Week 10: Revision

$$\boxed{P(A|E) = \frac{P(A \cap E)}{P(E)}}$$

Q1. An integer between 1 and 20 (inclusive) is generated uniformaly at random. Let X be the random variable that denotes the integer.
What is the value of $P[X \geq 8 | X \text{ is even}]$?

*

j	~~1~~	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	7	8	9	10	11	12	13	14	15	16
$P[X=j]$	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20	1/20
$P[X=j \mid X \geq 8]$	0	0	0	0	0	0	0	1/13	1/13	1/13	1/13	1/13	1/13	1/13	1/13	1/13
								1/7	0	1/7	0	1/7	0	1/7	0	1/7

j	17	18	19	20
$P[X=j]$	1/20	1/20	1/20	1/20
$P[X=j \mid X \geq 8]$	1/13	1/13	1/13	1/13
	0	1/7	0	1/7

- New information: X is ≥ 8
so, X cannot be 1, 2, 3, 4, 5, 6, 7.
Formally, conditioned on X is ≥ 8, the probability of these values is 0.
i.e. $P[X=j \mid X \geq 8] = 0$

* Another new information: X is an even number
so X cannot be 9, 11, 13, 15, 18, 19.

j	8	10	12	14	16	18	20
$P[X=j]$	1/20	1/20	1/20	1/20	1/20	1/20	1/20
$P[X \geq 8 \mid X \text{ is even}]$	1/7	1/7	1/7	1/7	1/7	1/7	1/7

~~$P[X \geq 8 \mid X \text{ is even}] = \frac{1}{7}+\frac{1}{7}+\frac{1}{7}+\frac{1}{7}+\frac{1}{7}+\frac{1}{7}+\frac{1}{7} = \frac{}{7}$~~

$\therefore \frac{1}{7} = 0.142857$

$$P(X \geq 8 \mid X \text{ is even}] = \frac{P(X \geq 8 \cap X \text{ is even})}{P(X \text{ is even})}$$

$$= \frac{\cancel{\;}}{1/\cancel{10}} = \cancel{\;} \; \frac{1/7}{1/10} = 1.428571$$

j	2	4	6	8	10	12	14	16	18	20
$P[X=j \mid X \text{ is even}$	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10

Q2. In a big company, 60% of the employees are female, and the rest are male. Among all the female employees, 70% are married. Among all the male employees, 40% are married.

Question: Among employees who are married, How many percentage of them are female?

Employees → 60% Female → 70% Married, 30% Not married

Employees → 40% Male → 40% Married, 60% Not married

60% = 100%

x   70%

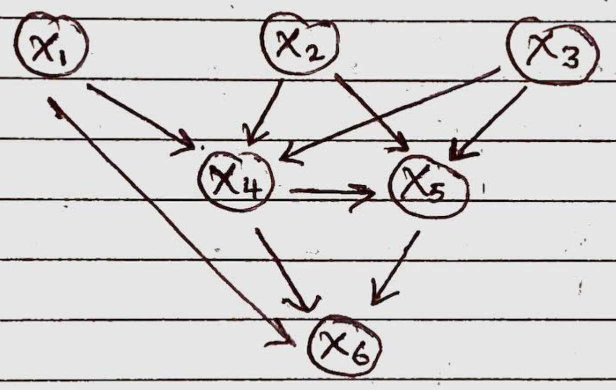
$x = \frac{60 \times 70}{100} = 42\%$

40% = 100%

x   40%

$x = \frac{40 \times 40}{100} = 16\%$

Q3. In the Bayesian network below, which random variables does X_4 directly depend on?



Ans: X_1, X_2 and X_3

Q4. How many parameters are needed to specify the Bayesian network above.

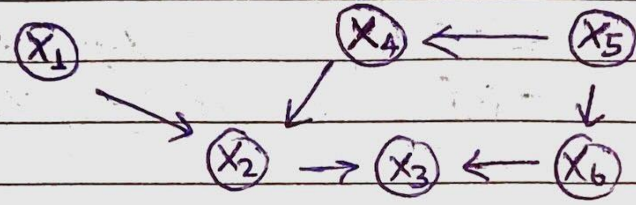
$2^0 \times 3$ X_1, X_2, X_3 do not depend on any other variable.

$2^3 \times 3$ X_4, X_5, X_6 depends on 3 other variable.

$\therefore\ 2^0 \times 3 + 2^3 \times 3 = 1 \times 3 + 8 \times 3 = 3 + 24 = \underline{27}$

Qs. At the final session of the lecture slides, we have presented a concrete example about COVID, which is a Bayesian network with 6 random variables, and the network has 14 parameters.

For this Bayesian network, compute $P[X_3=1 \mid X_1 \neq X_4]$. to 5d.p.



Step (1) Total number of parameters needed

$2^0 \times 2$. X_1, X_5 do not depend on any other variable.
$2^1 \times 2$ X_4, X_6 depends on 1 other variable.
$2^2 \times 2$ X_2, X_3 depends on 2 other variable.
$\therefore\ 2^0 \times 2 + 2^1 \times 2 + 2^2 + 2 = 2 + 4 + 8 = 14$

Step (2)
$P[X_1, X_2, X_3, X_4, X_5, X_6] = P[X_1] \cdot P[X_5] \cdot P[X_4 \mid X_5] \cdot P[X_6 \mid X_5]$
$\cdot P[X_2 \mid X_1, X_4] \cdot P[X_3 \mid X_2, X_6]$

Note: $P[X_4 \mid X_5]$ means X_4 depends on X_5

Step (3) the Bayes formula. $P(A \mid E) = \dfrac{P(A \cap E)}{P(E)}$

$$P[X_3 = 1 \mid X_1 \neq X_4] = \frac{P[X_3 = 1, X_1 \neq X_4]}{P[X_1 \neq X_4]}$$

$\cancel{P[X_3, X_1 \neq X_4] =}$

$$
\begin{aligned}
P[X_3 = 1,\ X_1 \neq X_4] = {} & P[X_1=0, X_2=0, X_3=1, X_4=1, X_5=0, X_6=0] \\
+ {} & P[X_1=0, X_2=0, X_3=1, X_4=1, X_5=0, X_6=1] \\
+ {} & P[X_1=0, X_2=0, X_3=1, X_4=1, X_5=1, X_6=0] \\
+ {} & P[X_1=0, X_2=0, X_3=1, X_4=1, X_5=1, X_6=1] \\
+ {} & P[X_1=0, X_2=1, X_3=1, X_4=1, X_5=0, X_6=0] \\
+ {} & P[X_1=0, X_2=1, X_3=1, X_4=1, X_5=0, X_6=1] \\
+ {} & P[X_1=0, X_2=1, X_3=1, X_4=1, X_5=1, X_6=0] \\
+ {} & P[X_1=0, X_2=1, X_3=1, X_4=1, X_5=1, X_6=1]
\end{aligned}
$$

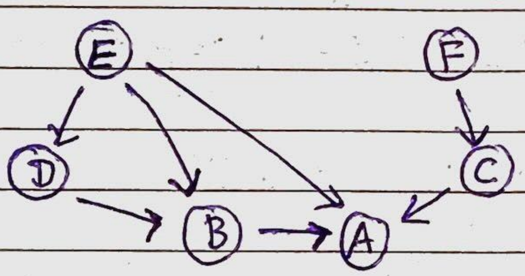
$$
\begin{aligned}
P[X_1, X_2, X_3, X_4, X_5, X_6] = {} & P[X_1] \cdot P[X_5] \\
& \cdot P[X_4 \mid X_5] \cdot P[X_6 \mid X_5] \cdot \\
& \cdot P[X_2 \mid X_1, X_4] \cdot P[X_3 \mid X_2, X_6]
\end{aligned}
$$

[redacted]

Q4. b) Suppose there are 6 binary random variables:
A, B, C, D, E and F.
The conditional independence among these variables enables the joint probability distribution
P(A, B, C, D, E, F) to be defined as

$$P(A,B,C,D,E,F) = P(A|B,C,E)\,P(B|D,E)\,P(C|F)\,P(D|E)\,P(E)\,P(F)$$

i) Draw the Bayes network that corresponds to this conditional independence. Is it a polytree? Justify your answer.



A polytree (directed tree, single connected network) — is a directed acyclic graph whose underlying undirected graph is a tree.

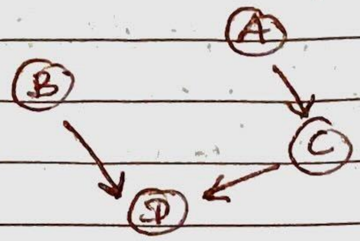
NO, It is not a ploytree because the is a cycl between the varriables.

ii) Specify the minimum number of parameters required to fully define the joint probability distribution.

$2^0 \times 2$	E, F	Do not depend on othe variable
$2^1 \times 2$	D, C	Do depend on 1 other varriable
$2^2 \times 1$	B	Does depend on 2 other vari able
$2^3 \times 1$	A	Does depend on 3 other variable.

$\therefore 2^0 \times 2 + 2^1 \times 2 + 2^2 \times 1 + 2^3 \times 1 = 2+4+4+8 = 18$

c) Consider the following Bayes network where each random variable can take two possible values {T, F}:



The associated probability distributions for the binary random variables A, B, C and D are

$P(A=T)=0.1$, $P(B=T)=0.8$, and

$P(C=T \mid A=T)=0.6$ $P(C=T \mid A=F)=0.3$

$P(D=T \mid C=T, B=T)=0.1$ $P(D=T \mid C=F, B=T)=0.5$

$P(D=T \mid C=T, B=F)=0.3$ $P(D=T \mid C=F, B=F)=0.8$

i) Write down an expression for full joint distribution of the random variables A, B, C and D.
Compute the probability that A and C are T while B and D are F.

$P(A, B, C, D) = P(D \mid B, C)\ P(C \mid A)\ P(A)\ P(B)$

$P(A \,\&\, C = T \mid B \,\&\, D = F) =$

ii) Explain why the representation of the joint probability distribution of A, B, C, and D using the Bayes network is preferable to a direct tabular representation of the full joint probability distribution.