

Task 1:

a. Because A have 8 values, and B_1, \dots, B_{10} have 5 possible values.

$$\therefore \text{the number of the joint distribution is } 8^1 \cdot 5^{10} - 1 = 78,124,999$$

these parameters need to sum to one, so we can drop one parameter.

b. Because each B_i is conditionally independent of all other 9 B_j variables given A.

And the use of conditional independence reduce the size of the representation of the joint distribution from exponential in n to linear in n.

$$\therefore P(B_j | A, B_i) = P(B_j | A) \Rightarrow P(A, B_1, B_2, \dots, B_n) = P(A) \cdot P(B_1 | A) \cdot P(B_2 | A) \cdots P(B_n | A)$$

After saving space via conditional independence relations, we need $8 \cdot 10 \cdot 5 - 1 = 399$ numbers.

c. Yes, because given A, then each B has the property of conditionally independent. We obtain

$$P(B_j | A, B_i) = P(B_j | A) \Rightarrow P(A, B_1, B_2, \dots, B_n) = P(A) \cdot P(B_1 | A) \cdot P(B_2 | A) \cdots P(B_n | A) \\ = P(A) \cdot \prod_i P(B_i | A), \text{ which is a naive bayes model.}$$

Task 2:

a. 0. $p_o(h_1) = 0.1, p_o(h_2) = 0.2, p_o(h_3) = 0.4, p_o(h_4) = 0.2, p_o(h_5) = 0.1$

C 1. $p_o(q_1 = C) = \sum_{i=1}^5 (p(q_i = C | h_i) \cdot P(h_i)) = 1 \cdot 0.1 + 0.75 \cdot 0.2 + 0.4 \cdot 0.5 + 0.2 \cdot 0.25 + 0 = 0.5$

C 2. $p_1(h_1) = \frac{p(q_1 = C | h_1) \cdot P_o(h_1)}{p_o(q_1 = C)} = \frac{1 \cdot 0.1}{0.5} = 0.2 \quad p_1(h_2) = \frac{0.75 \cdot 0.2}{0.5} = 0.3$

$$p_1(h_3) = \frac{0.5 \cdot 0.4}{0.5} = 0.4 \quad p_1(h_4) = \frac{0.25 \cdot 0.1}{0.5} = 0.1 \quad p_1(h_5) = 0$$

C 3. $p_1(q_2 = C) = 1 \cdot 0.2 + 0.75 \cdot 0.3 + 0.5 \cdot 0.4 + 0.25 \cdot 0.1 = 0.65$

$$2. p_2(h_1) = \frac{1 \cdot 0.2}{0.65} = 0.308 \quad p_2(h_2) = \frac{0.75 \cdot 0.3}{0.65} = 0.346 \quad p_2(h_3) = \frac{0.5 \cdot 0.4}{0.65} = 0.308$$

$$p_2(h_4) = \frac{0.25 \cdot 0.1}{0.65} = 0.038 \quad p_2(h_5) = 0$$

C 4. $p_2(q_3 = C) = 1 \cdot 0.308 + 0.75 \cdot 0.346 + 0.5 \cdot 0.308 + 0.25 \cdot 0.038 + 0 = 0.731$

$$3. p_3(h_1) = \frac{1 \cdot 0.308}{0.731} = 0.421 \quad p_3(h_2) = \frac{0.75 \cdot 0.346}{0.731} = 0.355$$

$$p_3(h_3) = \frac{0.5 \cdot 0.308}{0.731} = 0.211 \quad p_3(h_4) = \frac{0.25 \cdot 0.038}{0.731} = 0.013 \quad p_3(h_5) = 0$$

$$p_3(q_4 = C) = 1 \cdot 0.421 + 0.75 \cdot 0.355 + 0.5 \cdot 0.211 + 0.25 \cdot 0.013 = 0.796$$

4. $p_4(h_1) = \frac{1 \cdot 0.42}{0.796} = 0.529$ $p_4(h_2) = \frac{0.75 \cdot 0.355}{0.796} = 0.334$
 $p_4(h_3) = \frac{0.5 \cdot 0.211}{0.796} = 0.133$ $p_4(h_4) = \frac{0.25 \cdot 0.013}{0.796} = 0.004$ $p_4(h_5) = 0$
 $p_4(q_5=L) = 1 \cdot 0.529 + 0.75 \cdot 0.334 + 0.5 \cdot 0.133 + 0.25 \cdot 0.004 = 0.847$

5. $p_5(h_1) = \frac{0.529}{0.847} = 0.625$ $p_5(h_2) = \frac{0.75 \cdot 0.334}{0.847} = 0.296$ $p_5(h_3) = \frac{0.5 \cdot 0.133}{0.847} = 0.079$
 $p_5(h_4) = \frac{0.25 \cdot 0.004}{0.847} = 0.001$ $p_5(h_5) = 0$
 $p_5(q_6=L) = 1 \cdot 0.625 + 0.75 \cdot 0.296 + 0.5 \cdot 0.079 + 0.25 \cdot 0.001 = 0.8859$
 $\therefore p_5(q_6=L) = 0 \cdot 0.625 + 0.25 \cdot 0.296 + 0.5 \cdot 0.079 + 0.75 \cdot 0.001 + 0 = 0.1140$

6. $p_6(h_1) = 0$ $p_6(h_2) = \frac{0.25 \cdot 0.296}{0.1140} = 0.649$ $p_6(h_3) = \frac{0.5 \cdot 0.079}{0.1140} = 0.342$
 $p_6(h_4) = \frac{0.75 \cdot 0.001}{0.1140} = 0.008$ $p_6(h_5) = 0$

b. 0. $p_0(h_1) = 0.1$, $p_0(h_2) = 0.2$, $p_0(h_3) = 0.4$, $p_0(h_4) = 0.2$, $p_0(h_5) = 0.1$
 $p_0(q_1=C) = \sum_{i=1}^5 (P(q_i=C|h_0) \cdot P(h_i)) = 1 \cdot 0.1 + 0.75 \cdot 0.2 + 0.4 \cdot 0.5 + 0.2 \cdot 0.25 + 0 = 0.5$

1. $p_1(h_1) = \frac{P(q_1=C|h_1) \cdot P_0(h_1)}{P_0(q_1=C)} = \frac{1 \cdot 0.1}{0.5} = 0.2$ $p_1(h_2) = \frac{0.75 \cdot 0.2}{0.5} = 0.3$
 $p_1(h_3) = \frac{0.5 \cdot 0.4}{0.5} = 0.4$ $p_1(h_4) = \frac{0.25 \cdot 0.2}{0.5} = 0.1$ $p_1(h_5) = 0$
 $p_1(q_2=L) = 0 \cdot 0.2 + 0.25 \cdot 0.3 + 0.5 \cdot 0.4 + 0.75 \cdot 0.1 + 1 \cdot 0 = 0.35$

2. $p_2(h_1) = \frac{0 \cdot 0.2}{0.35} = 0$ $p_2(h_2) = \frac{0.25 \cdot 0.3}{0.35} = 0.214$ $p_2(h_3) = \frac{0.5 \cdot 0.4}{0.35} = 0.571$
 $p_2(h_4) = \frac{0.75 \cdot 0.1}{0.35} = 0.214$ $p_2(h_5) = 0$
 $p_2(q_3=C) = 1 \cdot 0 + 0.75 \cdot 0.214 + 0.5 \cdot 0.571 + 0.25 \cdot 0.214 + 0 = 0.5$

3. $p_3(h_1) = 1 \cdot 0 / 0.5 = 0$ $p_3(h_2) = \frac{0.75 \cdot 0.214}{0.5} = 0.321$
 $p_3(h_3) = \frac{0.5 \cdot 0.571}{0.5} = 0.571$ $p_3(h_4) = \frac{0.25 \cdot 0.214}{0.5} = 0.107$ $p_3(h_5) = 0$
 $p_3(q_4=L) = 0 \cdot 0 + 0.25 \cdot 0.321 + 0.5 \cdot 0.571 + 0.75 \cdot 0.107 + 1 \cdot 0 = 0.446$

4. $p_4(h_1) = 0 \cdot 0 / 0.446 = 0$ $p_4(h_2) = \frac{0.25 \cdot 0.321}{0.446} = 0.18$
 $p_4(h_3) = \frac{0.5 \cdot 0.571}{0.446} = 0.64$ $p_4(h_4) = \frac{0.75 \cdot 0.107}{0.446} = 0.18$ $p_4(h_5) = 0$
 $p_4(q_5=C) = 1 \cdot 0 + 0.75 \cdot 0.18 + 0.5 \cdot 0.64 + 0.25 \cdot 0.18 = 0.5$

$$5. p_s(h_1) = 1.0 / 0.5 = 0 \quad p_s(h_2) = 0.75 \cdot 0.18 / 0.5 = 0.27 \quad p_s(h_3) = 0.5 \cdot 0.64 / 0.5 = 0.64 \\ p_s(h_4) = 0.25 \cdot 0.18 / 0.5 = 0.09 \quad p_s(h_5) = 0$$

$$\therefore p_s(q_6=L) = 0.0 + 0.25 \cdot 0.27 + 0.5 \cdot 0.64 + 0.75 \cdot 0.09 + 0 = 0.455$$

$$6. p_b(h_1) = 0 \quad p_b(h_2) = 0.25 \cdot 0.27 / 0.455 = 0.148 \quad p_b(h_3) = 0.5 \cdot 0.64 / 0.455 = 0.703 \\ p_b(h_4) = 0.75 \cdot 0.09 / 0.455 = 0.148 \quad p_b(h_5) = 0$$

$$C. 0. p_o(h_1) = 0.1, p_o(h_2) = 0.2, p_o(h_3) = 0.4, p_o(h_4) = 0.2, p_o(h_5) = 0.1$$

$$p_o(q_1=C) = \sum_{i=1}^5 (p(q_i=C|h_i) \cdot P(h_i)) = 1 \cdot 0.1 + 0.75 \cdot 0.2 + 0.4 \cdot 0.5 + 0.2 \cdot 0.25 + 0 = 0.5$$

$$1. p_1(h_1) = \frac{P(q_1=C|h_1) \cdot P_o(h_1)}{p_o(q_1=C)} = \frac{1 \cdot 0.1}{0.5} = 0.2 \quad p_1(h_2) = \frac{0.75 \cdot 0.2}{0.5} = 0.3$$

$$p_1(h_3) = \frac{0.5 \cdot 0.4}{0.5} = 0.4 \quad p_1(h_4) = \frac{0.25 \cdot 0.2}{0.5} = 0.1 \quad p_1(h_5) = 0$$

$$p_1(q_2=C) = 1 \cdot 0.2 + 0.75 \cdot 0.3 + 0.5 \cdot 0.4 + 0.25 \cdot 0.1 = 0.65$$

$$2. p_2(h_1) = \frac{1 \cdot 0.2}{0.65} = 0.308 \quad p_2(h_2) = \frac{0.75 \cdot 0.3}{0.65} = 0.346 \quad p_2(h_3) = \frac{0.5 \cdot 0.4}{0.65} = 0.308$$

$$p_2(h_4) = 0.25 \cdot 0.1 / 0.65 = 0.038 \quad p_2(h_5) = 0$$

$$p_2(q_3=C) = 1 \cdot 0.308 + 0.75 \cdot 0.346 + 0.5 \cdot 0.308 + 0.25 \cdot 0.038 + 0 = 0.731$$

$$3. p_3(h_1) = 1 \cdot 0.308 / 0.731 = 0.421 \quad p_3(h_2) = 0.75 \cdot 0.346 / 0.731 = 0.355$$

$$p_3(h_3) = 0.5 \cdot 0.308 / 0.731 = 0.211 \quad p_3(h_4) = 0.25 \cdot 0.038 / 0.731 = 0.013 \quad p_3(h_5) = 0$$

$$p_3(q_4=L) = 0 \cdot 0.421 + 0.25 \cdot 0.355 + 0.5 \cdot 0.211 + 0.75 \cdot 0.013 + 1 \cdot 0 = 0.204$$

$$4. p_4(h_1) = 0 \cdot 0.421 / 0.204 = 0 \quad p_4(h_2) = 0.25 \cdot 0.355 / 0.204 = 0.435$$

$$p_4(h_3) = 0.5 \cdot 0.211 / 0.204 = 0.517 \quad p_4(h_4) = 0.75 \cdot 0.013 / 0.204 = 0.048 \quad p_4(h_5) = 0$$

$$p_4(q_5=L) = 0 \cdot 0 + 0.25 \cdot 0.435 + 0.5 \cdot 0.517 + 0.75 \cdot 0.048 + 0 \cdot 0 = 0.403$$

$$5. p_5(h_1) = 0 \cdot 0 / 0.403 = 0 \quad p_5(h_2) = 0.25 \cdot 0.435 / 0.403 = 0.270 \quad p_5(h_3) = 0.5 \cdot 0.517 / 0.403 = 0.641$$

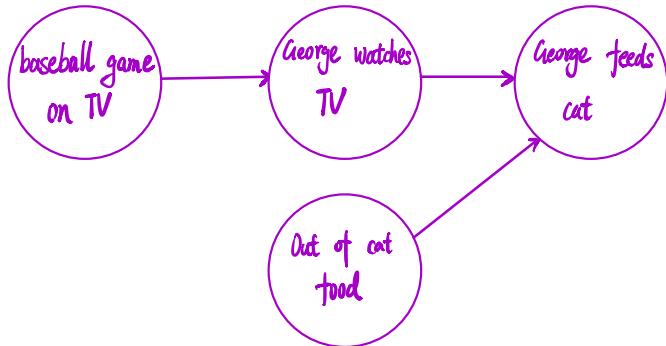
$$p_5(h_4) = 0.75 \cdot 0.048 / 0.403 = 0.089 \quad p_5(h_5) = 0$$

$$\therefore P_S(q_6=1) = 0 \cdot 0 + 0.25 \cdot 0.270 + 0.5 \cdot 0.641 + 0.75 \cdot 0.089 + 0 = 0.455$$

L6. $P_b(h_1) = 0$ $P_b(h_2) = \frac{0.25 \cdot 0.270}{0.455} = 0.148$ $P_b(h_3) = \frac{0.5 \cdot 0.641}{0.455} = 0.704$

$P_b(h_4) = \frac{0.75 \cdot 0.089}{0.455} = 0.146$ $P_b(h_5) = 0$

Task 3:



Task 4 :

$P(\text{baseball_game_on_TV})$	$: \frac{111}{365}$
0.304109	

basket game on TV	$P(\text{George_watches_TV})$
T	0.92793
F	0.11811

$\therefore \frac{103}{111} = \frac{30}{254}$

$P(\text{out_of_cat_food})$	$: \frac{62}{365}$
0.16986	

out_of_cat_food	George_watches_TV	$P(\text{George_feeds_cat})$
T	T	0.04167
T	F	0.31579
F	T	0.70642
F	F	0.95876

$24/1$
 $38/1$
 $109/1$
 $194/1$

```

The probability of P(baseball_game_on_TV) is 0.304109589041
(62L, 365.0)
The probability of P(out_of_cat_food) is 0.169863013699
The probability of P(George_watches_TV|basket_game_on_TV) is 0.927927927928
The probability of P(George_watches_TV|not(basket_game_on_TV)) is 0.405511811024
The probability of P(George_feeds_cat|out_of_cat_food, George_watches_TV) is 0.0416666666667
The probability of P(George_feeds_cat|out_of_cat_food, not(George_watches_TV)) is 0.315789473684
The probability of P(George_feeds_cat|not(out_of_cat_food), George_watches_TV) is 0.706422018349
The probability of P(George_feeds_cat|not(out_of_cat_food), not(George_watches_TV)) is 0.958762886598
  
```

```

import pandas as pd
data = pd.read_csv('./training_data.txt', names=['game', 'watch', 'wtoFood', 'feed'], sep='    ')
denom1 = float(data.shape[0])
pGame = data['game'].sum() / denom1
pWtoFood = data['wtoFood'].sum() / denom1
print(data['game'].sum(), denom1) # 111, 365
print('The probability of P(baseball_game_on_TV) is {}'.format(pGame)) # 0.30410
print(data['wtoFood'].sum(), denom1) # 62, 365
print('The probability of P(out_of_cat_food) is {}'.format(pWtoFood)) # 0.16986

denom2t = float(data[(data['game']==1)].shape[0]) # 111
denom2f = float(data[(data['game']==0)].shape[0]) # 254
watch = data[(data['game']>0)&(data['watch']>0)].sum()['watch'] # 103
pWatchTrue = watch / denom2t # 103/111
pWatchFalse = watch / denom2f # 103/254
print('The probability of P(George_watches_TV|basket_game_on_TV) is {}'.format(pWatchTrue))
print('The probability of P(George_watches_TV|not(basket_game_on_TV)) is {}'.format(pWatchFalse))

denom3tt = float(data[(data['wtoFood']==1)&(data['watch']==1)].shape[0]) # 24
denom3tf = float(data[(data['wtoFood']==1)&(data['watch']==0)].shape[0]) # 38
denom3ft = float(data[(data['wtoFood']==0)&(data['watch']==1)].shape[0]) # 109
denom3ff = float(data[(data['wtoFood']==0)&(data['watch']==0)].shape[0]) # 194
feedTT = data[(data['wtoFood']==1)&(data['watch']==1)&(data['feed']==1)].sum()['feed'] # 1
feedTF = data[(data['wtoFood']==1)&(data['watch']==0)&(data['feed']==1)].sum()['feed'] # 12
feedFT = data[(data['wtoFood']==0)&(data['watch']==1)&(data['feed']==1)].sum()['feed'] # 77
feedFF = data[(data['wtoFood']==0)&(data['watch']==0)&(data['feed']==1)].sum()['feed'] # 186
pFeedTT = feedTT / denom3tt
pFeedTF = feedTF / denom3tf
pFeedFT = feedFT / denom3ft
pFeedFF = feedFF / denom3ff
print('The probability of P(George_feeds_cat|out_of_cat_food, George_watches_TV) is {}'.format(pFeedTT))
print('The probability of P(George_feeds_cat|out_of_cat_food, not(George_watches_TV)) is {}'.format(pFeedTF))
print('The probability of P(George_feeds_cat|not(out_of_cat_food), George_watches_TV) is {}'.format(pFeedFT))
print('The probability of P(George_feeds_cat|not(out_of_cat_food), not(George_watches_TV)) is {}'.format(pFeedFF))

```

Task 5 :

Let $P(\text{not(George Feeds Cat)} \mid \text{Baseball Game on TV})$, take 'g', 'w', 'wto', 'f'.

as abbreviation. We obtain:

$$P(\neg f | g) = P(\neg f, g) / P(g)$$

$$= \frac{\text{count}(\neg f, g)}{\text{count}(g)} / \frac{\text{count}(g)}{\text{count}(\text{days})} = \frac{45}{365} / \frac{111}{365} = \frac{0.12328}{0.30411} = 0.40541$$

Task 6 :

a. Markovian Blanket of node L is: G, P, Q, K, M.

(parents + children + children's parents)

b. $\because P(C)=0.6, P(H|C)=0.6, P(C,H)=P(H|C) \cdot P(C)$

$$\therefore P(C,H)=0.6 \cdot 0.6 = 0.36$$

$$c. P(O|not(J), E) = \frac{P(O, not(J), E)}{P(not(J), E)} = \frac{P(O|not(J)) \cdot P(not(J)|E) \cdot P(E)}{P(E) \cdot P(not(J)|E)}$$
$$= P(O|not(J)) = 0.8$$