Staff Solutions to In-Class Problems Week 12, Fri.

STAFF NOTE: Conditional Probability, Ch. 17

In Spring ’13, most teams finished 20 or more minutes early, mainly because staff was not adequately briefed to get discussions going as directed in individual problems. Also, PS_conditional_probability_problem_errors did not make it into the actual handout used in class.

Problem 1.

There is a rare and serious disease called Beaver Fever which afflicts about 1 person in 1000. Victims of this disease start telling math jokes in social settings, believing other people will think they’re funny.

Doctor Meyer has some fairly reliable tests for this disease. In particular:

- If a person has Beaver Fever, the probability that Meyer diagnoses the person as having the disease is 0.99.
- If a person doesn’t have it, the probability that Meyer diagnoses that person as not having Beaver Fever is 0.97.

Let \( B \) be the event that a randomly chosen person has Beaver Fever, and \( Y \) be the event that Meyer’s diagnosis is “Yes, that person has Beaver Fever,” with \( \overline{B} \) and \( \overline{Y} \) the complements of these events.

(a) The description above explicitly gives the values of the following quantities. What are their values?

\[
\begin{align*}
\Pr(B) & = 0.001 \\
\Pr(Y \mid B) & = 0.99 \\
\Pr(\overline{Y} \mid \overline{B}) & = 0.97
\end{align*}
\]

**Solution.** \( \Pr(B) = 0.001 \quad \Pr(Y \mid B) = 0.99 \quad \Pr(\overline{Y} \mid \overline{B}) = 0.97 \)

(b) Write formulas for \( \Pr(\overline{B}) \) and \( \Pr(Y \mid \overline{B}) \) solely in terms of the explicitly given expressions. Literally use the expressions, not their numeric values.

**Solution.** \( \Pr(\overline{B}) = 1 - \Pr(B) \quad \Pr(Y \mid \overline{B}) = 1 - \Pr(\overline{Y} \mid \overline{B}) \)

(c) Write a formula for the probability that Doctor Meyer says a person has the disease solely in terms of \( \Pr(B) \), \( \Pr(\overline{B}) \), \( \Pr(Y \mid B) \) and \( \Pr(Y \mid \overline{B}) \).

**Solution.** By the Total Probability Law:

\[
\Pr(Y) = \Pr(Y \mid B) \Pr(B) + \Pr(Y \mid \overline{B}) \Pr(\overline{B})
\]

The values turn out to be 0.99(1/1000) + 0.03(1 – 1/1000) = 0.03096.
(d) Write a formula solely in terms of the expressions given in part (a) for the probability that a person has Beaver Fever given that Doctor Meyer says the person has it.

**Solution.**

\[
\Pr[B \mid Y] = \frac{\Pr[B \text{ and } Y]}{\Pr[Y]} = \frac{\Pr[Y \mid B] \Pr[B]}{\Pr[Y \mid B] \Pr[B] + \Pr[Y \mid \overline{B}] \Pr[\overline{B}]} = \frac{\Pr[Y \mid B] \Pr[B]}{\Pr[Y \mid B] \Pr[B] + (1 - \Pr[Y \mid \overline{B}]) (1 - \Pr[B])}.
\]

The values turn out to be

\[
\Pr[B \mid Y] = \frac{0.99(1/1000)}{0.03096} = \frac{99}{3096} \approx \frac{1}{32}.
\]

The low probability of actually having the Fever even though the (97% accurate) test says you do is because the false positive rate (3%) is much larger than the Fever probability (0.1%). This makes false positives more than 30 times more likely than true positives. So if the test says you have Beaver Fever, it's probably a false positive.

Of course Dr. Meyer has recourse to a 99.9% accurate test which has no false positives: simply tell everyone they don’t have the disease.

---

**Problem 2.**

There are three prisoners in a maximum-security prison for fictional villains: the Evil Wizard Voldemort, the Dark Lord Sauron, and Little Bunny Foo-Foo. The parole board has declared that it will release two of the three, chosen uniformly at random, but has not yet released their names. Naturally, Sauron figures that he will be released to his home in Mordor, where the shadows lie, with probability \(\frac{2}{3}\).

A guard offers to tell Sauron the name of one of the other prisoners who will be released (either Voldemort or Foo-Foo). If the guard has a choice of naming either Voldemort or Foo-Foo (because both are to be released), he names one of the two with equal probability.

Sauron knows the guard to be a truthful fellow. However, Sauron declines this offer. He reasons that if the guard says, for example, “Little Bunny Foo-Foo will be released”, then his own probability of release will drop to \(\frac{1}{2}\). This is because he will then know that either he or Voldemort will also be released, and these two events are equally likely.

Dark Lord Sauron has made a typical mistake when reasoning about conditional probability. Using a tree diagram and the four-step method, explain his mistake. What is the probability that Sauron is released given that the guard says Foo-Foo is released?

**Hint:** Define the events \(S\), \(F\), and “\(F\)” as follows:

- “\(F\)” = Guard says Foo-Foo is released
- \(F\) = Foo-Foo is released
- \(S\) = Sauron is released

**Solution.** Sauron’s mistake can be explained as his confusing the two different events \(F\) and “\(F\)”. His observation that \(\Pr[S \mid F] = \frac{1}{2}\) is correct, but that’s the wrong thing to calculate. He should be calculating \(\Pr[S \mid “F”]\).

To clarify the error and work out the proper probability, let’s begin by working out the sample space, noting events of interest, and computing outcome probabilities:
The outcomes in each of these events are noted in the tree diagram.

The tree shows how the event $F$ (Foo-foo will be released) is different from the event “$F$” (the guard says Foo-foo will be released). In particular, the probability that Sauron is released, given that Foo-foo is released, is indeed $\frac{1}{2}$:

$$
\Pr[S \mid F] = \frac{\Pr[S \cap F]}{\Pr[F]} \\
= \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6} + \frac{1}{6}} \\
= \frac{1}{2}
$$

But the probability that Sauron is released given that the guard actually says so is still $\frac{2}{3}$:

$$
\Pr[S \mid "F"] = \frac{\Pr[S \cap "F"]}{\Pr["F"]} \\
= \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}} \\
= \frac{2}{3}
$$

So Sauron’s probability of release is unchanged by the guard’s statement.

Problem 3.

There are two decks of cards. One is complete, but the other is missing the Ace of spades. Suppose you pick one of the two decks with equal probability and then select a card from that deck uniformly at random. What is the probability that you picked the complete deck, given that you selected the eight of hearts? Use the four-step method and a tree diagram.

STAFF NOTE: Try to get a brief discussion going on the issue “How could knowing about the eight of hearts be relevant to the presence of the Ace of spades?”

Solution. Let $C$ be the event that you pick the complete deck, and let $H$ be the event that you select the eight of hearts. In these terms, our aim is to compute:

$$
\Pr[C \mid H] = \frac{\Pr[C \cap H]}{\Pr[H]}
$$
A tree diagram is worked out below:
Now we can compute the desired conditional probability as follows:

$$\Pr[C \mid H] = \frac{\Pr[C \cap H]}{\Pr[H]}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{52}}{\frac{1}{2} \cdot \frac{1}{52} + \frac{1}{2} \cdot \frac{1}{51}}$$

$$= \frac{51}{103}$$

$$= 0.495146 \ldots$$

Thus, if you selected the eight of hearts, then the deck you picked is less likely to be the complete one. It’s worth stopping to think about how you might have arrived at this final conclusion without going through the detailed calculation—or better, how you might explain it to your 10-year-old niece.

The explanation is simple: drawing an eight of hearts from a small deck containing an eight of hearts is more likely than drawing one from a larger such deck. So if you see an eight of hearts, it’s more likely to have come from a smaller deck. The soundness of this intuitive explanation is proved in Problem 17.16. ■

**Problem 4.**
Suppose you repeatedly flip a fair coin until you see the sequence HHT or HHT. What is the probability you see the sequence HHT first?

*Hint:* Try to find the probability that HHT comes before HHT conditioning on whether you first toss an H or a T. The answer is not 1/2.

**Solution.** Let $A$ be the event that HHT appears before HHT, and let $p := \Pr[A]$.

Suppose our first toss is T. Since neither of our patterns starts with T, the probability that $A$ will occur from this point on is still $p$. That is, $\Pr[A \mid T] = p$.

Suppose our first toss is H. To find the probability that $A$ will now occur, that is, to find $r := \Pr[A \mid H]$, we consider different cases based on the subsequent throws.

Suppose the next toss is H, that is, the first two tosses are HH. Then neither pattern appears if we continue flipping H, and when we eventually toss a T, the pattern HHT will then have appeared first. So in this case, event $A$ will never occur. That is $\Pr[A \mid HH] = 0$. 
Suppose the first two tosses are $HT$. If we toss a $T$ again, then we have tossed $HTT$, so event $A$ has occurred. If we next toss an $H$, then we have tossed $HTH$. But this puts us in the same situation we were in after rolling an $H$ on the first toss. That is, $\Pr[A \mid HTH] = r$.

Summarizing this we have:

$$\Pr[A] = \Pr[A \mid T] \Pr[T] + \Pr[A \mid H] \Pr[H]$$

(Law of Total Probability)

$$p = p \frac{1}{2} + r \frac{1}{2}$$

Continuing, we have

$$\Pr[A \mid H] = \Pr[A \mid HT] \Pr[T] + \Pr[A \mid HH] \Pr[H]$$

(Law of Total Probability)

$$r = \Pr[A \mid HT] \frac{1}{2} + 0 \cdot \frac{1}{2}$$

(1)

$$\Pr[A \mid HT] = \Pr[A \mid HTT] \Pr[T] + \Pr[A \mid HTH] \Pr[H]$$

(Law of Total Probability)

$$\Pr[A \mid HT] = 1 \cdot \frac{1}{2} + r \frac{1}{2}$$

(2)

$$r = \left(\frac{1}{2} + r\right) \frac{1}{2}$$

by (1) & (2)

$$r = \frac{1}{3}$$

So $HTT$ comes before $HHT$ with probability

$$p = r = \frac{1}{3}.$$

These kind of events are have an amazing intransitivity property: if you pick any pattern of three tosses such as $HTT$, then I can pick a pattern of three tosses such as $HHT$. If we then bet on which pattern will appear first in a series of tosses, the odds will be in my favor. In particular, even if you instead picked the “better” pattern $HHT$, there is another pattern I can pick that has a more than even chance of appearing before $HHT$. Watch out for this intransitivity phenomenon if somebody proposes that you bet real money on coin flips.

Supplemental problem

Problem 5.

There is a subject—naturally not Math for Computer Science—in which 10% of the assigned problems contain errors. If you ask a Teaching Assistant (TA) whether a problem has an error, then they will answer correctly 80% of the time, regardless of whether or not a problem has an error. If you ask a lecturer, he will identify whether or not there is an error with only 75% accuracy.

We formulate this as an experiment of choosing one problem randomly and asking a particular TA and Lecturer about it. Define the following events:

$$E := \text{the problem has an error},$$

$$T := \text{the TA says the problem has an error},$$

$$L := \text{the lecturer says the problem has an error}.$$  

(a) Translate the description above into a precise set of equations involving conditional probabilities among the events $E$, $T$, and $L$.  

Solution. The assumptions above tell us:

\[ \Pr[E] = \frac{10}{100} = \frac{1}{10}. \]
\[ \Pr[T \mid E] = \Pr[T \mid \overline{E}] = \frac{80}{100} = \frac{4}{5}, \]
\[ \Pr[L \mid E] = \Pr[L \mid \overline{E}] = \frac{75}{100} = \frac{3}{4}. \]

Also, \( T \) and \( L \) are independent given \( E \), and given \( \overline{E} \):

\[ \Pr[T \cap L \mid E] = \Pr[T \mid E] \Pr[L \mid E] \]
\[ \Pr[T \cap L \mid \overline{E}] = \Pr[T \mid \overline{E}] \Pr[L \mid \overline{E}] \]

Note that while we know that \( T \) and \( L \) are independent given \( E \) or given \( \overline{E} \), they are not independent by themselves, see part (c).

(b) Suppose you have doubts about a problem and ask a TA about it, and they tell you that the problem is correct. To double-check, you ask a lecturer, who says that the problem has an error. Assuming that the correctness of the lecturers’ answer and the TA’s answer are independent of each other, regardless of whether there is an error\(^1\), what is the probability that there is an error in the problem?

Solution. We want to calculate \( \Pr[E \mid \overline{T} \cap L] \).

From the definition of conditional probability (this is known as Bayes’ rule):

\[ \Pr[E \mid T \cap L] = \frac{\Pr[E \mid \overline{T} \cap L]}{\Pr[\overline{T} \cap L]} \]

By the independence assumptions, we have:

\[ \Pr[\overline{T} \cap L \mid E] = \Pr[\overline{T} \mid E] \Pr[L \mid E] = \frac{13}{20} \cdot \frac{3}{4} = \frac{3}{20}, \]
\[ \Pr[\overline{T} \cap L \mid \overline{E}] = \Pr[\overline{T} \mid \overline{E}] \Pr[L \mid \overline{E}] = \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5}, \]
\[ \Pr[\overline{T} \cap L] = \Pr[\overline{T} \cap L \mid E] \Pr[E] + \Pr[\overline{T} \cap L \mid \overline{E}] \Pr[\overline{E}] \]
\[ = \frac{3}{20} \cdot \frac{1}{10} + \frac{1}{5} \cdot \frac{9}{10} = \frac{39}{200}. \]

Substituting these values in equation (3), we get

\[ \Pr[E \mid \overline{T} \cap L] = \frac{1}{10} \cdot \frac{3/20}{39/200} = \frac{1}{13} \approx 0.077. \]

So this contradictory information has decreased the probability of an error from 10% to about 7.7%.

The calculations here support the common-sense rule that when two people make contradictory statements, you should be influenced more by the most “authoritative” person—the one who is right more often. But note that this does not mean that you should believe in what the most authoritative person says, since the probability of an error remains uncomfortably large.

---

\(^1\)This assumption is questionable: by and large, we would expect the lecturer and the TA’s to spot the same glaring errors and to be fooled by the same subtle ones.
(c) Is event $T$ independent of event $L$?

**Solution.** The answer is no. Because the TA is usually right, when the TA says that the problem has an error, the likelihood that there really is an error is increased. But the lecturer is also usually right, so increasing the likelihood of there being an error also increases the likelihood that the lecturer will report an error.

We verify this informal argument by actually calculating the probability of each of these events and their conjunction, and observing that the probability that the two events occur is different from the product of the probabilities. Let events $E, T, L$ be as above.

\[
\Pr[T] = \Pr[T \cap E] + \Pr[T \cap \overline{E}] \\
= \Pr[T \mid E] \Pr[E] + \Pr[T \mid \overline{E}] \Pr[\overline{E}] \\
= \frac{4}{5} \cdot \frac{1}{10} + (1 - \frac{4}{5})(1 - \frac{1}{10}) = \frac{13}{50},
\]

\[
\Pr[L] = \Pr[L \cap E] + \Pr[L \cap \overline{E}] \\
= \frac{3}{4} \cdot \frac{1}{10} + (1 - \frac{3}{4})(1 - \frac{1}{10}) = \frac{3}{10},
\]

\[
\Pr[L \cap T] = \Pr[L \cap T \cap E] + \Pr[L \cap T \cap \overline{E}] \\
= \Pr[L \cap T \mid E] \Pr[E] + \Pr[L \cap T \mid \overline{E}] \Pr[\overline{E}] \\
= \Pr[L \mid E] \Pr[T \mid E] \Pr[E] + \Pr[L \mid \overline{E}] \Pr[T \mid \overline{E}] \Pr[\overline{E}] \\
= \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{1}{10} + (1 - \frac{3}{4})(1 - \frac{4}{5})(1 - \frac{1}{10}) = \frac{105}{1000} = 0.105,
\]

which is higher than

\[
\Pr[L] \Pr[T] = \frac{3}{10} \cdot \frac{13}{50} = 0.078.
\]

-