Solutions: Problem Set #1

(1) The following table gives the joint probability distribution \( p(X, Y) \) of random variables \( X \) and \( Y \).

<table>
<thead>
<tr>
<th>Y</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.02</td>
<td>.04</td>
<td>.08</td>
</tr>
<tr>
<td>2</td>
<td>.03</td>
<td>.18</td>
<td>.04</td>
</tr>
<tr>
<td>3</td>
<td>.04</td>
<td>.04</td>
<td>.08</td>
</tr>
<tr>
<td>4</td>
<td>.09</td>
<td>.18</td>
<td>.18</td>
</tr>
</tbody>
</table>

Determine the following:

(a) Do the entries of the table satisfy the conditions for a bivariate density function?

**ANSWER**: Yes, since all the entries are non-negative, and they sum to unity (one).

(b) The marginal (or unconditional) probability distributions of \( X \) and \( Y \). [Note: These will be a collection of probabilities: the probabilities associated with the 3 values of \( X \) and the probabilities associated with the 4 values of \( Y \).] **ANSWER**

\[
\begin{align*}
\Pr(X = 1) &= .18, \quad \Pr(X = 2) = .44, \quad \Pr(X = 3) = .38 \\
\Pr(Y = 1) &= .14, \quad \Pr(Y = 2) = .25, \quad \Pr(Y = 3) = .16, \quad \Pr(Y = 4) = .45.
\end{align*}
\]

(c) The conditional probability distributions \( p(X|Y = 3) \) and \( p(Y|X = 1) \). (Note: The first conditional probability distribution is the collection of three numbers, \( \Pr(X = 1|Y = 3), \Pr(X = 2|Y = 3), \Pr(X = 3|Y = 3) \).)

**ANSWER**: Applying our formulas for calculating a conditional from a joint:

\[
\begin{align*}
\Pr(X = 1|Y = 3) &= .04/.16 = 1/4, \quad \Pr(X = 2|y = 3) = .04/.16 = 1/4,
\end{align*}
\]
and

\[ \Pr(X = 3|Y = 3) = .08/.16 = 1/2. \]

As for the remaining conditional

\[ \Pr(Y = 1|X = 1) = .02/.18 = 1/9, \quad \Pr(Y = 2|X = 1) = .03/.18 = 1/6 \]

and

\[ \Pr(Y = 3|X = 1) = .04/.18 = 2/9, \quad \Pr(Y = 4|X = 1) = .09/.18 = 1/2. \]

(2) Using straightforward manipulations of the conditional probability (see, e.g., Down’s Syndrome example):

\[
\Pr(D = 1|S = 0) = \frac{\Pr(S = 0|D = 1)\Pr(D = 1)}{\Pr(S = 0)} = \frac{\Pr(S = 0|D = 1)\Pr(D = 1)}{\Pr(S = 0, D = 1) + \Pr(S = 0, D = 0)} = \frac{\Pr(S = 0|D = 1)\Pr(D = 1)}{\Pr(S = 0|D = 1)\Pr(D = 1) + \Pr(S = 0|D = 0)\Pr(D = 0)}.
\]

The second line simply notes that the marginal probability can be obtained by summing over all of the corresponding joint probabilities.

Based on what is given in the problem, \( \Pr(S = 1|D = 1) = .5 \), and therefore \( \Pr(S = 0|D = 1) = .5 \), since these two numbers have to add up to one. Similarly, \( \Pr(D = 1) = .2 \) and thus \( \Pr(D = 0) = .8 \). Finally, \( \Pr(S = 0|D = 0) = 1 \), since if the father does not have the disease, the son cannot have the disease either.

Putting these numbers into the expression above gives

\[ \Pr(D = 1|S = 0) = .5(.2)/[.5(.2) + 1(.8)] = 1/9. \]

Extra Credit
Using similar manipulations, we obtain
\[
\Pr(D = 1|S_1 = 0, S_2 = 0) = \frac{\Pr(S_1 = 0, S_2 = 0|D = 1)\Pr(D = 1)}{\Pr(S_1 = 0, S_2 = 0)}.
\]

What is given in the problem is that the outcome of each son is independent given the disease status of the father. Thus,

\[
\Pr(S_1 = 0, S_2 = 0|D = 1) = \Pr(S_1 = 0|D = 1)\Pr(S_2 = 0|D = 1).
\]

We thus substitute this into the numerator of our expression above, and continue to simplify the denominator, to obtain

\[
\Pr(D = 1|S_1 = 0, S_2 = 0) = \frac{\Pr(S_1 = 0|D = 1)\Pr(S_2 = 0|D = 1)\Pr(D = 1)}{\Pr(S_1 = 0, S_2 = 0, D = 1) + \Pr(S_1 = 0, S_2 = 0, D = 0)}
\]

which simplifies to

\[
\Pr(S_1 = 0|D = 1)\Pr(S_2 = 0|D = 1)\Pr(D = 1)
\]

\[
\frac{\Pr(S_1 = 0|D = 1)\Pr(S_2 = 0|D = 1)\Pr(D = 1)}{\Pr(S_1 = 0|D = 1)\Pr(S_2 = 0|D = 1)\Pr(D = 1) + \Pr(S_1 = 0|D = 0)\Pr(S_2 = 0|D = 0)\Pr(D = 0)}
\]

\[
\frac{.5(.5)(.2)}{.5(.5)(.2) + (1)(1)(.8)},
\]

which equals .059.