Exercise 9-1  Posterior Distribution

On a day when an assignment is due (A=a), the newsgroup tends to be busy (B=b), and the computer lab tends to be full (C=c). Consider the following conditional probability tables for the domain, where $A \in \{a, \neg a\}$, $B \in \{b, \neg b\}$ and $C \in \{c, \neg c\}$. In this problem, we assume that $B$ and $C$ are conditionally independent given $A$.

\[
\begin{array}{c|cc}
A & b & \neg b \\
\hline
a & 0.8 & 0.2 \\
\neg a & 0.3 & 0.7 \\
\end{array}
\quad
\begin{array}{c|cc}
C & c & \neg c \\
\hline
a & 0.4 & 0.6 \\
\neg a & 0.2 & 0.8 \\
\end{array}
\]

- Construct the full joint distribution $P(A,B,C)$ out of these conditional probabilities. Note that this is only possible because of the assumption that $B$ and $C$ are conditionally independent given $A$.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>P(A,B,C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>0.032</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>\neg c</td>
<td>0.048</td>
</tr>
<tr>
<td>a</td>
<td>\neg b</td>
<td>c</td>
<td>0.008</td>
</tr>
<tr>
<td>a</td>
<td>\neg b</td>
<td>\neg c</td>
<td>0.012</td>
</tr>
<tr>
<td>\neg a</td>
<td>b</td>
<td>c</td>
<td>0.054</td>
</tr>
<tr>
<td>\neg a</td>
<td>b</td>
<td>\neg c</td>
<td>0.216</td>
</tr>
<tr>
<td>\neg a</td>
<td>\neg b</td>
<td>c</td>
<td>0.126</td>
</tr>
<tr>
<td>\neg a</td>
<td>\neg b</td>
<td>\neg c</td>
<td>0.504</td>
</tr>
</tbody>
</table>

- What is the marginal distribution $P(B,C)$? Are these two variables independent in this distribution? Justify your answer using the actual probabilities, not your intuitions.
Possible Solution:

\[
P(B, C) = \begin{array}{c|cc}
   & c & \sim c \\
 b & 0.086 & 0.264 \\
 \sim b & 0.134 & 0.516 \\
\end{array}
\]

\[
P(B = b) = 0.35, P(C = c) = 0.22
\]
\[
P(B = b, C = c) = 0.086 
eq P(B = b)P(C = c) = 0.077
\]

So \(B, C\) are not independent.

• What is the posterior distribution over \(A\) given that \(B = b\), that is \(P(A \mid B = b)\)?

Possible Solution:

\[
P(A = a \mid B = b) = \frac{P(A = a, B = b)}{P(B = b)} = \frac{0.08}{0.35} = 0.229
\]
\[
P(A = \sim a \mid B = b) = 0.771
\]

• What is the posterior distribution over \(A\) given that \(B = b\) and \(C = c\), that is \(P(A \mid B = b, C = c)\)?

Possible Solution:

\[
P(A = a \mid B = b, C = c) = \frac{P(A = a, B = b, C = c)}{P(B = b, C = c)} = \frac{0.032}{0.086} = 0.372
\]
\[
P(A = \sim a \mid B = b, C = c) = 0.628
\]

• Briefly explain why the pattern amongst \(P(A)\), \(P(A \mid B = b)\), and \(P(A \mid B = b, C = c)\) makes intuitive sense.

Possible Solution:

\[
P(A = a) < P(A = a \mid B = b) < P(A = a \mid B = b, C = c)
\]

Observing more evidence increases belief in \(A = a\).
Consider the following network, in which a mouse agent is reasoning about the behavior of a cat. The mouse really wants to know whether the cat will attack (A), which depends on whether the cat is hungry (H) and whether the cat is sleepy (S). The mouse can observe two things, whether the cat is sleepy (S) and whether the cat has a collar (C). The cat is more often sleepy (S) when it’s either full (f) or starved (v) than when it is peckish (p) and the collar (C) tends to indicate that the cat is not starved. Note that entries are omitted, such as $P(C = \neg c)$, when their complements are given.

\[
\begin{array}{c|c|c}
C & P(C) & \\
\hline
f & c & 0.7 \\
v & c & 0.1 \\
p & c & 0.2 \\
f & \neg c & 0.2 \\
v & \neg c & 0.5 \\
p & \neg c & 0.3 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
H & C & P(S | H) & P(A | H,S) \\
\hline
s & f & 0.9 & a f s 0.01 \\
v & s & 0.6 & a f \neg s 0.1 \\
p & s & 0.3 & a v s 0.4 \\
v & \neg s & 0.9 & a v \neg s 0.9 \\
p & \neg s & 0.7 & a p s 0.2 \\
p & \neg s & 0.7 & a p \neg s 0.7 \\
\end{array}
\]

(a) Draw the Bayesian network corresponding to the above joint probability distribution on $C,H,S,A$.

Possible Solution:

(b) Compute the following probabilities:

- $P(A=a, C=c, S=s, H=f)$
- $P(A=a, C=c, S=s)$
- $P(C=c, S=s)$
- $P(A=a \mid C=c, S=s)$

Possible Solution:

- $P(A = a, C = c, S = s, H = f) = A \cdot .7 \cdot .9 \cdot .01 = .00252$
- $P(A=a, C=c, S=s)$
  
\[
P(A = a, C = c, S = s, H = f) + P(A = a, C = c, S = s, H = v) + P(A = a, C = c, S = s, H = p) = 
(0.4 \cdot 0.7 \cdot 0.9 \cdot 0.01) + (0.4 \cdot 0.1 \cdot 0.6 \cdot 0.4) + (0.4 \cdot 0.2 \cdot 0.3 \cdot 0.2) = 0.01692
\]

- $P(C=c, S=s)$
  
\[
P(C = c, S = s, H = f) + P(C = c, S = s, H = v) + P(C = c, S = s, H = p) = 
(0.4 \cdot 0.7 \cdot 0.9) + (0.4 \cdot 0.1 \cdot 0.6) + (0.4 \cdot 0.2 \cdot 0.3) = 0.3
\]

- $P(A = a \mid C = c, S = s) = \frac{P(A=a,C=c,S=s)}{P(C=c,S=s)} = \frac{0.01692}{0.3} = 0.0564$

(c) The mouse is trying to figure out whether it should run out its hole and eat the cheese (E) or do nothing (N). If the mouse hides, nothing happens but it stays hungry. If the mouse runs out to eat the cheese and
the cat attacks, the mouse dies (which has a low utility). Otherwise, if the mouse tries to eat the cheese and the cat does not attack, it gets to eat tasty cheese (high utility).

<table>
<thead>
<tr>
<th>Utilities</th>
<th>Cat ready to attack (A)</th>
<th>Mouse’s action</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>E</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>¬a</td>
<td>E</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>N</td>
<td>-2</td>
<td></td>
</tr>
</tbody>
</table>

- Suppose in the above table $x$ is $-10$. The mouse sees that the cat has a collar on and is sleepy. What is the utility of trying to eat the cheese? What about doing nothing? Which option should the mouse choose?
- What should the utility of dying ($x$ in the above table) be in order for the mouse to be ambivalent between running for the cheese and doing nothing? Again, the cat is wearing a collar and is sleepy.

**Possible Solution:**

- The utility of doing nothing is $-2$. The utility of going for the cheese is $P(A = a \mid C = c, S = s) \times x + P(A = \neg a \mid C = c, S = s) \times 5 = 4.154$. The mouse should go for the cheese.
- We need to find an $x$ such that the utility of going for the food is equal to the utility of doing nothing. Thus, $x$, must solve

$$P(A = a \mid C = c, S = s) \times x + P(A = \neg a \mid C = c, S = s) \times 5 = -2.$$  

$$\rightarrow x = -119.113.$$  

(d) You may have noticed that one of the variables in the network is “collar”, which according to the CPTs (conditional probability table) causes hunger. However, the real relationship is of correlation, not causation. Introduce a new node $O$ for “owner” and draw a network which better models the true relationship between the variables. $C$ and $H$ should be independent conditioned on $O$.

**Possible Solution:**
Exercise 9-3  Markov Property

For each of the following definitions of the state $X_k$ at time $k$ (for $k = 1,2,\ldots$), determine whether the Markov property is satisfied by the sequence $X_1,X_2,\ldots$

A fair six-sided die (with sides labeled 1,2,3,4,5,6) is rolled repeatedly and independently.

(a) Let $X_k$ denote the largest number obtained in the first $k$ rolls. Does the sequence $X_1,X_2,\ldots$ satisfy the Markov property?

(b) Let $X_k$ denote the number of 6’s obtained in the first $k$ rolls, up to a maximum of ten. (That is, if ten or more 6’s are obtained in the first $k$ rolls, then $X_k = 10$.) Does the sequence $X_1,X_2,\ldots$ satisfy the Markov property?

(c) Let $Y_k$ denote the result of the $k^{th}$ roll. Let $X_1 = Y_1$ and for $k \geq 2$, let $X_k = Y_k + Y_{k-1}$. Does the sequence $X_1,X_2,\ldots$ satisfy the Markov property?

(d) Let $Y_k = 1$ if the $k^{th}$ roll results is an odd number; and $Y_k = 0$ otherwise. Let $X_1 = Y_1$ and for $k \geq 2$ let $X_k = Y_k \cdot X_{k-1}$. Does the sequence $X_1,X_2,\ldots$ satisfy the Markov property?
Possible Solution:

(a) Since the state $X_k$ is the largest number obtained in $k$ rolls, the set of states is $S = \{1,2,3,4,5,6\}$. Given the largest number obtained in the first $k$ rolls, the probability distribution of the largest number obtained in the first $k+1$ rolls no longer depends on what the largest number obtained was in the first $k-1$ rolls (or in the first $k-2$ rolls, etc.). Therefore the Markov property is satisfied.

For $i, j \in \{1,2,3,4,5,6\}$, the transition probabilities are:

$$p_{i,j} = \begin{cases} 
0, & \text{if } j < i \\
\frac{i}{6}, & \text{if } j = i \\
1 - \frac{i}{6}, & \text{if } j > i 
\end{cases}$$

(b) Since the state $X_k$ is the number of 6’s in the first $k$ rolls, the set of states is $S = \{0,1,2,\ldots,10\}$. The probability of getting a 6 in a given trial is $\frac{1}{6}$. Given the number of 6’s in the first $k$ rolls, the probability distribution of the number of 6’s in the first $k+1$ rolls no longer depends on the number of 6’s in the first $k-1$ rolls (or in the first $k-2$ rolls, etc.). Therefore the Markov property is satisfied. Thus $p_{10,10} = 1$, and for $i \leq 9$, the transition probabilities are:

$$p_{i,j} = \begin{cases} 
\frac{1}{6}, & \text{if } j = i + 1 \\
\frac{5}{6}, & \text{if } j = i \\
0, & \text{otherwise} 
\end{cases}$$

(c) We have:

$$P(X_3 = 2 \mid X_2 = 3, X_1 = 1) = P(Y_2 + Y_3 = 2 \mid Y_1 = 1, Y_2 = 2)$$
$$P(Y_3 = 0 \mid Y_1 = 1, Y_2 = 2) = 0$$

but

$$P(X_3 = 2 \mid X_2 = 3, X_1 = 2) = P(Y_2 + Y_3 = 2 \mid Y_1 = 2, Y_2 = 1)$$
$$P(Y_3 = 1 \mid Y_1 = 2, Y_2 = 1) = P(Y_3 = 1) = 1/6$$

and therefore the Markov property is violated.

(d) At each stage, $Y_k$ has equal probability of being 0 or 1. Since $X_k = Y_k \cdot X_{k-1}$, and we assume independent rolls, clearly $X_k$ depends only on the $k^{th}$ roll and the value of $X_{k-1}$. Therefore the Markov property is satisfied.

The transition probabilities are $p_{00} = 1$, $p_{01} = 0$, $p_{10} = \frac{1}{2}$, and $p_{11} = \frac{1}{2}$.