This note includes simple questions on probabilities in joint distributions and Bayesian networks that are useful for self study. Please feel free to consult me on these. (You are not expected to hand in solutions.)

1 Joint Distributions

The management committee of the carnival in Brazil decided to gather statistics on participants in the performances. The scheme this year prescribed that each person who wanted to participate put forward an application through one of the many dance schools in the country. Applications were then examined individually and a decision made in each case.

A pilot study was started by recording 3 properties for each applicant:

- $P$ is a Boolean variable recording whether the applicant participated in the carnival in the previous year.
- $S$ is the size of the school through which the application was put forward; possible values recorded were small, medium, large, denoted below as $s, m, l$.
- $D$ is the decision, 1 for accepted and 0 otherwise.

The distribution over applicants is summarized in the table.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$S$</th>
<th>$D$</th>
<th>$\Pr{}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$s$</td>
<td>0</td>
<td>0.07595</td>
</tr>
<tr>
<td>0</td>
<td>$s$</td>
<td>1</td>
<td>0.07905</td>
</tr>
<tr>
<td>0</td>
<td>$m$</td>
<td>0</td>
<td>0.1475</td>
</tr>
<tr>
<td>0</td>
<td>$m$</td>
<td>1</td>
<td>0.1475</td>
</tr>
<tr>
<td>0</td>
<td>$l$</td>
<td>0</td>
<td>0.015</td>
</tr>
<tr>
<td>0</td>
<td>$l$</td>
<td>1</td>
<td>0.035</td>
</tr>
<tr>
<td>1</td>
<td>$s$</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>1</td>
<td>$s$</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>1</td>
<td>$m$</td>
<td>0</td>
<td>0.1825</td>
</tr>
<tr>
<td>1</td>
<td>$m$</td>
<td>1</td>
<td>0.1825</td>
</tr>
<tr>
<td>1</td>
<td>$l$</td>
<td>0</td>
<td>0.0405</td>
</tr>
<tr>
<td>1</td>
<td>$l$</td>
<td>1</td>
<td>0.0945</td>
</tr>
</tbody>
</table>

Questions: Recall from the slides that $\Pr\{}$ denotes the probability of a particular event and $\Pr()$ denotes the marginal distribution over a variable or a set of variables.

1. Compute $\Pr\{D = 1|P = 0\}$
   (the probability that a new applicant is accepted)

2. Compute $\Pr(D)$

3. Is the event $D = 1$ “the applicant is accepted” independent of the event $P = 0$ “a new applicant”?

4. Is $D$ independent of $P$?

5. Compute $\Pr(D|S = l)$.

6. Compute $\Pr(D|S = l, P)$.
   (For each value $v$ of $P$, compute the conditional distribution $\Pr(D|S = l, P = v)$.)

7. Verify that $D$ conditionally independent of $P$ given $S = l$.

8. In fact, for the distribution we have, $D$ is conditionally independent of $P$ given $S$.
   What would you need to check in order to verify that?
2 Ad Hoc Reasoning in Bayes Networks

Encouraged by the previous study the management committee of the carnival in Brazil decided to gather further statistics on participants in the performances. The scheme this year prescribed that each person who wanted to participate put forward an application through one of the many dance schools in the country. Applications were then examined individually and a decision made in each case.

In the current study 5 properties were recorded for each applicant:

- \( P \) is a Boolean variable recording whether the applicant participated in the carnival in the previous year.
- \( S \) is the size of the school through which the application was put forward; possible values recorded were small, medium, large, denoted below as \( s, m, l \).
- \( D \) is the decision, 1 for accepted and 0 otherwise.
- \( W \) is a Boolean variable recording whether the applicant took a preparation workshop in the school before applying.
- \( C \) records the home city of the applicant (with values: \( a, b, c \)).

Since the joint distribution table is not small it was decided to use a Bayesian Network to represent it. The distribution over these variables was analyzed and found to conform with the network drawn on the right. The conditional distribution tables are as follows.

\[
\begin{array}{c|c}
\text{Pr}\{ W = 0 \} & \text{Pr}\{ W = 1 \} \\
0.3 & 0.7 \\
\end{array}
\]

\[
\begin{array}{c|c}
\text{Pr}\{ P = 0 \} & \text{Pr}\{ P = 1 \} \\
0.5 & 0.5 \\
\end{array}
\]

\[
\begin{array}{c|c|c}
\text{Pr}\{ C = a \} & \text{Pr}\{ C = b \} & \text{Pr}\{ C = c \} \\
0.35 & 0.3 & 0.35 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\text{Pr}\{ S = s \} & \text{Pr}\{ S = m \} & \text{Pr}\{ S = l \} \\
0 & 0 & 0.8 & 0.1 & 0.1 \\
0 & 1 & 0.1 & 0.8 & 0.1 \\
1 & 0 & 0 & 0.8 & 0.2 \\
1 & 1 & 0 & 0.7 & 0.3 \\
\end{array}
\]


<table>
<thead>
<tr>
<th>C</th>
<th>S</th>
<th>Pr(D = 0)</th>
<th>Pr(D = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>s</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>a</td>
<td>m</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>a</td>
<td>l</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>b</td>
<td>s</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>b</td>
<td>m</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>b</td>
<td>l</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>c</td>
<td>s</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>c</td>
<td>m</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>c</td>
<td>l</td>
<td>0.3</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Questions: Recall from the slides that \(\text{Pr}\{\}\) denotes the probability of a particular event and \(\text{Pr}()\) denotes the marginal distribution over a variable or a set of variables.

1. What is the size of the table needed to describe the joint distribution directly?

2. What is the probability that a new applicant \((P = 0)\) from city \(a\) \((C = a)\) who took the preparation workshop \((W = 1)\) is accepted \((D = 1)\)?

3. Compute a table for \(\text{Pr}(S|P)\).

4. Compute a table for \(\text{Pr}(D|S)\).

5. Use the previous two parts to compute a table for \(\text{Pr}(D|P)\).
   
   You may want to use the following formula as the basis for the computation:
   \[
   \text{Pr}\{D = v_1|P = v_2\} = \sum_{v_3} \text{Pr}\{S = v_3|P = v_2\}\text{Pr}\{D = v_1|P = v_2, S = v_3\}
   \]

6. What is the probability that an applicant from city \(a\) who was accepted, participated in the carnival in the previous year? (Compute \(\text{Pr}\{P = 1|C = a \land D = 1\}\))
   
   You can use Bayes’ Rule as in the formula below and use normalization by computing similar value with \(P = 0\) as well.
   \[
   \text{Pr}\{P = 1|D = 1 \land C = a\} = \frac{\text{Pr}\{D = 1|P = 1 \land C = a\}\text{Pr}\{P = 1|C = a\}}{\text{Pr}\{D = 1|C = a\}}
   \]

3 Inference Algorithms

1. Solve the last question of the previous part using variable elimination.

2. Solve the last question of the previous part using the message passing algorithm.

3. Simulate a few steps of Gibbs sampling for the the last question of the previous part.