Josh Bell | May 12th 2022 | Intro to Artificial Intelligence

Written Assignment #3

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).**

Toothache	Cavity	Catch	P(Toothache, Cavity, Catch)
F	F	F	0.582
F	F	T	0.128
F	T	F	0.009
F	T	T	0.071
T	F	F	0.066
T	F	T	0.014
T	T	F	0.013
T	T	T	0.117

   a. **P(Catch)**

$$P(Catch = True) = 0.128 + 0.071 + 0.014 + 0.117 = \mathbf{0.33}$$
$$P(Catch = False) = 1 - 0.33 = \mathbf{0.67}$$

   b. **P(Toothache, Catch)**

$$P(Toothache = \boldsymbol{True}, Catch = \boldsymbol{True}) = 0.014 + 0.117 = \mathbf{0.131}$$
$$P(Toothache = \boldsymbol{True}, Catch = \boldsymbol{False}) = 0.066 + 0.013 = \mathbf{0.079}$$
$$P(Toothache = \boldsymbol{False}, Catch = \boldsymbol{True}) = 0.128 + 0.071 = \mathbf{0.199}$$
$$P(Toothache = \boldsymbol{False}, Catch = \boldsymbol{False}) = 0.582 + 0.009 = \mathbf{0.591}$$

   c. **P(Toothache | Catch)**

$$P(Toothache = \boldsymbol{True} \mid Catch = \boldsymbol{True}) = \frac{P(Toothache = \boldsymbol{True}, Catch = \boldsymbol{True})}{P(Catch = \boldsymbol{True})} = \frac{0.131}{0.33} = \mathbf{0.397}$$
$$P(Toothache = \boldsymbol{True} \mid Catch = \boldsymbol{False}) = \frac{P(Toothache = \boldsymbol{True}, Catch = \boldsymbol{False})}{P(Catch = \boldsymbol{False})} = \frac{0.079}{0.67} = \mathbf{0.118}$$
$$P(Toothache = \boldsymbol{False} \mid Catch = \boldsymbol{True}) = \frac{P(Toothache = \boldsymbol{False}, Catch = \boldsymbol{True})}{P(Catch = \boldsymbol{True})} = \frac{0.199}{0.33} = \mathbf{0.603}$$
$$P(Toothache = \boldsymbol{False} \mid Catch = \boldsymbol{False}) = \frac{P(Toothache = \boldsymbol{False}, Catch = \boldsymbol{False})}{P(Catch = \boldsymbol{False})} = \frac{0.591}{0.67} = \mathbf{0.882}$$

2. **Consider a different variant of the Monty Hall problem we did in class. In this variant, there are 4 doors; one door reveals a car, and the other four doors reveal goats. After you select a door, the host will show you 2 goats. The host will not open the door you chose, nor will they reveal the car. Let $C \in \{1,2,3,4\}$ be the location of the car, $D \in \{1,2,3,4\}$ be your initial choice of door, $H1 \in \{1,2,3,4\}$ be the first door opened by the host, and $H\mathbf{2} \in \{1,2,3,4\}$ be the second door opened by the host. Suppose $\boldsymbol{P}(C) = \boldsymbol{P}(D) = \{¼, ¼, ¼, ¼\}$.**
   a. **What is $\boldsymbol{P}(H\mathbf{1} \mid C,D)$?**

C	D	P(H1 = 1| C, D)	P(H1 = 2| C, D)	P(H1 = 3 | C, D)	P(H1 = 4 | C, D)
1	1	0	1/3	1/3	1/3
1	2	0	0	1/2	1/2
1	3	0	1/2	0	1/2
1	4	0	1/2	1/2	0
2	1	0	0	1/2	1/2
2	2	1/3	0	1/3	1/3
2	3	1/2	0	0	1/2
2	4	1/2	0	1/2	0
3	1	0	1/2	0	1/2
3	2	1/2	0	0	1/2
3	3	1/3	1/3	0	1/3
3	4	1/2	1/2	0	0
4	1	0	1/2	1/2	0
4	2	1/2	0	1/2	0
4	3	1/2	1/2	0	0
4	4	1/3	1/3	1/3	0

   b. **What is $\boldsymbol{P}(H\mathbf{2} \mid H\mathbf{1} = 1, C, D = 3)$?**

C	$P(H2=1 \mid H1 = 1, C, D = 3)$	$P(H2=2 \mid H1 = 1, C, D = 3)$	$P(H2=3 \mid H1 = 1, C, D = 3)$	$P(H2=4 \mid H1 = 1, C, D = 3)$
1	0	1/2	0	1/2
2	0	0	0	1/2
3	0	1/3	0	1/3
4	0	1/2	0	0

   c. **Suppose you select door 1. The host opens doors 2 and 3. What are $P(C = 1 \mid D = 1, H\mathbf{1} = 2, H\mathbf{2} = 3)$ and $P(C = 4 \mid D = 4, H\mathbf{1} = 2, H\mathbf{2} = 3)$? (At the point, the car cannot be behind doors 2 or 3 by the rules of the game.) Should you stick with door 1 or switch to door 4?**

$$P(C = 1 \mid D = 1, H1 = 2, H2 = 3) = \frac{1}{4}$$
$$P(C = 4 \mid D = 1, H1 = 2, H2 = 3) = \frac{3}{4}$$

You should **ALWAYS** switch for the better chance because the probability of switching are always higher, in this case ¾ or a 75%, so switching is the way to go.

3. **A breast cancer test has a sensitivity of 92% and a specificity of 97.7%. Sensitivity means the probability of a positive result, given that you have the disease. Specificity means the probability of a negative result, given that you do NOT have the disease. The American breast cancer rate is 13%.**
   a. **Based on these numbers, compute the probability that a patient has breast cancer, given that they get a positive test.**

Question Provides:

$$P(Test = True \mid Disease = True) = \mathbf{92}\%$$

$$P(Test = False \mid Disease = False) = \mathbf{97.7}\%$$

$$P(Disease = True) = \mathbf{13}\%$$

Question Implies:

$$P(Test = False \mid Disease = True) = \mathbf{8}\%$$

$$P(Test = True \mid Disease = False) = \mathbf{2.3}\%$$

$$P(Disease = False) = \mathbf{87}\%$$

Need to Solve:

$$P(Disease = True \mid Test = True) =$$

$$= \frac{P(Test = True \mid Disease = True) * P(Disease = True)}{P(Test = True \mid Disease = True) * P(Disease = True) + P(Test = True \mid Disease = False) * P(Disease = False)}$$

$$= \frac{0.92 * 0.13}{(0.92 * 0.13) + (0.023 * 0.87)}$$

$$= \frac{0.1196}{0.1196 + 0.02001} = \mathbf{0.85667} \cong \mathbf{85.67}\%$$

   b. **What if the breast cancer rate is actually 8%? How does your answer to part (a) change?**

$$= \frac{0.92 * 0.08}{(0.92 * 0.08) + (0.023 * 0.92)}$$

$$= \frac{0.0736}{0.0736 + 0.02001} = \mathbf{0.77669} \cong \mathbf{77.67}\%$$

It lowers the result, it goes from about 85.67% to about 77.67%, which is about 8 percent.