CSE-151A - Discussion 01

Problem 1.

Given the following joint distribution of random variables $X \in \{1, 2, 3\}$ and $Y \in \{a, b\}$:

		Y a	b
	1	1/4	1/8
X	2	1/8	3/8
	3	1/16	1/16

calculate:

a) the marginal distribution of X and Y.

Solution: Recall that marginal distributions are the totals of probabilities.

		Y a	b	P(X)
	1	1/4	1/8	**3/8**
X	2	1/8	3/8	**1/2**
	3	1/16	1/16	**1/8**
	P(Y)	**7/16**	**9/16**	

b) the conditional distribution of X given $Y = a$.

Solution: $P(X|Y = a) = P(X, Y = a)/P(Y = a)$. From 1a, we know that $P(Y = a) = 7/16$. We will then divide the distribution of X when $Y = a$ by 7/16 to obtain conditional distribution of X given $Y = a$.

		Y a	b	P($X, Y = a$)/P($Y = a$)
	1	1/4	1/8	**4/7**
X	2	1/8	3/8	**2/7**
	3	1/16	1/16	**1/7**

Problem 2.

A disease is found in 1/8 of a population. There are two tests for detecting this disease: Test A and Test B. The probabilities that A and B give correct results are 3/4 and 5/6, respectively. A person is randomly chosen and tested. What is the probability that the person actually has the disease in the following cases:

a) Both tests, A and B, are positive.

Solution: Here, we will be using Bayes Theorem:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

and our probability of interest is:

$$P(\text{disease}|\text{both A and B positive}) = \frac{P(\text{both A and B positive}|\text{disease})P(\text{disease})}{P(\text{both A and B positive})}$$

Since we already know that $P(disease) = 1/8$, so we only have to compute the other two probabilities.

$$\begin{aligned} P(\text{both A and B positive}|\text{disease}) &= P(\text{A positive}|\text{disease}) \cdot P(\text{B positive}|\text{disease}) \\ &= \frac{3}{4} \cdot \frac{5}{6} \end{aligned}$$

$$\begin{aligned} P(\text{both A and B positive}) &= P(\text{both A and B positive}|\text{disease})P(\text{disease}) \\ &+ P(\text{both A and B positive}|\text{no disease})P(\text{no disease}) \\ &= P(\text{A positive}|\text{disease}) \cdot P(\text{B positive}|\text{disease}) \cdot P(\text{disease}) \\ &+ P(\text{A positive}|\text{no disease}) \cdot P(\text{B positive}|\text{no disease}) \cdot P(\text{no disease}) \\ &= (\frac{3}{4} \cdot \frac{5}{6})(\frac{1}{8}) + (\frac{1}{4} \cdot \frac{1}{6})(\frac{7}{8}) \end{aligned}$$

Putting all the numbers together, we get:

$$\begin{aligned} P(\text{disease}|\text{both A and B positive}) &= \frac{P(\text{both A and B positive}|\text{disease})P(\text{disease})}{P(\text{both A and B positive})} \\ &= \frac{(\frac{3}{4} \cdot \frac{5}{6})(\frac{1}{8})}{(\frac{3}{4} \cdot \frac{5}{6})(\frac{1}{8}) + (\frac{1}{4} \cdot \frac{1}{6})(\frac{7}{8})} \end{aligned}$$

b) Test A is positive but Test B is negative.

Solution: Similar to part (a), this time our probability of interest is:

$$P(\text{disease}|\text{A positive and B negative}) = \frac{P(\text{A positive and B negative}|\text{disease})P(\text{disease})}{P(\text{A positive and B negative})}$$

We again know that $P(disease) = 1/8$, and have to compute the other two probabilities.

$$\begin{aligned} P(\text{A positive and B negative}|\text{disease}) &= P(\text{A positive}|\text{disease}) \cdot P(\text{B negative}|\text{disease}) \\ &= \frac{3}{4} \cdot \frac{1}{6} \end{aligned}$$

$$
\begin{aligned}
P(\text{A positive and B negative}) &= P(\text{A positive and B negative}|\text{disease})P(\text{disease}) \\
&+ P(\text{A positive and B negative}|\text{no disease})P(\text{no disease}) \\
&= P(\text{A positive}|\text{disease}) \cdot P(\text{B negative}|\text{disease}) \cdot P(\text{disease}) \\
&+ P(\text{A positive}|\text{no disease}) \cdot P(\text{B negative}|\text{no disease}) \cdot P(\text{no disease}) \\
&= (\frac{3}{4} \cdot \frac{1}{6})(\frac{1}{8}) + (\frac{1}{4} \cdot \frac{5}{6})(\frac{7}{8})
\end{aligned}
$$

Finally, we get:

$$
\begin{aligned}
P(\text{disease}|\text{A positive and B negative}) &= \frac{P(\text{A positive and B negative}|\text{disease})P(\text{disease})}{P(\text{A positive and B negative})} \\
&= \frac{(\frac{3}{4} \cdot \frac{1}{6})(\frac{1}{8})}{(\frac{3}{4} \cdot \frac{1}{6})(\frac{1}{8}) + (\frac{1}{4} \cdot \frac{5}{6})(\frac{7}{8})}
\end{aligned}
$$

Problem 3.

You arrive to campus in one of three ways: on the bus, in your car, or on foot.

The buses often run late and have long lines, so when you take the bus, you have a 50% chance of being late for class. When you drive, you sometimes hit traffic or have trouble finding a parking spot, so you have a 30% chance of being late for class. When you walk to campus, you only have a 5% of arriving late. One day, you arrive late for your midterm, and your professor wonders how you got to school that day.

a) If your professor assumes that you are equally likely to use all three modes of transportation, what will the professor calculate for the probability that you took the bus on the day of your midterm?

Solution: We will use Bayes' Theorem, which says that

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

Let B be the event that you took the bus, and A be the event that you were late, so that Bayes theorem becomes

$$P(\text{bus}|\text{late}) = \frac{P(\text{late}|\text{bus})P(\text{bus})}{P(\text{late})}.$$

We are given that

$$
\begin{aligned}
P(\text{late}|\text{bus}) &= 0.5 \\
P(\text{bus}) &= \frac{1}{3}
\end{aligned}
$$

To calculate $P(\text{late})$, notice that since you take exactly one of the three modes of transportation,

$$
\begin{aligned}
P(\text{late}) &= P(\text{late and bus}) + P(\text{late and drive}) + P(\text{late and walk}) \\
&= P(\text{late}|\text{bus})P(\text{bus}) + P(\text{late}|\text{drive})P(\text{drive}) + P(\text{late}|\text{walk})P(\text{walk}) \\
&= 0.5 * \frac{1}{3} + 0.3 * \frac{1}{3} + 0.05 * \frac{1}{3} \\
&= 0.2833
\end{aligned}
$$

Plugging in the values above gives that

$$\begin{aligned}
P(\text{bus}|\text{late}) &= \frac{P(\text{late}|\text{bus})P(\text{bus})}{P(\text{late})} \\
&= \frac{0.5 * \frac{1}{3}}{0.2833} \\
&= 0.5882.
\end{aligned}$$

Thus, the professor would estimate that there is approximately a 59% chance that you took the bus.

b) If your professor happens to know that you take the bus 20% of the time, drive 20% of the time, and walk 60% of the time, what will the professor calculate for the probability that you took the bus on the day of your midterm?

Solution:

We use Bayes Theorem as in part (a), except now

$$\begin{aligned}
P(\text{late}|\text{bus}) &= 0.5 \\
P(\text{bus}) &= 0.2 \\
P(\text{late}) &= P(\text{late and bus}) + P(\text{late and drive}) + P(\text{late and walk}) \\
&= P(\text{late}|\text{bus})P(\text{bus}) + P(\text{late}|\text{drive})P(\text{drive}) + P(\text{late}|\text{walk})P(\text{walk}) \\
&= 0.5 * 0.2 + 0.3 * 0.2 + 0.05 * 0.6 \\
&= 0.19
\end{aligned}$$

Plugging in the values above gives that

$$\begin{aligned}
P(\text{bus}|\text{late}) &= \frac{P(\text{late}|\text{bus})P(\text{bus})}{P(\text{late})} \\
&= \frac{0.5 * 0.2}{0.19} \\
&= 0.5263.
\end{aligned}$$

Thus, the professor would estimate that there is approximately a 53% chance that you took the bus.