Homework 1
Inference and Representation

Aodong Li
NetID: al5350
al5350@nyu.edu

June 9, 2018

1. $P(+|disease) = 0.98$, $P(-|no\,disease) = 0.98$, $P(disease) = 1/40,000$.

$$
\begin{aligned}
P(disease|+) &= \frac{P(+|disease)P(disease)}{P(+)} \\
&= \frac{P(+|disease)P(disease)}{P(+|disease)P(disease) + P(+|no\ disease)P(no\ disease)} \\
&= \frac{0.98/40,000}{0.98/40,000 + 0.02} \\
&= 0.0012 \\
P(disease|-) &= \frac{P(-|disease)P(disease)}{P(-)} \\
&= \frac{P(-|no\ disease)P(no\ disease)}{P(-|no\ disease)P(no\ disease) + P(-|disease)P(disease)} \\
&= \frac{0.02/40,000}{0.98 * 39,999/40,000 + 0.02/40,000} \\
&\approx 0
\end{aligned}
$$

Because the low prior probability makes this situation less likely. In other words, if we want this situation happens, we may need stronger evidence to compensate the low prior probability. The chances that you really have the disease is 0.0012.

2. Since every conditional probability distribution is a valid one, i.e., $\sum_{x_v \in Val(X_v)} f_v(x_v|pa(x_v)) = 1$. If there exist a directed cycle, to see if f defines a valid distribution, we test if

$$\sum_{v\in V} \prod_{v\in V} f_v(x_v|pa(x_v)) = 1.$$

By exchanging the order of summation and product, for any particular node v in the graph, given a cycle, we can no longer get the a single summation term where x_v only appears in the summation position, instead, we get a summation over at least two terms since v must be a parent for some other nodes. This kind of summation does not ensure the summation equals to 1.

In particular, consider two fully connected binary nodes X_1, X_2:

$$\begin{aligned} P(x_1 = 0|x_2 = 0) = 0, P(x_1 = 1|x_2 = 0) = 1, \\ P(x_1 = 0|x_2 = 1) = 1, P(x_1 = 1|x_2 = 1) = 0, \\ P(x_2 = 0|x_1 = 0) = 1, P(x_2 = 1|x_1 = 0) = 0, \\ P(x_2 = 0|x_1 = 1) = 0, P(x_2 = 1|x_1 = 1) = 1. \end{aligned}$$

By the definition of probabilistic graphical model, $f(x_1, x_2) = P(x_1|x_2)P(x_2|x_1)$,

$$\sum_{x_1,x_2} f(x_1, x_2) = 0.$$

Although every conditional distribution is valid, the joint distribution is not a valid one.

3. (a)

$$\begin{gathered} 1 \perp \{2, 3, 5, 7, 8, 9, 10\} \\ 2 \perp \{7, 8\} \\ 3 \perp \{6, 7, 8\} \\ 4 \perp 8 \\ 6 \perp \{7, 8\} \\ 7 \perp 8. \end{gathered}$$

(b) $A = \{3, 5, 7, 8, 10\}$.

4. By calculating the marginal probability, it can be seen that for each variable, the probability of taking 1 or 0 is equally 1/2. 1) Given no evidence, all three variables are dependent because $P(x, y, z) \neq P(x)P(y)P(z)$; 2) Conditioned on any single variable, the other two are not independent $P(x, y|z) \neq P(x|z)P(y|z) = 1/4$. These facts indicate the Bayesian network is fully connected. $I(G) = \emptyset$.

   On the other hand, every two variables are independent because $P(x, y) = P(x)P(y)$. However, the fully connected network can not represent this kind of independence. $I(P) \neq \emptyset$.

5. Check skeleton and v-structure.

   1) If two networks have the same skeleton and v-structure, then the two networks are equivalent. But the reverse is not necessarily true.

   2) If two networks have the same skeleton and immoralities if and only if the two networks are equivalent.