Overview of Section Materials
The warmup questions provided will help students practice concepts introduced in lectures. The section problems are meant to apply these concepts in more complex scenarios similar to what you will see in problem sets and exams.

1 Warmups

1. Cite Bayes’ Theorem and explain why we might want to use it.

2. True or False? Note that true means true for ALL cases. Don’t forget to show proof to support your answer.
   (a) In general, $P(AB|C) = P(B|C)P(A|BC)$
   (b) If $A$ and $B$ are independent, so are $A$ and $B^C$.
   (c) If $A$ and $B$ are independent, so are $A^C$ and $B$.  

1 Bayes’ Theorem: $P(E|F) = \frac{P(F|E)P(E)}{P(F)}$
Bayes’ theorem provides a way to convert a conditional probability from one direction, say $P(E|F)$, to the other direction, $P(F|E)$. It turns out to be the ubiquitous way to answer the question: "how can I update a belief about something, which is not directly observable, given evidence."

2. (a) True
   Using Chain Rule or Definition of Conditional Probability seems hard with three events, but remember that the intersection of two events is also an event. Therefore, expressions with three events are analogous to two events if we are
clever about how we group things.

\[
P(AB|C) = P((AB)|C) \quad \text{Group}
\]
\[
= \frac{P((AB)C)}{P(C)} \quad \text{Def'n Conditional Prob}
\]
\[
= \frac{P(A(BC))}{P(C)} \quad \text{Group}
\]
\[
= \frac{P(A|BC)P(BC)}{P(C)} \quad \text{Chain Rule}
\]
\[
= \frac{P(A|BC)P(B|C)P(C)}{P(C)} \quad \text{Chain Rule}
\]
\[
= P(A|BC)P(B|C) \quad \text{Simplify}
\]
\[
= P(B|C)P(A|BC) \quad \text{Rearrange}
\]

(b) True
Start with some known equation which relates \( A \) and \( B^c \). Is there a way we can exploit the assumed independence between \( A \) and \( B \) and then massage the equation into a form which proves \( A \) and \( B^c \) are independent?

\[
P(A) = P(AB) + P(AB^c) \quad \text{Law of Total Probability}
\]
\[
P(A) = P(A)P(B) + P(AB^c) \quad \text{Independence}
\]
\[
P(A) - P(A)P(B) = P(AB^c) \quad \text{Subtract}
\]
\[
P(A)(1 - P(B)) = P(AB^c) \quad \text{Factor}
\]
\[
P(A)P(B^c) = P(AB^c) \quad \text{Complement}
\]
\[
P(AB^c) = P(A)P(B^c) \quad \text{Rearrange}
\]

(c) True
We know that if \( A \) and \( B \) are independent, then \( B \) and \( A^c \) are independent (as was shown in lecture). Then, we can use what we proved in part (b) to show that \( B \) and \( A^c \) must also be independent.

\[2\]

2 Problems

2.1 Probability Misunderstood: The Sally Clark Case

Preamble: Conditional probabilities are hard to interpret, especially if they are extremely close to zero or one. You should be careful about how you convey meaning to your audience, whether they are on a jury (as below) or are users of software that you have written.
**Problem:** Sally Clark was a British lawyer who was wrongly sentenced to life in prison in 1999 for the deaths of her two infant children. Her elder son Christopher died at age 11 weeks in December 1996 and her younger son Harry at 8 weeks in January 1998. At her trial, the defence argued that the deaths were due to sudden infant death syndrome (SIDS). Clark was convicted on the basis of testimony by pediatrician Sir Roy Meadow, who made the following argument:

- Hospital records show that the ratio of SIDS deaths to live births in affluent non-smoking families is about \( \frac{1}{8500} \). (A live birth is a birth in which a child is born alive; not a still birth.)
- The chance of two SIDS deaths occurring in the same family is about \( \frac{1}{8500} \times \frac{1}{8500} = \frac{1}{73000000} \).
- It is therefore extremely unlikely that Clark is innocent.

As a result of this prosecution, Clark spent more than 3 years in prison and was finally exonerated in 2003 after it was determined that Meadow’s expert testimony was flawed. Two other women against whom Meadow provided expert testimony had their convictions overturned as well. In today’s day, we now believe that underlying genetic and environmental risk factors predispose certain families to SIDS, which can make multiple SIDS deaths in a single family more likely.

a. Identify a flaw in Meadow’s \( \frac{1}{73000000} \) figure. (Hint: think about what assumptions are being made to produce this value)

The probability of two SIDS deaths occurring in the same family is actually likely to be much higher than \( \frac{1}{73000000} \). The flaw in Meadow’s reasoning is that SIDS deaths within the same family are unlikely to occur as independent events. As stated in the question, there are risk factors that predispose certain families to SIDS, which suggests that SIDS deaths cannot be treated as independent events.

Let’s understand these ideas in the notation of probability. Consider an affluent non-smoking family with two children. Let \( C_1 \) and \( C_2 \) be the events that each child dies of SIDS, respectively. According to Meadow’s data,

\[
P(C_1) = P(C_2) \approx \frac{1}{8500}.
\]

But SIDS deaths within a family are not independent, so the probability of both children dying of SIDS is not the product of the probabilities of each of them dying of SIDS:

\[
P(C_1 \cap C_2) \neq P(C_1)P(C_2) \approx \frac{1}{73000000}.
\]

Since there are underlying factors that elevate the risk for certain families, it is likely that

\[
P(C_1 \cap C_2) > \frac{1}{73000000}.
\]

We can restate this comparison using conditional probability:

\[
P(C_2|C_1) = \frac{P(C_1 \cap C_2)}{P(C_1)} > \frac{1/73000000}{1/8500} \approx \frac{1}{8500}.
\]
In words, the probability that the second child dies of SIDS given that the first child has died of SIDS is greater than $\frac{1}{8500}$.

b. Even if we accept Meadow’s $\frac{1}{73000000}$ calculation as valid, what is wrong with a juror interpreting it as the probability of Clark’s innocence?

Let’s be generous to Meadow and accept his claim that the probability of two SIDS deaths in the same family is approximately $\frac{1}{73000000}$. It is still wrong for the jury to interpret this value as the probability that Clark is innocent. Let’s understand the difference by defining some events.

Let $D$ be the event that two children in family die (of any cause). Let $I$ be the event that their mother is innocent of murder. According to Meadow, the probability of the two deaths given that the mother is innocent is miniscule:

$$P(D|I) \approx \frac{1}{73000000}.$$  

But what Clark’s jury should consider instead is $P(I|D)$, the probability that the mother is innocent given that her children have died. These two quantities are related by Bayes rule:

$$P(I|D) = \frac{P(D|I)P(I)}{P(D)}.$$  

Notice that $P(I)$, the prior probability of the mother’s innocence, should be very close to 1; the vast majority of mothers do not murder their children. Likewise, $P(D)$, the prior probability of two children in a family dying, should be very close to 0. Therefore,

$$P(I|D) \gg P(D|I).$$  

In the context of criminal trials, mixing up conditional probabilities in this way is known as the prosecutor’s fallacy. The jury should have disregarded Meadow’s argument. It says almost nothing about Clark’s innocence.

### 2.2 Combining Concepts: Sock Drawer Shenanigans

You wake up at the crack of dawn for your unreasonably early CS109 Final. You stumble around in the dark, throwing on a (hopefully) lucky outfit. You open one of your three sock drawers uniformly at random and hastily pull out two socks, again uniformly at random.

a. You quietly wonder to yourself how likely it is that you put on a matching pair. Before the start of the exam, you decide to calculate this value using information about the content of your drawers:
Let $M$ be the event of a match, and let $A$, $B$, and $C$ be the event that you opened each drawer. We can then use the law of total probability:

$$P(M) = P(M|A)P(A) + P(M|B)P(B) + P(M|C)P(C)$$

$$= \frac{1}{3} \left( \binom{7}{2} + \binom{4}{2} + \binom{8}{2} \right) + \frac{1}{3} \left( \binom{3}{2} + \binom{2}{2} + \binom{9}{2} \right) + \frac{1}{3} \left( \binom{10}{2} + \binom{12}{2} + \binom{5}{2} \right)$$

To get each of the three large fractions above, we treat the socks in each drawer as distinct and divide the number of possible matching pairs in the drawer by the number of pairs (matching or non-matching) that we could have drawn.

b. During the exam, you look down and find that you are wearing two matching socks. What is the probability that those socks came from drawer $B$?

We apply Bayes:

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

and look! We already know the values of those three terms from the first part, which we can simply plug in.

3 Previous Exam Questions

3.1 Spring 2021: Midterm 1

Jerry and Doris have abandoned their respective careers in education and puppy paw modeling and have taken up new work as art curators. As luck would have it, the San Francisco Museum of Modern Art has hired them both to independently appraise modern paintings. Doris has a keen eye for art and can spot a forgery - that is, identifies fake painting as fake - with probability 0.91, whereas Jerry is more easily fooled and only spots a fake painting as fake with probability 0.82. Doris and Jerry also do an excellent good job at certifying authentic art as authentic. When Doris sees an authentic modern art painting, she certifies it as authentic with probability 0.99. Jerry does the same thing, but with probability 0.84.

Those working at SFMoMA are adamant that, based on past experience, the probability that each painting they consider is authentic with probability 0.6. Assuming a single painting is fake, Jerry and Doris independently identify that painting as fake. Similarly, Jerry and Doris independently identify authentic paintings as authentic.

For all of the questions below, in addition to providing an expression, please compute a numeric answer.
a. Assuming a painting is authentic, what is the probability that neither Jerry nor Doris believe it to be authentic?

b. Assuming a painting is a forgery, what is the probability that exactly one of Doris and Jerry identify it as a forgery?

c. Given that both Doris and Jerry certify a painting as authentic, what is the probability the painting is really a forgery?

a. Let's define some events for the problem. Let $A$ be the event that a painting is authentic, let $D$ be the event that Doris says a painting is authentic, and let $J$ be the event that Jerry says a painting is authentic. We have:

\[
P(A) = 0.6 \text{ so that } P(A^C) = 0.4
\]

\[
P(D|A) = 0.99 \text{ and } P(D^C|A) = 0.01
\]

\[
P(D^C|A^C) = 0.91 \text{ and } P(D|A^C) = 0.09
\]

\[
P(J|A) = 0.84 \text{ and } P(J^C|A) = 0.16
\]

\[
P(J^C|A^C) = 0.82 \text{ so } P(J|A^C) = 0.18
\]

The problem is asking for $P(D^C J^C | A)$, which is $P(D^C | A) P(J^C | A)$, or $0.01 \cdot 0.16 = 0.0016$.

b. The approach is similar to that for part a, except that there are more mutually exclusive conditional events that need to be union-ed. In particular, we need to compute $P(J^C | A^C) + P(D^C J | A^C) = P(A^C) P(J^C | A^C) + P(D^C | A^C) P(J | A^C)$. Numerically, this is $0.09 \cdot 0.82 + 0.91 \cdot 0.18 = 0.2376$.

c. Here, we're asking for $P(A^C | DJ)$. The best approach is to employ Bayes' Theorem, in my presentation below, to effectively derive it from scratch.

\[
P(A^C | DJ) = \frac{P(A^C DJ)}{P(DJ)} = \frac{P(DJ | A^C)P(A^C)}{P(DJ)}
\]

\[
= \frac{P(D | A^C) P(J | A^C) P(A^C)}{P(DJ)}
\]

\[
= \frac{P(D | A^C) P(J | A^C) P(A^C)}{P(DJ | A)P(A) + P(DJ | A^C)P(A^C)}
\]

\[
= \frac{P(D | A^C) P(J | A^C) P(A^C)}{0.09 \cdot 0.18 \cdot 0.4 + 0.99 \cdot 0.84 \cdot 0.6 + 0.09 \cdot 0.18 \cdot 0.4}
\]

\[
= \frac{0.99 \cdot 0.84 \cdot 0.6}{0.99 \cdot 0.84 \cdot 0.6 + 0.09 \cdot 0.18 \cdot 0.4}
\]

\[
= 0.01282
\]