1. Using the full joint probability distribution below, write out what the following probability distributions look like. Notice that the P is in boldface to emphasize that these are distributions (i.e., probability tables). This means you have to write out the probability distributions for all uninstantiated random variables e.g., for (a), write out P(Catch=true) and P(Catch=false).

<table>
<thead>
<tr>
<th>Toothache</th>
<th>Cavity</th>
<th>Catch</th>
<th>P(Toothache, Cavity, Catch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>false</td>
<td>false</td>
<td>false</td>
<td>0.576</td>
</tr>
<tr>
<td>false</td>
<td>false</td>
<td>true</td>
<td>0.144</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>0.008</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>true</td>
<td>0.072</td>
</tr>
<tr>
<td>true</td>
<td>false</td>
<td>false</td>
<td>0.064</td>
</tr>
<tr>
<td>true</td>
<td>false</td>
<td>true</td>
<td>0.016</td>
</tr>
<tr>
<td>true</td>
<td>true</td>
<td>false</td>
<td>0.012</td>
</tr>
<tr>
<td>true</td>
<td>true</td>
<td>true</td>
<td>0.108</td>
</tr>
</tbody>
</table>

a) \( P(\text{Catch}) \) [2 points]
\[
P(\text{Catch} = \text{false}) = 0.576 + 0.008 + 0.064 + 0.012 = 0.66
P(\text{Catch} = \text{true}) = 1 - 0.66 = 0.34
\]

b) \( P(\text{Toothache}, \text{Cavity}) \) [4 points]
\[
P(\text{Toothache}=\text{false}, \text{Cavity}=\text{false}) = 0.576 + 0.144 = 0.72
P(\text{Toothache}=\text{false}, \text{Cavity}=\text{true}) = 0.008 + 0.072 = 0.08
P(\text{Toothache}=\text{true}, \text{Cavity}=\text{false}) = 0.064 + 0.016 = 0.08
P(\text{Toothache}=\text{true}, \text{Cavity}=\text{true}) = 0.012 + 0.108 = 0.12
\]
c) \( P(\text{Catch} \mid \text{Cavity}) \) [4 points]
\[
P(\text{Catch} = \text{false} \mid \text{Cavity} = \text{false}) \\
= P(\text{Catch} = \text{false}, \text{Cavity} = \text{false}) / P(\text{Cavity} = \text{false}) \\
= (0.576+0.064)/(0.576+0.144+0.064+0.016) \\
= 0.8
\]
\[
P(\text{Catch} = \text{true} \mid \text{Cavity} = \text{false}) \\
= P(\text{Catch} = \text{true}, \text{Cavity} = \text{false}) / P(\text{Cavity} = \text{false}) \\
= (0.144+0.016)/(0.576+0.144+0.064+0.016) \\
= 0.2
\]
\[
P(\text{Catch} = \text{false} \mid \text{Cavity} = \text{true}) \\
= P(\text{Catch} = \text{false}, \text{Cavity} = \text{true}) / P(\text{Cavity} = \text{true}) \\
= (0.144+0.016)/ (0.144 + 0.072 + 0.016 + 0.108) \\
= 0.471
\]
\[
P(\text{Catch} = \text{true} \mid \text{Cavity} = \text{true}) \\
= P(\text{Catch} = \text{true}, \text{Cavity} = \text{true}) / P(\text{Cavity} = \text{true}) \\
= (0.072 + 0.108)/ (0.144 + 0.072 + 0.016 + 0.108) \\
= 0.529
\]

2. Show that the three forms of independence below are equivalent: [4 points]
\[
P(X,Y) = P(X)P(Y) \quad \ldots(1)
\]
\[
P(X \mid Y) = P(X) \quad \ldots(2)
\]
\[
P(Y \mid X) = P(Y) \quad \ldots(3)
\]

Assume that \( P(X,Y) = P(X)P(Y) \)

We now show that you can derive (2) from (1)
\[
P(X \mid Y) = \frac{P(X,Y)}{P(Y)} = \frac{P(X)P(Y)}{P(Y)} = P(X)
\]

Similarly, you can derive (3) from (1)
\[
P(Y \mid X) = \frac{P(Y,X)}{P(X)} = \frac{P(Y)P(X)}{P(X)} = P(Y)
\]

3. Suppose you are given a coin that lands heads with probability \( x \) and tails with probability \( 1-x \).

a) Are the outcomes of successive flips of the coin independent of each other given that you know the value of \( x \)? Justify your answer. [2 points]

Yes they are. More precisely, they are conditionally independent of each other.
\[
P(\text{flip1}, \text{flip2} \mid x) = P(\text{flip1} \mid x)P(\text{flip2} \mid x) = x^2
\]

b) Are the outcomes of successive flips of the coin independent of each other if you do not know the value of \( x \)? Justify your answer. [2 points]
No they are not. In fact, the probability of each successive flip depends on all previous flips because it gives you more information about x.

4. After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 97% accurate (i.e. the probability of testing positive when you do have the disease is 0.97, as is the probability of testing negative when you don’t have the disease). The good news is that this is a rare disease, striking only 1 in 5,000 people of your age. Why is it good news that the disease is rare? What are the chances that you actually have the disease? [5 points]

Let us define the following random variables:
D = Disease
T= Test

\[ P(T = \text{true} | D = \text{true}) = 0.97 \]
\[ P(T = \text{false} | D = \text{false}) = 0.97 \Rightarrow P(T = \text{true} | D = \text{false}) = 0.03 \]
\[ P(D = \text{true}) = 0.0002 \]
\[ P(D = \text{false}) = 0.9998 \]

\[
P(D = \text{true} | T = \text{true}) = \frac{P(T = \text{true} | D = \text{true})P(D = \text{true})}{P(T = \text{true} | D = \text{true})P(D = \text{true}) + P(T = \text{true} | D = \text{false})P(D = \text{false})}
\]
\[
= \frac{(0.97)(0.0002)}{(0.97)(0.0002) + (0.03)(0.9998)} = \frac{0.000194}{0.000194 + 0.029994} = 0.000194 / 0.030188 = 0.009804
\]

The probability you actually have the disease is 0.009804. It is good news that the disease is rare because \( P(\text{Disease=\text{true}} | \text{Test=\text{true}}) \) is proportional to \( P(\text{Disease=\text{true}}) \). The lower \( P(\text{Disease=\text{true}}) \) is, the lower \( P(\text{Disease=\text{true}} | \text{Test=\text{true}}) \) will be.