Probability Theory for EOR 2020/2021

—

Solutions to Exercises

—

Week 2

November 25, 2021

Conditioning on evidence

1. Define the events. S: message is spam. F: the phrase "free money" appears. We are given that

$$\begin{aligned} P(S) &= 0.8, \\ P(F|S) &= 0.1, \\ P(F|S^C) &= 0.01. \end{aligned}$$

   What is $P(S|F)$? Since we have $P(F|S)$ and are asked for $P(S|F)$, this immediately points you to Bayes' rule:

$$P(S|F) = \frac{P(F|S)P(S)}{P(F|S)P(S) + P(F|S^C)P(S^C)} = \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.01 \cdot 0.2} = 0.9756$$

2. Define the events. I: identical twins, BB: two boys. We are given that

$$\begin{aligned} P(I) &= \frac{1}{3}, \\ P(BB|I) &= \frac{1}{2}, \\ P(BB|F) &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}. \end{aligned}$$

What is $P(I|BB)$? Again, we use Bayes' rule:

$$
\begin{aligned}
P(I|BB) &= \frac{P(BB|I)P(I)}{P(BB|I)P(I) + P(BB|I^C)P(I^C)} \\
&= \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{2}\frac{1}{3} + \frac{1}{4}\frac{2}{3}} \\
&= \frac{1}{2}.
\end{aligned}
$$

4a. We are given in the question that

$$
\begin{aligned}
P(R|K) &= 1, \\
P(R|K^C) &= \frac{1}{n}, \\
P(K) &= p.
\end{aligned}
$$

What is $P(K|R)$? What do you think, maybe Bayes' rule?

$$
\begin{aligned}
P(K|R) &= \frac{P(R|K)P(K)}{P(R|K)P(K) + P(R|K^C)P(K^C)} \\
&= \frac{p}{p + \frac{1-p}{n}} \\
&= \frac{n}{(n-1)p + 1}p.
\end{aligned}
$$

4b. From (a)

$$
P(K|R) = \frac{n}{(n-1)p + 1}p.
$$

Since $\frac{n}{(n-1)p+1} \geq \frac{n}{(n-1)+1} = 1$, we have $P(K|R) \geq p$. Given the evidence that Fred has the right answer you update your belief that he actually knows the answer upwards. Of course, if $p = 0$, then $P(K) = P(K|R) = 0$. If we Fred can never know the answer, then any evidence that he got an answer right must be interpreted as due to random chance. (Thank Quinten Huisman for pointing this case out) Or if $p = 1$, then Fred always knows the answer, so event R does not update the probability of the event K either.

6. Define as DH the event that the chosen coin is double headed, by H_7 the event that

the chosen coin lands heads 7 times. Then

$$
\begin{aligned}
P(DH) &= \frac{1}{100}, \\
P(H_7|DH) &= 1, \\
P(H_7|DH^C) &= \frac{1}{2}^7.
\end{aligned}
$$

What is $P(DH|H_7)$? Bayes' rule anyone?

$$
\begin{aligned}
P(DH|H_7) &= \frac{P(H_7|DH)P(DH)}{P(H_7|DH)P(DH) + P(H_7|DH^C)P(DH^C)} \\
&= \frac{\frac{1}{100}}{\frac{1}{100} + \frac{1}{2^7}\frac{99}{100}} \\
&= 0.5639
\end{aligned}
$$

7. Denote by D the event that there is a double headed coin.

$$
\begin{aligned}
P(D|H_7) &= \frac{P(H_7|D)P(D)}{P(H_7|D)P(D) + P(H_7|D^C)P(D^C)} \\
&= \frac{P(H_7|D)\frac{1}{2}}{P(H_7|D)\frac{1}{2} + \frac{1}{2}^8}.
\end{aligned}
$$

What remains is $P(H_7|D)$. Denote by C the event that you select the double headed coin, then, using the law of total probability

$$
\begin{aligned}
P(H_7|D) &= P(H_7|D,C)P(C|D) + P(H_7|D,C^C)P(C^C|D) \\
&= \frac{1}{100} + \frac{1}{2}^7 \frac{99}{100}.
\end{aligned}
$$

Plugging in the numbers into our expression for $P(D|H_7)$, we find that

$$
P(D|H_7) = \frac{227}{327} = 0.6942.
$$

9a. The thing to realize is that A_1 implies B in set notation is $A_1 \subseteq B$. Therefore, $P(A_1 \cap B) = P(A_1)$. The same is true for A_2. Then,

$$
\begin{aligned}
P(A_1|B) &= \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1)}{P(B)}, \\
P(A_2|B) &= \frac{P(A_2 \cap B)}{P(B)} = \frac{P(A_2)}{P(B)}.
\end{aligned}
$$

If $P(A_1) = P(A_2)$, then the right hand sides are equal, so the left hand sides must also be equal.

9b. Suppose the probability of studying Econometrics (A_1) is the same as the probability of studying Mathematics (A_2). Denote by B the probability of going to the university. Of course, A_1 and A_2 imply B. The problem is completely symmetric in A_1 and A_2. If I now observe that someone is going to the university, I have to update the probabilities for A_1 and A_2 in the same way.

10. Use LOTP with extra conditioning on A_2, and then the conditional independence of A_3 and A_1 given A_2 or A_2^C that is given in the question.

$$\begin{aligned}
P(A_3|A_1) &= P(A_3|A_1, A_2)P(A_2|A_1) + P(A_3|A_1, A_2^C)P(A_2^C|A_1) \\
&= P(A_3|A_2)P(A_2|A_1) + P(A_3|A_2^C)P(A_2^C|A_1) \\
&= 0.8^2 + 0.3 \cdot 0.2 \\
&= 0.7.
\end{aligned}$$

And also,

$$\begin{aligned}
P(A_3|A_1^C) &= P(A_3|A_1^C, A_2)P(A_2|A_1^C) + P(A_3|A_1^C, A_2^C)P(A_2^C|A_1^C) \\
&= P(A_3|A_2)P(A_2|A_1^C) + P(A_3|A_2^C)P(A_2^C|A_1^C) \\
&= 0.8 \cdot 0.3 + 0.3 \cdot 0.7 \\
&= 0.45.
\end{aligned}$$

10b. Using the law of total probability (with A_1 as conditioning event) and the results from (a), we find that

$$\begin{aligned}
P(A_3) &= P(A_3|A_1)P(A_1) + P(A_3|A_1^C)P(A_1^C) \\
&= 0.7 \cdot 0.75 + 0.45 \cdot 0.25 \\
&= 0.6375.
\end{aligned}$$

14a. If Peter's home is burglarized, he will probably install a burglar alarm system after. So $P(A|B) > P(A|B^c)$.

14b. If Peter installs an alarm system, he is less likely to be burglarized, so $P(B|A) < P(B|A^c)$.

14c. By the LOTP

$$\begin{aligned}P(A) &= P(A|B)P(B) + P(A|B^C)P(B^C)\\ &< P(A|B)P(B) + P(A|B)P(B^C)\\ &= P(A|B)\\ &= \frac{P(B|A)P(A)}{P(B)}\end{aligned}$$

Rearranging gives $P(B|A) > P(B)$. Using this and again using the LOTP

$$\begin{aligned}P(B) &= P(B|A)P(A) + P(B|A^C)P(A^C)\\ &> P(B)P(A) + P(B|A^C)P(A^C)\end{aligned}$$

Rearranging this, we find $P(B) > P(B|A^C)$. Since we already got $P(B|A) > P(B)$, we now have $P(B|A) > P(B|A^C)$.

Perhaps it's more intuitive to make a Venn diagram to see the symmetry that leads to the correct answer. $P(A|B) > P(A|B^c)$ is the same as saying that $P(A \cap B) \cdot P(A^C \cap B^C) > P(A \cap B^C) \cdot P(A^C \cap B)$. Since this inequality is symmetric in A and B (if I change A to B it stays the same), this must also mean that $P(B|A) > P(B|A^C)$.

14d. In a, the argument is that Peter waits until he's burglarized before making a decision on the alarm system. If this were true, then in B, observing that he has an alarm system should increase the probability of the event that he was at some point burglarized. So the answers in a. and b. are inconsistent. You mind however flips the time ordering in the argument in a compared to b, and this leads to an inconsistency between the answers.

15. You want to maximize $P(A \cap B|I)$, so the probability that both A and B occur given some information I. Now I can be A, B, or $A \cup B$. By the definition of conditional probability

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

Think about the second equality sign. In order to maximize this probability, you want to minimize $P(I)$. Given the ordering of the probabilities in the question, you would like to know A occurred.

17a. We start with the definition of conditional probability.

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

Since $P(B|A) = 1$, we have that $P(A) = P(B \cap A)$. We can split the event A into the disjoint sets $B \cap A$ and $B^C \cap A$. Then, by the second Axiom of Probability

$$P(A) = P(B \cap A) + P(B^C \cap A).$$

Since $P(A) = P(B \cap A)$, we infer that $P(B^C \cap A) = 0$. Using again the definition of conditional probability

$$P(A|B^C) = \frac{P(B^c \cap A)}{P(B^C)}.$$

Since we found that $P(B^C \cap A) = 0$, we see that $P(A|B^C) = 0$. Using that $P(A|B^C) + P(A^C|B^C) = 1$ (why?), we arrive at our conclusion that $P(A^C|B^C) = 1$.

Alternatively (thanks to Wietze Koops for suggesting this solution.)
Assume $P(B|A) = 1$. Then, $P(B^C|A) = 1 - P(B|A) = 0$. (The first equality [called the complement rule] follows from the conditional versions of the Axioms of Probability.)
Hence, $P(A|B^C) = \frac{P(B^C|A)P(A)}{P(B^C)} = 0$. (Bayes' rule)
Hence, $P(A^C|B^C) = 1 - P(A|B^C) = 1$. (Again the complement rule)

17b. Suppose you throw two n sided die (with i eyes on the ith side for $i = 1, \dots, n$). Let A be the event that the number of eyes on the first die is strictly smaller than n. Let B be the event that the number of eyes of the second die is strictly smaller than n. Now, $P(B|A) = P(B) = \frac{n-1}{n}$, and $P(A^C|B^C) = P(A^C) = \frac{1}{n}$. The reason that this does not contradict (a) is in the requirement that the events A and B cannot have probability equal to 0 or 1. Suppose we change the events A and B to the number of eyes smaller or equal than n. Clearly $P(B|A) = P(B) = 1$. However, $P(A^C|B^C) = \dots$ not defined, because the definition of conditional probability requires $P(B^C) > 0$.

21a. Denote by A_i the event that we toss i heads. we are looking for

$$\begin{aligned}
P(A_3|A_2 \cup A_3) &= \frac{P(A_3 \cap (A_2 \cup A_3))}{P(A_2 \cup A_3)} && \text{(Definition of Conditional Probability)} \\
&= \frac{P((A_3 \cap A_2) \cup (A_3 \cap A_3))}{P(A_2 \cup A_3)} && \text{(Distributive law)} \\
&= \frac{P(A_3)}{P(A_2 \cup A_3)} \\
&= \frac{P(A_3)}{P(A_2) + P(A_3) - P(A_2 \cap A_3)} && \text{(Inclusion-exclusion)} \\
&= \frac{P(A_3)}{P(A_2) + P(A_3)} \\
&= \frac{(1/2)^3}{3 \cdot (1/2)^3 + (1/2)^3} \\
&= \frac{1}{4}.
\end{aligned}$$

21b. Denote by C the event that the slips show heads. We are looking for

$$\begin{aligned}
P(A_3|C) &= \frac{P(C|A_3)P(A_3)}{P(C)} \\
&= \frac{P(C|A_3)P(A_3)}{P(C|A_3)P(A_3) + P(C|A_3^C)P(A_3^C)} \\
&= \frac{1 \cdot \frac{1}{2^3}}{1 \cdot \frac{1}{2^3} + P(C|A_3^C)(1 - \frac{1}{2^3})}
\end{aligned}$$

We now only need $P(C|A_3^C)$. Now using again the LOTP, we have

$$\begin{aligned}
P(C|A_3^C) &= P(C|A_3^C, A_2)P(A_2|A_3^C) + P(C|A_3^C, A_2^C)P(A_2^C|A_3^C) \\
&= P(C|A_2)P(A_2|A_3^C) + 0 \cdot P(A_2^C|A_3^C) \\
&= \frac{\binom{2}{2}}{\binom{3}{2}} \cdot \frac{3}{7} \\
&= \frac{1}{7},
\end{aligned}$$

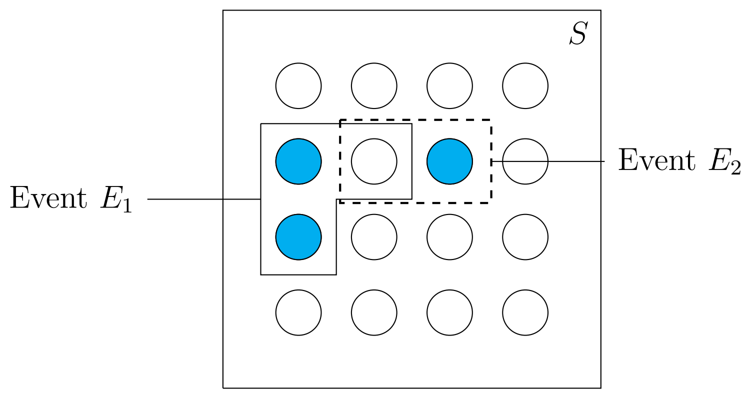
where on the second to last line, both probabilities are calculated using the naive definition of probability. For the first, conditional on A_2, we know there are two slips indicating H, from which we need to choose 2. There are $\binom{2}{2}$ ways to do so. The total number of ways to draw the 2 slips from 3 is $\binom{3}{2} = 3$. To find the probability of $A_2|A_3^C$, simply list all 7 possible options (which are equally likely) and count the number of outcomes which contain 2 times an H.

Now we are ready to get the final answer

$$\begin{aligned} P(A_3|C) &= \frac{1 \cdot \frac{1}{2^3}}{1 \cdot \frac{1}{2^3} + \frac{1}{7}(1 - \frac{1}{2^3})} \\ &= \frac{1}{1+1} = \frac{1}{2}. \end{aligned}$$

Intuitively, in (a) we have information that there are at least two slips with heads, but we don't know which. There are 8 possible outcomes, 4 of which satisfy the criterion that at least two are heads. These are all equally likely, and only one outcome is three heads, so the answer to (a) has to be 1/4. In (b), consider randomly arranging the slips and then picking the first two. It is given that these two are heads. This leaves only two possible, equally likely, outcomes for the final slip. One of these two outcomes results in three heads, so the answer must be 1/2.

23. It's easiest to make a drawing. Let blue denote the event G.



We see that $P(G) = 3/16$, $P(G|E_1) = 2/3$, $P(G|E_2) = 1/2$, yet $P(G|E_1, E_2) = 0$.

Independence and Conditional Independence

32a. You can calculate these probabilities by conditioning on the outcome of one of the two dice involved.

$$\begin{aligned}
P(A > B) &= P(A > B|A = 4)P(A = 4) + P(A > B|A = 0)P(A = 0) \\
&= 1 \cdot \frac{2}{3} + 0 \\
&= \frac{2}{3}, \\
P(B > C) &= \frac{2}{3}, \\
P(C > D) &= P(C > D|C = 6)P(C = 6) + P(C > D|C = 2)P(C = 2) \\
&= 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} \\
&= \frac{2}{3}, \\
P(D > A) &= P(D > A|D = 5)P(D = 5) + P(D > A|D = 1)P(D = 1) \\
&= 1 \cdot \frac{1}{2} + \frac{1}{3}\frac{1}{2} \\
&= \frac{2}{3}.
\end{aligned}$$

32b. If $A > B$, then you know $A = 4$. This doesn't tell you anything about the event $B > C$, so $A > B$ is independent of $B > C$. If $B > C$, then you know that $C = 2$, so $P(C > D|B > C) = P(D < 2) = \frac{1}{2}$. We see that $P(C > D|B > C) \neq P(C > D)$, so $C > D$ is not independent of $B > C$.

34a. Suppose you have an accident in year 1. Then, you are more likely to be a bad driver, and hence, also more likely to have an accident in year 2 as well. This means that $P(B|A) \neq P(B)$ and hence, A and B are not independent.

34b. We find this probability using Bayes' rule,

$$\begin{aligned}
P(G|A^C) &= \frac{P(A^C|G)P(G)}{P(A^C)} \\
&= \frac{P(A^C|G)P(G)}{P(A^C|G)P(G) + P(A^C|G^C)P(G^C)} \\
&= \frac{(1 - p_1)g}{(1 - p_1)g + (1 - p_2)(1 - g)}.
\end{aligned}$$

34c. We are given that given G, the events A and B are independent. Then, using the

LOTP,

$$\begin{aligned}
P(B|A^C) &= P(B|A^C, G)P(G|A^C) + P(B|A^C, G^C)P(G^C|A^C) \\
&= P(B|G)P(G|A^C) + P(B|G^C)P(G^C|A^C) \\
&= p_1 \frac{(1-p_1)g}{(1-p_1)g + (1-p_2)(1-g)} + p_2 \left(1 - \frac{(1-p_1)g}{(1-p_1)g + (1-p_2)(1-g)}\right) \\
&= \frac{p_1(1-p_1)g + p_2(1-p_2)(1-g)}{(1-p_1)g + (1-p_2)(1-g)}.
\end{aligned}$$

35a. The probability of winning the first game is found by applying the LOTP as follows,

$$\begin{aligned}
P(W_1) &= P(W_1|B)P(B) + P(W_1|I)P(I) + P(W_1|M)P(M) \\
&= \frac{9}{10} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + \frac{3}{10} \cdot \frac{1}{3} \\
&= \frac{17}{30}.
\end{aligned}$$

35b. Denote by B, I, M the events of having a beginner, intermediate or master opponent, respectively.

$$\begin{aligned}
P(W_2|W_1) &= P(W_2|W_1, B)P(B|W_1) + P(W_2|W_1, I)P(I|W_1) + P(W_1|W_1, M)P(M|W_1) \\
&= P(W_2|B)P(B|W_1) + P(W_2|I)P(I|W_1) + P(W_1|M)P(M|W_1) \\
&\qquad \text{(Using that given skill, outcomes are independent)}
\end{aligned}$$

Now, using Bayes' rule

$$\begin{aligned}
P(B|W_1) &= \frac{P(W_1|B)P(B)}{P(W_1)} = \frac{9}{17}, \\
P(I|W_1) &= \frac{P(W_1|I)P(I)}{P(W_1)} = \frac{5}{17}, \\
P(M|W_1) &= \frac{P(W_1|M)P(M)}{P(W_1)} = \frac{3}{17}.
\end{aligned}$$

We can conclude that

$$\begin{aligned}
P(W_2|W_1) &= \frac{9}{10}\frac{9}{17} + \frac{1}{2}\frac{5}{17} + \frac{3}{10}\frac{3}{17} \\
&= \frac{23}{34}.
\end{aligned}$$

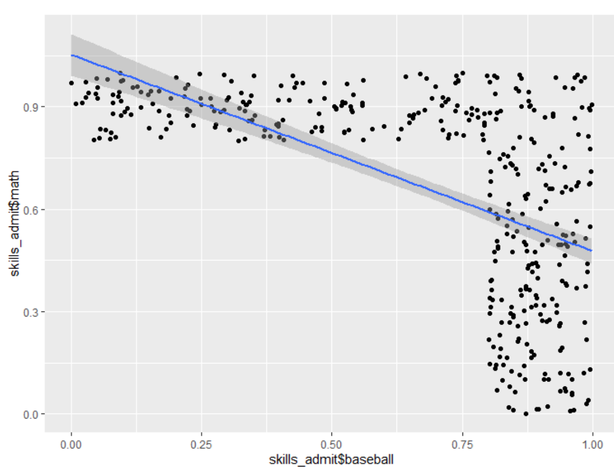
Now, look back at your answer to (a). There you found that winning the first game had probability 0.567. Given that you won the first round, you would expect the probability of winning the second round to go up. Indeed, the answer to (b) is 0.676.

35c. Knowing the outcome of one of the matches, we adjust our beliefs on the quality of our opponent. This changes our beliefs on the outcome of the other match. However, if the quality of the opponent is fixed, then outcome of the first match does not affect the outcome of the second match.

36 This question will of course be superimportant for later in the studies when issues like selection bias are being discussed. Perhaps you have a nice alternative example as well. Also, Tom has written some R code to illustrate what's going on.

```
# Author: Tom Boot
# Date: 22 October 2020
# Introduction to Probability Question 2.36

library('ggplot2')

# Take a population of size ng
n <- 1000

# With math skills and baseball skills some number between 0 and 1
# [runif() gives continuous uniform random variables,
# which will be discussed later in the course]
# and these skills are independent
skills <- data.frame("baseball"=runif(n),"math"=runif(n))

# Scatter plot of baseball and math skills for the whole population
ggplot(skills, aes(x=baseball, y=math)) + geom_point() +
geom_smooth(method="lm",formula=y~x)

# Only admit students that are good at at least one of baseball and math
admit <- ((skills$baseball>0.8) + (skills$math>0.8)>0)
skills_admit <- skills[admit,]

# Scatter plot of baseball and math skills for admitted students
ggplot(skills_admit,aes(x=skills_admit$baseball,y=skills_admit$math)) +
geom_point()+ geom_smooth(method="lm",formula=y~x)
```

36a. Within the group admitted to the college, if you are bad at baseball, you must be good at math, otherwise you would not be admitted. This creates a negative association between baseball and math skills. See also the R illustration provided:



36b.

$$P(A|C) = P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} > P(A).$$

Using this result, we see that

$$P(A|B, C) = P(A|B \cap (A \cup B)) = P(A|B) = P(A) < P(A|C)$$

We conclude that $P(A|B, C) \neq P(A|C)$, hence A and B are conditionally dependent given C.

Monty Hall

41. It's annoying that we do not know whether Monty flipped Heads H or Tails, so we are going to condition on this event.

    Let C_i be the event that the car is behind door i and G_i the event there is a goat behind door i.. Denote by D_j the event that Monty opens door j. We now apply the LOTP

$$\begin{aligned} P(C_3|D_2, G_2) &= P(C_3|H, D_2, G_2)P(H|D_2, G_2) + P(C_3|H^C, D_2, G_2)P(H^C|D_2, G_2) \\ &= \frac{2}{3} \cdot P(H|D_2, G_2) + P(C_3|H^C, D_2, G_2)P(H^C|D_2, G_2) \end{aligned}$$

    The second line uses the result from the standard Monty Hall problem that the probability of winning the car given the switching strategy is 2/3. Now, we use the definition of conditional probability

$$P(C_3|H^C, D_2, G_2) = \frac{P(D_2, G_2|C_3, H^C)P(C_3|H^C)}{P(D_2, G_2|H^C)} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{2}{3}} = \frac{1}{2}.$$

    Here we used that $P(D_2, G_2|C_3, H^C) = \frac{1}{2}$, since (1) if there is a car behind door 3, then there is a goat behind door 2, so $P(D_2, G_2|C_3, H^C) = P(D_2|C_3, H^C)$ and (2) Monty is randomly opening one of the two doors, so $P(D_2|C_3, H^C) = P(D_2|H^C) = \frac{1}{2}$. Also $P(C_3|H^C) = P(C_3) = \frac{1}{3}$, since the location of the car is independent of whether Monty opens doors at random or only opens doors with goats. Finally $P(D_2, G_2|H^C) = P(D_2|H^C)P(G_2|H^C) = \frac{1}{2} \cdot \frac{1}{3}$, since Monty randomly opens a door, this is independent of whether there is a goat there.

The only unknown quantity that remains is

$$
\begin{aligned}
P(H|D_2, G_2) &= \frac{P(D_2, G_2|H)P(H)}{P(D_2, G_2|H)P(H) + P(D_2, G_2|H^C)P(H^C)} \\
&= \frac{(P(D_2, G_2|C_1, H)P(C_1|H) + P(D_2, G_2|C_3, H)P(C_3|H))P(H)}{P(D_2, G_2|H)P(H) + P(D_2, G_2|H^C)P(H^C)} \qquad \text{(LOTP)} \\
&= \frac{(\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3})p}{(\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3})p + (\frac{1}{2} \cdot \frac{2}{3})(1-p)}
\end{aligned}
$$

where we used that $P(D_2, G_2|C_1, H) = \frac{1}{2}$ (if there is a car behind door 1, Monty decides between the two remaining doors with probability $\frac{1}{2}$). Also, $P(C_1|H) = P(C_3|H) = P(C_1) = \frac{1}{3}$, and $P(D_2, G_2|C_3, H) = P(D_2|C_3, H) = 1$ (note that the term involving C_2 not appearing when applying the LOTP, why?).

Substituting all of this into the first equation of the answer gives

$$P(C_3|D_2, G_2) = \frac{1+p}{2+p}.$$

First-step analysis and gambler's ruin

48a. Suppose, we current have $n - i$ for $i = 1, \ldots, 6$, then we have a probability of $\frac{1}{6}$ of achieving p_n. So

$$p_n = \frac{1}{6}\sum_{i=1}^{6} p_{n-i}$$

Assuming we start at 0, it is reasonable to set $p_0 = 1$ and $p_k = 0$ for $k < 0$.

48b.

$$
\begin{aligned}
p_1 &= \frac{1}{6}, \\
p_2 &= \frac{1}{6}(p_0 + p_1) = \frac{1}{6}\left(1 + \frac{1}{6}\right), \\
p_3 &= \frac{1}{6}(p_0 + p_1 + p_2) = p_2(1 + \frac{1}{6}) = \frac{1}{6}\left(1 + \frac{1}{6}\right)^2, \\
p_4 &= \frac{1}{6}(p_0 + p_1 + p_2 + p_3) = p_3\left(1 + \frac{1}{6}\right) = \frac{1}{6}\left(1 + \frac{1}{6}\right)^3, \\
p_5 &= \frac{1}{6}(p_0 + p_1 + p_2 + p_3 + p_4) = p_4\left(1 + \frac{1}{6}\right) = \frac{1}{6}\left(1 + \frac{1}{6}\right)^4, \\
p_6 &= \frac{1}{6}(p_0 + p_1 + p_2 + p_3 + p_4 + p_5) = p_5\left(1 + \frac{1}{6}\right) = \frac{1}{6}\left(1 + \frac{1}{6}\right)^5, \\
p_7 &= \frac{1}{6}(p_1 + p_2 + p_3 + p_4 + p_5 + p_6) = p_6\left(1 + \frac{1}{6}\right) - \frac{1}{6}p_0 = \frac{1}{6}\left(1 + \frac{1}{6}\right)^6 - \frac{1}{6}.
\end{aligned}
$$

In decimals, $p_7 = 0.2536$.

48c. Skip.

50a. First use the LOTP where you condition on the outcome of the first match,

$$P(W) = P(W|W_1)p + P(W|W_1^C)(1-p).$$

Now use extra conditioning on the outcome of the second match

$$
\begin{aligned}
P(W|W_1) &= P(W|W_1, W_2)P(W_2|W_1) + P(W|W_1, W_2^C)P(W_2^C|W_1) \\
&= 1 \cdot p + P(W)(1-p).
\end{aligned}
$$

Similarly, we find that

$$
\begin{aligned}
P(W|W_1^C) &= P(W|W_1^C, W_2)P(W_2|W_1^C) + P(W|W_1^C, W_2^C)P(W_2^C|W_1^C) \\
&= P(W)p.
\end{aligned}
$$

Substituting the last two results into the first equation, we have

$$P(W) = p^2 + P(W)(1-p)p + P(W)p(1-p) = p^2 + 2p(1-p)P(W).$$

Solving for $P(W)$ gives

$$P(W) = \frac{p^2}{1 - 2p(1-p)}.$$

50b. Skip.

Simpson's paradox

56a. (From the sample solutions) Let H be the event that the man will hurt Stampy, let L be the event that a man has lots of ivory, and let D be the event that the man is an ivory dealer.

56b. Lisa observes that L is true. She suggests (reasonably) that this evidence makes D more likely, i.e., $P(D|L) > P(D)$. Implicitly, she suggests that this makes it likely that the man will hurt Stampy, i.e., $P(H|L) > P(H|L^C)$. Homer argues that $P(H|L) < P(H|L^C)$.

56c. Homer does not realize that observing that Blackheart has so much ivory makes it much more likely that Blackheart is an ivory dealer, which in turn makes it more likely that the man will hurt Stampy. This is an example of Simpson's paradox. It may be true that, controlling for whether or not Blackheart is a dealer, having high ivory supplies makes it less likely that he will harm Stampy: $P(H|L, D) < P(H|L^C, D)$ and $P(H|L, D^C) < P(H|L^C, D^C)$. However, this does not imply that $P(H|L) < P(H|L^C)$.

57. Suppose C_1 is small, but has a high percentage of green gummy bears. M_1 is large, and has a lower percentage of green gummy bears. C_2 is large and has a higher percentage green gummy bears than M_2, but lower than M_1. M_2 is small. Because C_1 and M_2 are small, the percentage of green gummy bears after merging will be close to the percentage in C_2 for the $C_1 + C_2$ mixture, and close to the percentage in M_2 for the $M_1 + M_2$ mixture. Since the percentage in M_2 is higher than in C_2, the $M_1 + M_2$ mixture will also have a higher percentage of green gummi bears. In numbers, $C_1 : 9G, 1R$. $M_1 : 500G, 500R$, $C_2 : 300G, 700R$, $M_2 : 1G, 9R$. The percentage of green gummy bears in C is 309/1010, while in M it is 501/1010.