Problem 2

a) 1) $X \quad Y$ 2) $X \to Y$ 3) $X \leftarrow Y$

b) 1) $P(X=x)P(Y=y)$ 2) $P(Y=y|X=x)P(X=x)$ 3) $P(X=x|Y=y)P(Y=y)$

c) 1) $P(Y=y)$ 2) $P(Y=y|X=x)$ 3) $\dfrac{P(X=x|Y=y)P(Y=y)}{\sum_y P(X=x|Y=y)P(Y=y)}$

d) 1) $P(Y=y)$ 2) $P(Y=y|X=x)$ 3) $P(Y=y)$

e) 1) $P(Y=y|X=x) \Rightarrow$ means that we are interested (computing prob) in knowing the probability of the event Y taking value y, given that we observed that X has the value of x. Note that $Y=y$ and $X=x$ are events.

2) $P(Y=y|do(X=x)) \Rightarrow$ means that we are computing the probability of the event $Y=y$, given that we forced X to take the value of x, e.g. forced smoking.

The former predicts the prob. of events given observations, and the latter given active interventions on the system.

Problem 3

1.a. $P(R|D) = \frac{20}{40} = 0.5$ $\quad P(R^c|D) = 0.5$, R - recovery, D - drug usage

$P(R|D^c) = \frac{16}{40} = 0.4$ $\quad P(R^c|D^c) = \frac{24}{40} = 0.6$

1.b. Since $P(R|D) > P(R|D^c)$ a doctor would recommend the drug.

2.a. M - male, M^c - female

$P(R|D,M) = \frac{18}{30} = 0.6$ $\quad P(R^c|D,M) = 0.4$ } males

$P(R|D^c,M) = \frac{7}{10} = 0.7$ $\quad P(R^c|D^c,M) = 0.3$

$P(R|D,M^c) = \frac{2}{10} = 0.2$ $\quad P(R^c|D,M^c) = 0.8$ } females

$P(R|D^c,M^c) = \frac{9}{30} = 0.3$ $\quad P(R^c|D^c,M^c) = 0.7$

2.b. Since $P(R|D,M) < P(R|D^c,M)$ and $P(R|D,M^c) < P(R|D^c,M^c)$ a doctor would not recommend the drug.

3. In hindsight I would not recommend the drug because of 2.b. Yes it contradicts earlier recommendation

The flip of the sign is due to the Simpson's paradox, but let's spend a bit time why does it happen:

$$P(R|D) = P(R|D,M)\,\underline{\underline{P(M|D)}} + P(R|D,M^c)\,\underline{\underline{P(M^c|D)}} = 0.6 \cdot \frac{30}{40} + 0.2 \cdot \frac{10}{40} =$$
$$= 0.5$$

$$P(R|D^c) = P(R|D^c,M)\,\underline{\underline{P(M|D^c)}} + P(R|D^c,M^c)\,\underline{\underline{P(M^c|D^c)}} = 0.7 \cdot \frac{10}{40} + 0.3 \cdot \frac{30}{40} =$$
$$= 0.4.$$

From [Simpson's paradox analysis] we know that the cause of the flip can be the "weights" ($P(M|D^c)$, etc.), let's $P(M|D) = P(M^c|D) = P(M|D^c) = P(M^c|D^c) = \frac{1}{2}$ and run the same calculations:

$$P(R|D) = 0.6 \cdot 0.5 + 0.2 \cdot 0.5 = 0.3 \cdot 0.1 = 0.4$$
$$P(R|D^c) = 0.7 \cdot 0.5 + 0.2 \cdot 0.5 = 0.5$$

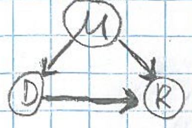
Thus, $P(R|D) < P(R|D^c)$ and $P(R|D,M) < P(R|D^c,M)$
$P(R|D,M^c) < P(R|D^c,M^c)$

Problem 3 (skipped part)

6.d. $P(R|D) = 0.5$ and $P(R|D^c) = 0.4$ (from 1.a)

Thus, a doctor would recommend the drug.

4a.



$\Rightarrow \overline{M}_D$

We are interested in D affect on R, so we have a path $R \leftarrow M \rightarrow D$ which can be blocked by M. Thus,

$$p(R|do(D)) = \sum_{M} p(R|M,D)\,p(M)$$

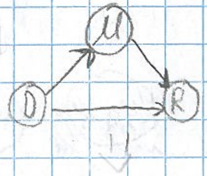
4.b.
$$p(R|D) = \sum \frac{P(D,R)}{P(D)} = \frac{\sum_{M} p(D|M)\,p(R|D,M)\,p(M)}{\sum_{R}\sum_{M} p(D|M)\,p(R|D,M)\,p(M)}$$
, so it's not the same.

4.c.
$$p(R|do(D)) = P(R|M,D)\,p(M) + p(R|M^c,D)\,p(M^c) =$$
$$= 0.6 \cdot \frac{1}{2} + 0.2 \cdot \frac{1}{2} = 0.3 + 0.1 = 0.4$$

$$P(R|do(D^c)) = P(R|M,D^c)\,p(M) + p(R|M^c,D^c)\,p(M^c) = 0.7 \cdot \frac{1}{2} + 0.3 \cdot \frac{1}{2} = 0.5$$

So I would not recommend the drug.

5a.



Again we exploring the affect on D on R, and we have no incomings edges to D, so $\emptyset$ is admissible only.

$$p(R|do(D)) = \sum_{M} p(R|D,M)\,p(M|D)$$

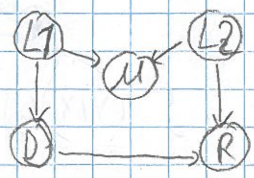
5b. $p(R|D) = \sum_{M} p(R|D,M)\,p(M|D)$ so it's the same

5.c.
$$p(R|do(D)) = p(R|D,M)\,p(M|D) + p(R|D,M^c)\,P(M^c|D) = 0.6 \cdot \frac{30}{40} + 0.2 \cdot \frac{10}{40} = 0.5$$
$$P(R|do(D^c)) = P(R|D^c,M)\,p(M|D^c) + P(R|D^c,M^c)\,p(M^c|D^c) = 0.4$$

So as $p(R|do(D)) > p(R|do(D^c))$ I would recommend the drug.

6.a. We shall assume that M is the blood preasure, and L1 - general health status e.g. $L1 = l_1$, where $l_1 \in \{Bad, average, good\}$, and $L2$ = genetics, i.e. people with good genes have lower blood preasure and recover faster apriori.

6.b



$S = \emptyset$ is the backdoor.

Here we have a situation where the back-door path $R \leftarrow L2 \rightarrow M \leftarrow L1 \rightarrow D$ is blocked by V-structure, thus we can use $\emptyset$ as our admissible node. see slides p 47. Which is a special case.

$$p(R|do(D)) = P(R|D)$$

6.c.
$$P(R|D) = \frac{P(R,D)}{P(D)} = \frac{\sum_{l_2}\sum_{l_1} p(R|L2=l_2,D)\,p(D|L1=l_1)\,p(L1=l_1)\,p(L2=l_2)}{\sum_{R}\sum_{l_2}\sum_{l_1} p(R|L2=l_2,D)\,p(D|L1=l_1)\,p(L1=l_1)\,p(L2=l_2)}$$

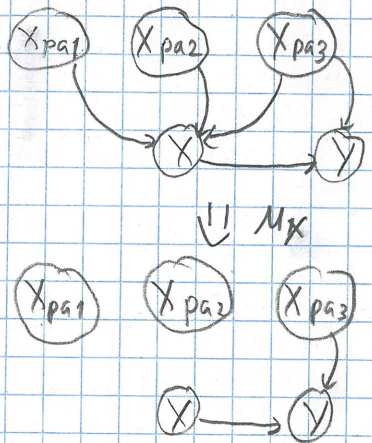
Problem 5

1) $$P(Y|do(X), \underline{X}_{pa(x)}) = \frac{P(Y, \underline{X}_{pa(x)} | do(X))}{P(\underline{X}_{pa(x)} | do(X))} = \frac{\prod_{i \in \underline{X}_{pa(x)} \cup Y} P(x_i | \underline{X}_{pa(x_i)})}{\prod_{i \in \underline{X}_{pa(x)}} P(x_i | \underline{x}_{pa(x_i)})} =$$

$$= \frac{P(Y| \underline{X}_{pa(x)}, X) \cancel{\prod_{i \in \underline{X}_{pa(x)}} P(x_i | \underline{X}_{pa(x_i)})}}{\cancel{\prod_{i \in \underline{X}_{pa(x)}} P(x_i | \underline{X}_{pa(x_i)})}} = P(Y| \underline{X}_{pa(x)}, X)$$

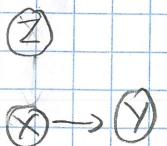
Note that Y can be in general indep. of $\underline{X}_{pa(x)}$ and X, but then we just rewrite it from the equation we obtained. In the third step we've been a perfect intervention rule from the slides p. 37.

2) To explore our findings let's look into an example with binary variables.



$$P(Y|do(X), X_{pa1}, X_{pa2}, X_{pa3}) = P(Y|X, \cancel{X_{pa1}}, \cancel{X_{pa2}}, X_{pa3}) =$$
$$= P(Y|X, X_{pa3})$$, here we used the independence property.

2) To prove the equality we shall ask two subsequent questions, where the first one is „why is it valid to marginalize over nodes that potentially do not influence Y?" i.e. $\underline{X}\,pa(x)$. Let's analyze an example, where Y is not influenced by Z.



(binary vars. graph)

$$P(Y=y) = \sum_x P(Y=y, X=x) = \sum_x P(Y=y|X=x) P(X=x)$$

$$P(Y=y) = \sum_x \sum_z P(Y=y, X=x, Z=z) = \sum_x \sum_z P(Y=y|X=x) P(X=x) P(Z=z)$$

$$= \sum_x P(Y=y|X=x) P(X=x) \sum_z P(Z=z) = \sum_x P(Y=y|X=x) P(X=x)$$

Thus, even though we marginalized over a node that did not affect nor Y neither X, we arrived to the valid result. Notice that the graph is close to what we would get by performing $do(X)$.

Now let's step back to the initial problem, as we know already about the marginalization.

$$P(Y=y|do(X=x)) = \int P(Y=y, \underline{X}_{pa(x)} = \underline{x}' | do(X=x)) \, d\underline{x}' = \int P(Y=y|\underline{X}_{pa(x)} = x', do(X=x)) \times P(\underline{X}_{pa(x)} = x' | do(X=x)) \, d\underline{x}'$$

From (1) we know that $P(Y=y| \underline{X}_{pa(x)}, do(X=x)) =$
$= P(Y=y| \underline{X}_{pa(x)}, X=x)$.

Now we are 1 step away from having the proof, we have to prove only that
$$P(\underline{X}_{pa(x)} | do(X=x)) = P(\underline{X}_{pa(x)})$$

The proof is very simple $P(\underline{X}_{pa(x)} | do(X)) = P(\underline{X}_{pa(x)})$, because $\underline{X}_{pa(x)}$ are the causes and X is the effect., and by forcing the effect, we do not affect the probability of the causes.

Let's take an example; Y - rooster crow, X - sun rises, $X \rightarrow Y$

if we force the rooster to crow it should not affect by any means the sun rise.

Thus, we have proved that:

$$P(Y=y \mid do(X=x)) = \int P(Y=y \mid \underline{X}_{pa(x)} = x', X=x)\, P(\underline{X}_{pa(x)})\, d\underline{x}'$$

Simpon's paradox analysis

Let R be an event of the recovery, D - drug has been used, M - male and we have: $P(R|D) > P(R|D^c)$, let's expand both sides using the sum rule.

$$P(R|D,M)P(M|D) + P(R|D,M^c)P(M^c|D) > P(R|D^c,M)P(M|D^c) + P(R|D^c,M^c)P(M^c|D^c)$$

and let's assume that $P(M|D) = P(M^c|D) = \frac{1}{2}$ and $P(M^c|D^c) = P(M|D^c) = \frac{1}{2}$

so we get:

$$P(R|D,M)\cancel{\tfrac{1}{2}} + P(R|D,M^c)\cancel{\tfrac{1}{2}} > P(R|D^c,M)\cancel{\tfrac{1}{2}} + P(R|D^c,M^c)\cdot\cancel{\tfrac{1}{2}}$$

Now let's assume that we still have Simpson's paradox, thus:

$P(R|D,M) < P(R|D^c,M)$ and $P(R|D,M^c) < P(R|D^c,M^c)$, which we can write as:

$a < b$ and $c < d$, but $a + c > b + d$

This is clearly impossible, and thus Simpon's paradox has been canceled by setting the "priors" (weights) to have an equal probability of $\frac{1}{2}$.