
2	2	4.
$\Sigma X = 0$	$\Sigma Y = 6$	$\Sigma XY = 0$.

$$\bar{x} = \frac{\Sigma x}{5} = 0.$$

$$\bar{y} = \frac{\Sigma y}{5} = \frac{6}{5}.$$

$$\Sigma xy = 0.$$

$$Cov(x, y) = \frac{1}{n} \Sigma xy - \bar{x} \cdot \bar{y}.$$

$$= 0 - 0 = 0.$$

$$Corr(x, y) = \frac{Cov(x, y)}{\sigma_x \sigma_y}.$$

$$= 0.$$

$P(SM) = \frac{3}{5}$

$P(\text{not } SM) = \frac{2}{5}$

$P(\text{obese}) = \frac{4}{5}$

$P(\text{not obess}) = \frac{1}{5}$

$\therefore$

$P(SM \text{ and obese}) = P(SM) \times P(\text{obese})$

$= \frac{3}{5} \times \frac{4}{5} = \frac{12}{25}.$

c) Are these two distributions independent?

ans) For this we should check :-

$P(\text{obese} | SM) = P(O).$

$P(\text{obes}/SM) = \frac{3}{3} = 1$

$P(\text{obese}) = \frac{4}{5}.$ $\therefore P(O/SM) \neq P(O).$

~~similarly~~, $\therefore$ Two are independent.

d) Based on your answer to part b), can you explain the result in part (c).

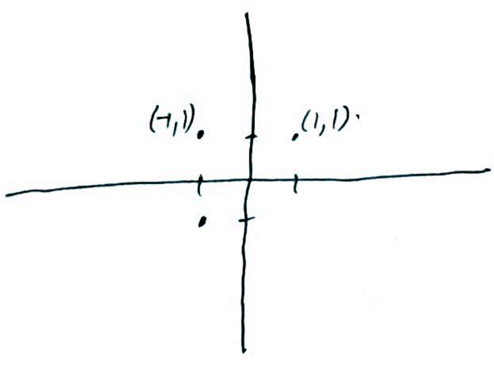
ans.).

Correlation can vary from -1 to +1.

When -1 it means they are linear opposite related.

When 1 it means they are linearly positive related

Since $Y = |X|$. ~~X can b~~ Y has 1 value but X can be negative & also could be positive with Y having the same value.



We can see by plotting in the graph with X value negative or positive Y takes a possitive value.