

Q1

1)

a

b

c

d

2)

a

b

c

d

3)

a

b

c

d

4)

a

b

c

d

Q2 a) $i(t_1) = \frac{D}{35} \times \frac{45}{35} = \frac{450}{3025} = \frac{18}{121}$

b) $g(t_1) = \frac{R(t_1) - R(T_1)}{|T_1| - 1} = \frac{\frac{1}{4} - 0}{4 - 1} = \frac{\frac{1}{4}}{3} = \frac{1}{12} \times \frac{1}{5} = \frac{1}{60} = \frac{2}{120} = 0,01\bar{6}$

$g(t_2) = \frac{R(t_2) - R(T_2)}{|T_2| - 1} = \frac{\frac{1}{6} - 0}{2 - 1} = \left(\frac{1}{20}\right) = 0,05$

$g(t_3) = \frac{R(t_3) - R(T_3)}{|T_3| - 1} = \frac{\frac{1}{2} - 0}{2 - 1} = \frac{1}{2} = 0,5$

$g(t_4) = \frac{R(t_4) - R(T_4)}{|T_4| - 1} = \frac{\frac{1}{3} - \frac{1}{6}}{3 - 1} = \frac{\frac{1}{6}}{2} = \frac{1}{12} = \frac{1}{6} \times \frac{2}{2} = 0,1\bar{6}$

$g(t_5) = \frac{R(t_5) - R(T_5)}{|T_5| - 1} = \frac{\frac{1}{10} - 0}{2 - 1} = 0,05$

Let SMS denote the smallest minimising subtree.

1. $T_1 = T_{root}$ does not appear to be a SMS for any a (the minimising subtree is T_2 for $a \in (0, 0,05)$ and T_3 for $a \in (0,05, 0,5)$).
2. Prune on t_2 to obtain T_2 , which is the SMS for $a \in (0, 0,05)$.
3. The root node is the SMS for $a \in (0,05, \infty)$.

c)

Q3 a) Level 1

itemset	support	frequent	→ generator
A	3	✓	A
B	5	✓	B
C	5	✓	C
D	3	✓	D
E	1	X	(E is pruned as it's not frequent)

Level 2

itemset	support	frequent	→ generator
AB	3	✓	(AB is pruned, subset with the same support)
AC	3	✓	(AC is pruned, subset with the same support)
AD	0	X	(AD is pruned, it is infrequent)
BC	5	✓	(BC is pruned, subset with the same support)
BD	2	✓	BD
CD	2	✓	CD

Level 3

itemset	support	frequent	→ generator
BCD	2	✓	(BCD is pruned, subset with the same support)

b) Generator → closure

itemset	support	closure	support
A	3	AB, AC	3
B	5	BC	5
C	5	BC	5
D	3	D	3
BD	2	BD, BCD	2
CD	2	BCD	2

⇒

- Q4 a) Yes. In the moral graph "physical" and "bp" are not connected by an edge as they are not parents.
- b) No. In the moral graph "physical" and "bp" are connected by an edge as they are both parents of "mental", thus married.
- c) Calculate "lpo" contribution to the log-likelihood score for the updated model.

$$28 \ln \frac{224}{1130} + 406 \ln \frac{106}{1130} + 337 \ln \frac{337}{211} + 371 \ln \frac{274}{24} = -1229.77 \approx -1230$$

The change in log-likelihood score = $-1230 + 1208 = 22$

d) Initial model = . . .

Q5 a) $\hat{P}(\text{good} | \text{Positive}) = \frac{2+1}{6+10} = \frac{3}{16}$

$\hat{P}(\text{good} | \text{Negative}) = \frac{0+1}{2+10} = \frac{1}{12}$

$\hat{P}(\text{teacher} | \text{Positive}) = \frac{1+1}{6+10} = \frac{2}{16}$

$\hat{P}(\text{teacher} | \text{Negative}) = \frac{2+1}{2+10} = \frac{3}{12}$

b) $\hat{P}(\text{Positive}) \hat{P}(\text{good teacher} | \text{Positive}) = \frac{1}{2} \cdot \frac{3}{16} = \frac{3}{32} = \frac{1 \cdot 3 \cdot 2}{2 \cdot 16} = \frac{3}{32}$