Causal inference HW1

Pablo Ruiz de Velasco
261406

Sébastien Jeanfavre
269551

Pål Forr Austnes
289508

March 28, 2022

Problem 1

a)

$$p(x_1; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_1-1)^2} \qquad \iff \qquad X_1 \sim \mathcal{N}(1,1) \tag{1}$$

b)

Thanks to linearity of expectation, we have:

$$E\{2X_1 + X_2 + X_3\} = 2 \cdot E\{X_1\} + E\{X_2\} + E\{X_3\} = 9 \tag{2}$$

The variance can be decomposed as follows:

$$\begin{aligned} Var\{2X_1 + X_2 + X_3\} &= 2^2 \cdot Var\{X_1\} + Var\{X_2\} + Var\{X_3\} + ... \\ &\quad 2 \cdot (2 \cdot Cov\{X_1, X_2\} + 2 \cdot Cov\{X_1, X_3\} + Cov\{X_2, X_3\}) \\ &= 21 \end{aligned} \tag{3}$$

With $Y = 2X_1 + X_2 + X_3$, this gives:

$$p(y; \mu, \sigma^2) = \frac{1}{\sqrt{21 \cdot 2\pi}} e^{-\frac{1}{2}\frac{(y-9)^2}{21}} \qquad \iff \qquad Y \sim \mathcal{N}(9, 21) \tag{4}$$

c)

From the covariance matrix we observe that X_3 is independent from X_1 and X_2. It follows that:

$$p(x_3|x_1, x_2) = p(x_3) = \frac{1}{3 \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x_3-2)^2}{9}} \qquad \iff \qquad X_3 \sim \mathcal{N}(2, 9) \tag{5}$$

d)

With Bayes theorem, we have that

$$p(x_2, x_3|x_1) = \frac{p(x_2, x_3, x_1)}{p(x_1)} \tag{6}$$

where

$$p(x_2, x_3, x_1) = \frac{1}{\sqrt{(2\pi)^3|\mathbf{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{7}$$

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} 4/3 & -1/3 & 0 \\ -1/3 & 1/3 & 0 \\ 0 & 0 & 1/9 \end{bmatrix} \tag{8}$$

$$|\mathbf{\Sigma}| = (4-1)\cdot 9 = 27 \tag{9}$$

$$(\boldsymbol{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = (x_1-1)^2 - \frac{2}{3}(x_1-1)(x_2-5) + \frac{1}{3}(x_2-5)^2 + \frac{1}{3}(x_3-2)^2 \tag{10}$$

Therefore,

$$p(x_2, x_3|x_1) = \frac{\frac{1}{\sqrt{21(2\pi)^3}}\exp(-\frac{1}{2}(x_1-1)^2 - \frac{2}{3}(x_1-1)(x_2-5) + \frac{1}{3}(x_2-5)^2 + \frac{1}{3}(x_3-2)^2)}{\frac{1}{\sqrt{2\pi}}\exp-\frac{1}{2}\left(x-1\right)^2} \tag{11}$$

Problem 2

1)

We explore all the possibilities of the form:

$$X_i \perp\!\!\!\perp X_j|\boldsymbol{S}, \qquad \boldsymbol{S} \subseteq \{X_1, X_2, X_3, X_4\} \setminus \{X_i, X_j\}, \quad 1 \leq i < j \leq 4 \tag{12}$$

1. $X_1 \perp\!\!\!\perp X_2|X_3, X_4$? X_1 and X_2 are not d-separated by $\{X_3, X_4\}$, there is a collider-free path traversing no member of $\boldsymbol{S}$.

2. $X_1 \perp\!\!\!\perp X_2|X_3$? X_1 and X_2 are not d-separated by X_3, there is a collider-free path traversing no member of $\boldsymbol{S}$.

3. $X_1 \perp\!\!\!\perp X_2|X_4$? X_1 and X_2 are not d-separated by X_4, there is a collider-free path traversing no member of $\boldsymbol{S}$.

4. $X_1 \perp\!\!\!\perp X_3|X_2, X_4$? X_1 and X_3 are d-separated by $\{X_2, X_4\}$ because X_2 is not a collider.

5. $X_1 \perp\!\!\!\perp X_3|X_2$? X_1 and X_3 are not d-separated by X_2. Indeed, We can follow the path (1,2,4,3) since 2 is a collider and hence is not blocking.

6. $X_1 \perp\!\!\!\perp X_3|X_4$? X_1 and X_3 are not d-separated by X_4, there is a collider-free path traversing no member of $\boldsymbol{S}$.

7. $X_1 \perp\!\!\!\perp X_4|X_2, X_3$? X_1 and X_4 are not d-separated by $\{X_2, X_3\}$ since X_2 is a collider in $\boldsymbol{S}$ hence is not blocking.

8. $X_1 \perp\!\!\!\perp X_4|X_2$? X_1 and X_4 are not d-separated by X_2 since X_2 is a collider in $\boldsymbol{S}$ hence is not blocking.

9. $X_1 \perp\!\!\!\perp X_4|X_3$? X_1 and X_4 are not d-separated by X_3 since X_3 is a collider in $\boldsymbol{S}$ hence is not blocking.
10. $X_2 \perp\!\!\!\perp X_3|X_1, X_4$? X_2 and X_3 are not d-separated by $\{X_1, X_4\}$, there is a collider-free path traversing no member of $\boldsymbol{S}$.
11. $X_2 \perp\!\!\!\perp X_3|X_1$? X_2 and X_3 are not d-separated by X_1, there is a collider-free path traversing no member of $\boldsymbol{S}$.
12. $X_2 \perp\!\!\!\perp X_3|X_4$? X_2 and X_3 are not d-separated by X_4, there is a collider-free path traversing no member of $\boldsymbol{S}$.
13. $X_2 \perp\!\!\!\perp X_4|X_1, X_3$? X_2 and X_4 are not d-separated by $\{X_1, X_3\}$, there is a collider-free path traversing no member of $\boldsymbol{S}$.
14. $X_2 \perp\!\!\!\perp X_4|X_1$? X_2 and X_4 are not d-separated by X_1, there is a collider-free path traversing no member of $\boldsymbol{S}$.
15. $X_2 \perp\!\!\!\perp X_4|X_3$? X_2 and X_4 are not d-separated by X_3, there is a collider-free path traversing no member of $\boldsymbol{S}$.
16. $X_3 \perp\!\!\!\perp X_4|X_1, X_2$? X_3 and X_4 are not d-separated by $\{X_1, X_2\}$, there is a collider-free path traversing no member of $\boldsymbol{S}$.
17. $X_3 \perp\!\!\!\perp X_4|X_1$? X_3 and X_4 are not d-separated by X_1, there is a collider-free path traversing no member of $\boldsymbol{S}$.
18. $X_3 \perp\!\!\!\perp X_4|X_2$? X_3 and X_4 are not d-separated by X_2, there is a collider-free path traversing no member of $\boldsymbol{S}$.

2)

A possible SEM could be:

$$\begin{aligned}
X_1 &= N_{X_1} \\
X_2 &= X_1 + X_4 + N_{X_2} \\
X_3 &= X_2 + X_4 + N_{X_3} \\
X_4 &= N_{X_4} \\
N_{X_i} &\overset{iid}{\sim} \mathcal{N}(0, \sigma_i^2)
\end{aligned}$$

This SEM gives the following joint distribution.

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1, X_4)P(X_3|X_2, X_4)P(X_4) \tag{13}$$

The only d-separation is $X_1 \perp\!\!\!\perp X_3|X_2, X_4$ which implies $P(X_1, X_3|X_2, X_4) = P(X_1|X_2, X_4)P(X_3|X_2, X_4)$. By definition of the conditional probability, we have:

$$\begin{aligned}
P(X_1, X_3|X_2, X_4) &= \frac{P(X_1, X_2, X_3, X_4)}{P(X_2, X_4)} && (14) \\
&= \frac{P(X_1)P(X_2|X_1, X_4)P(X_3|X_2, X_4)P(X_4)}{P(X_2, X_4)} && (15) \\
&= \frac{P(X_1)P(X_2, X_1, X_4)P(X_3|X_2, X_4)P(X_4)}{P(X_2, X_4)P(X_1, X_4)} && (16) \\
& && (17)
\end{aligned}$$

Since nodes 1 and 4 are independent, it gives

$$
\begin{aligned}
P(X_1, X_3|X_2, X_4) &= \frac{P(X_1)P(X_2, X_1, X_4)P(X_3|X_2, X_4)P(X_4)}{P(X_2, X_4)P(X_1)P(X_4)} && (18)\\
&= \frac{P(X_2, X_1, X_4)P(X_3|X_2, X_4)}{P(X_2, X_4)} && (19)\\
&= P(X_1|X_2, X_4)P(X_3|X_2, X_4) && (20)
\end{aligned}
$$

Therefore, the conditional independence $X_1 \perp\!\!\!\perp X_3|X_2, X_4$ holds in the distribution of the SEM.

Problem 3

$$
P(H|E_1, E_2) = \frac{P(H, E_1, E_2)}{P(E_1, E_2)} = \frac{P(E_1, E_2|H)P(H)}{P(E_1, E_2)} \tag{21}
$$

We need $P(E_1, E_2|H), P(H), P(E_1, E_2)$ so answer b) is correct.
Since $P(E_1, E_2, |H) = P(E_1|H) \cdot P(E_2|H)$ only holds when E_1 and E_2 are conditionally independent, neither a) nor c) can be correct.

Problem 4

(a)

In the following, R, S and W respectively correspond to rain, sprinkler and grass wet.

$$
P(R = T|W = T) = \frac{P(W = T|R = T)P(R = T)}{P(W = T)} \tag{22}
$$

$$
\begin{aligned}
P(W = T|R = T) &= \frac{P(W = T, R = T)}{P(R = T)}\\
&= \sum_S \frac{P(W = T, R = T, S)}{P(R = T)}\\
&= \sum_S \frac{P(W = T|R = T, S)P(R = T, S)}{P(R = T)}\\
&= \sum_S \frac{P(W = T|R = T, S)P(S|R = T)P(R = T)}{P(R = T)}\\
&= P(W = T|R = T, S = F)P(S = F|R = T) + P(W = T|R = T, S = T)P(S = T|R = T)\\
&= 0.8 * 0.99 + 0.99 * 0.01 = 0.8019
\end{aligned} \tag{23}
$$

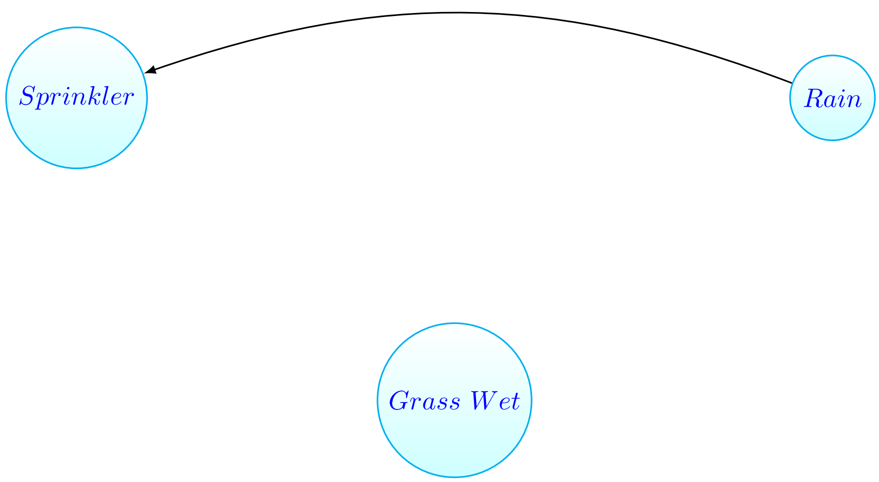
$$
\begin{aligned}
P(W=T) &= \sum_{S,R} P(W=T,R,S) \\
&= \sum_{S,R} P(W=T|R,S)P(R,S) \\
&= \sum_{S,R} P(W=T|R,S)P(S|R)P(R) \\
&= P(W=T|R=F,S=F)P(S=F|R=F)P(R=F) + P(W=T|R=T,S=F)P(S=F|R=T)P(R=T) \\
&+ P(W=T|R=F,S=T)P(S=T|R=F)P(R=F) + P(W=T|R=T,S=T)P(S=T|R=T)P(R=T) \\
&= 0 + 0.8*0.99*0.2 + 0.9*0.4*0.8 + 0.99*0.01*0.2 \\
&= 0.44888
\end{aligned}
\tag{24}
$$

$$
\implies P(R=T|W=T) = \frac{P(W=T|R=T)P(R=T)}{P(W=T)} = \frac{0.8019*0.2}{0.44888} = 0.3573 \tag{25}
$$

The probability that it is raining given that the grass is wet is thus 0.3573.

(b)

We want to find $P(R=T|do(W=T))$. Looking at the graph, if we perform an intervention on W, all the edges connecting W to its predecessors are removed and we obtain:



It follows that the probability of rain is independent from the intervention on W:

$$
P(R=T|do(W=T)) = P(R=T) = 0.2 \tag{26}
$$

(c)

To be a valid adjustment set, Z must contain all nodes that d-separates the sprinkler and the grass being wet. This is true, since Rain is a fork in the path between the sprinkler and the grass being wet.

Otherwise, according to the definition, $Z = \{R\}$ is a valid adjustment set for predicting the effect of sprinkler is turned on on grass being wet if the following holds:

$$P(W|do(S=T)) = \sum_R P(W|R, S=T)P(R) \tag{27}$$

Under the intervention $do(S = T)$, the joint distribution becomes:

$$P(R, W|do(S=T)) = P(W|R, S=T)P(R) \tag{28}$$

By marginalizing $P(R, W|do(S = T))$ over R to get $P(W|do(S = T))$, it is then straightforward to see that Eq. (27) holds and thus $Z = \{R\}$ is a valid adjustment set.