1. (13 pts.) Colorful coins

(a) (3 pts.) Describe the basic random variables and the outcomes in the sample space, and give their probabilities.

Answer: We have one random variable $C$ which denotes the coin chosen (1, 2 and 3, with 1 being the fair coin), two random variables $F_1$ and $F_2$ denoting the face that comes up for the first and second flip, and two random variables $X_1$ and $X_2$ denoting the color of the first and second flip. There are two stages to the experiment: the selection of a coin to flip coins and the two flips of the coin. Then the possible outcomes are shown in the tree in Figure 1. The probability of each outcome is found by multiplying the probabilities on each branch.

![Tree diagram of possible outcomes](image)

Figure 1: Possible outcomes of the colorful coin tossing experiment

(b) (3 pts.) Now suppose two coins are chosen randomly with replacement and each flipped once. Describe the outcomes in the sample space in this new experiment, and give their probabilities. Are they the same as in part (a)?

Answer: There are nine possible outcomes of the choice of coins in this case, each occurring with probability $\frac{1}{9}$. Corresponding to each of the nine choices, there are four possible outcomes of the coin
tosses \((HH, HT, TH, TT)\). The probability of each outcome can again be calculated by multiplying heads or tails probabilities of the appropriate coins in each case.

If we don’t care whether coin 2 or 3 was chosen, since they are identical, then the nine possible outcomes of choosing two coins reduces to the four outcomes of choosing either the fair coin (1), or a biased coin (2 or 3) for each flip. Writing out these four cases we have:

- \(\frac{1}{9}\) probability of choosing two fair coins, and probabilities to get \((HH, HT, TH, TT)\) of \((\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})\), respectively.

- \(\frac{2}{9}\) probability of choosing a fair coin followed by a biased coin, and probabilities to get \((HH, HT, TH, TT)\) of \((\frac{1}{2}p, \frac{1}{2}(1-p), \frac{1}{2}p, \frac{1}{2}(1-p))\), respectively.

- \(\frac{2}{9}\) probability of choosing a biased coin followed by a fair coin, and probabilities to get \((HH, HT, TH, TT)\) of \((\frac{1}{2}p, \frac{1}{2}p, \frac{1}{2}(1-p), \frac{1}{2}(1-p))\), respectively.

- \(\frac{4}{9}\) probability of choosing two biased coins, and probabilities to get \((HH, HT, TH, TT)\) of \((p^2, p(1-p), p(1-p), (1-p)^2)\), respectively.

Note that the last three cases for the coins chosen expand into eight different cases when we care whether coin 2 or coin 3 is chosen.

The outcomes are not the same as in part (a) because now it is possible to have one toss being from the blue-white coin and one toss from a red-blue coin, which was not a possible outcome in part (a). For example, when coins 1 and 2 are chosen, we can get an outcome which is \((white, red)\), which was not possible previously.

(c) (3 pts.) Now suppose two coins are chosen randomly without replacement and each flipped once. Describe the outcomes in the sample space in this new experiment, and give their probabilities. Are they the same as in parts (a) or (b)?

**Answer:** There are \(\binom{3}{2} = 3\) different ways that two coins can be selected, and each of these choices is equally likely. For each chosen pair of coins, the flips have four possible outcomes: \((heads, heads)\), \((heads, tails)\), \((tails, heads)\), \((tails, tails)\). We enumerate all the elements in the space with their associated probabilities in Figure 2.

![Figure 2: Possible outcomes of the colorful coin tossing experiment](image-url)
The probabilities are not the same as in either (a) or (b). Again, we can have both the fair and an unbiased coin chosen, so unlike in (a), we can get a white and a blue face for the two coins flipped. Also, there is 0 probability in this case of choosing two fair coins, so (white, white) is not a possible outcome, unlike in (b).

(d) **(4 pts.)** Suppose the probability that the two sides that land face up are the same color is $\frac{29}{96}$ in the experiment in part (c). What does this tell you about the possible values of $p$?

**Answer:** We see that four of the possible outcomes lie in the event we are interested in (both sides having the same color); these are marked with an asterisk in Figure 2.

Adding up the probabilities of these four outcomes, we get that the probability that the sides that land face up are the same color is:

$$P(\text{same}) = \frac{1}{3} \left\{ \frac{1}{2} (1-p) + \frac{1}{2} (1-p) + p^2 + (1-p)^2 \right\}$$

$$= \frac{1}{3} (2p^2 - 3p + 2).$$

Setting this equal to $\frac{29}{96}$ and simplifying, we obtain the quadratic equation $64p^2 - 96p + 35 = 0$. Solving this equation for $p$ yields the two possible solutions $p = \frac{5}{8}$ or $\frac{7}{8}$.

2. **(12 pts.) Beat Deep Blue**

You’ve been booked to play a chess tournament where to win the tournament, you have to win two consecutive games of chess, out of three games. You have the choice of playing David Tse, then Deep Blue, then David Tse (TBT) – or Deep Blue, then David Tse, then Deep Blue (BTB). David Tse is a lousy chess player; Deep Blue is very difficult to beat. Which schedule should you choose, to maximize your chances of winning the tournament? (NOTE: If you win the first and third game of chess but lose the second, you lose the tournament!)

**Answer:** Start by letting $p$ be the probability of beating Deep Blue and $q$ be the probability of beating David Tse. By the problem statement we necessarily have $p < q$.

1. What is the probability that you win the tournament, if you choose the TBT schedule?

The sample space is

$$\Omega = \{LLL, LLW, LWL, LWL, WLL, WLL, WLL, WLL, WWW\},$$

where, for instance, $LWL$ indicates that we lose the first game, win the second game, and lose the third game.

We can win in three ways: if we win all three games ($WWW$), if we win the first and second game and lose the third ($WWL$), or if we lose the first game and win the second and third game ($LWW$). These outcomes have probability $qpq$, $qp(1-q)$, and $(1-q)pq$, respectively. Thus, the probability of winning on the TBT schedule is:

$$P(\text{winning}) = qpq + qp(1-q) + (1-q)pq$$

$$= pq + 1 - q + 1 - q$$

$$= pq(2 - q).$$

2. What is the probability that you win the tournament, if you choose the BTB schedule?

The sample space is the same as in part 1, but with a different probability assignment. We can win in the same three ways: win all three games ($WWW$, with probability $pqp$), win the first and second
game and lose the third (WWL, probability $pq(1-p)$), lose the first and win the second and third game (LWW, probability $(1-p)qp$). Thus, the probability of winning on the BWB schedule is:

$$P(\text{winning}) = pqp + pq(1-p) + (1-p)qp$$

$$= pq(p + 1 - p + 1 - p)$$

$$= pq(2 - p)$$

The BTB schedule offers us a better chance of winning. The specific values of $p, q$ don’t matter. The problem statement implies $p < q$, from which it follows that $2-p > 2-q$. Since we know $0 < pq < 1$, we can multiply both sides by $pq$ to obtain $pq(2-p) > pq(2-q)$. Therefore the probability of winning under the BTB schedule is always strictly larger than the probability of winning under the TBT schedule.

**Comment:** This might seem surprising, because the math tells us to play chess-champion Deep Blue twice and play crummy amateur David Tse once, instead of the other way around. Surely it must be better to play the weaker player twice, right? Wrong.

If this seems counter-intuitive, here is what is going on. Due to the special victory condition, to win the tournament, you absolutely must win the second game; assuming you do so, you only need to win one out of the other two games. Under the TBT schedule, you have to win your one game against Deep Blue, and you get two chances to beat David Tse (you only need to beat him once). Under the BTB schedule, you get two chances to beat Deep Blue (it’s enough to beat Deep Blue once, you don’t have to win both games against Deep Blue), and then you have to win your one game against David Tse. Which would you rather have: two chances to beat Deep Blue and just one chance to beat David Tse, or two chances to beat David Tse and just one chance to beat Deep Blue? It’s better to give yourself as many chances to beat Deep Blue as possible.

The trick here is the special rule that to win the tournament you have to win two consecutive games. In comparison, under the standard rule that victory goes to whoever wins any two out of three games, then we get a more intuitive result: you are indeed better off minimizing the number of times you have to play Deep Blue and thus choosing the TBT schedule.

3. **(12 pts.) Conditional probability**

(a) **(6 pts.)** I have a bag containing either a $1 or $5 bill (with probability 1/2 for each of these two possibilities). I then add a $1 bill to the bag, so it now contains two bills. The bag is shaken, and you randomly draw a bill from the bag (without looking). Suppose it turns out to be a $1 bill. If a second student draws the remaining bill from the bag, what is the chance that it too is a $1 bill? Show your calculations.

**Answer 1:** The sample space is

$$\Omega = \{S1, D1, D5\},$$

with the following interpretation. The outcome $S1$ represents the case where the bag initially contains two $1 bills (same denomination) and you draw a $1 bill out of the bag. The outcome $D1$ represents the case where the bag initially contains a $1 and a $5 bill (different denomination) and you draw a $1 bill out of the bag. Finally, $D5$ represents the case where the bag initially contains $1 and $5, and you draw a $5 bill. By the assumption that the bag originally contains either $1 or $5 with equal probability,

$$P(S1) = \frac{1}{2},$$

Since the probabilities add up to 1 and $D1, D5$ have the same probability,

$$P(D1) = \frac{1}{4} \quad P(D5) = \frac{1}{4}.$$
Let \( A \) be the event that you draw out a $1 bill, and \( B \) be the event that the second student draws a $1 bill. The problem is to compute the conditional probability \( P(B|A) \). To do this, we will need to know \( P(A \cap B) \) and \( P(A) \).

The event \( A \) occurs for outcomes \( S1 \) and \( D1 \). Therefore,

\[
P(A) = P(S1) + P(D1) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.
\]

The event \( A \cap B \) occurs only for outcome \( S1 \), so

\[
P(A \cap B) = P(S1) = \frac{1}{2}.
\]

Now we can calculate

\[
P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}.
\]

**Answer 2:** We can calculate it yet another way. Let the events \( A, B \) be defined as above. Both students would draw $1 bills iff the first bill added was a $1 bill. Thus \( P(A \cap B) = 1/2 \). Event \( A \) occurs if either (a) the first bill added was $1 (with probability 1/2), and you draw a $1 (with probability 1), or (b) the first bill was $5 (with probability 1/2) and you draw the $1 (with probability 1/2). So \( P(A) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = 3/4 \). The rest follows as above.

(b) (6 pts.) Your gambling buddy found a website online where he could buy trick coins that are heads or tails on both sides. He puts three coins into a bag: one coin that is heads on both sides, one coin that is tails on both sides, and one that is heads on one side and tails on the other side. You shake the bag, draw out a coin at random, put it on the table without looking at it, then look at the side that is showing. Suppose you notice that the side that is showing is heads. What is the probability that the other side is heads? Show your work.

**Answer:** An outcome consists of us selecting a coin, and then choosing one of its sides. Since there are 3 coins each with 2 sides, our sample space \( \Omega \) has 6 elements

\[
\Omega = \{(HH, H_1), (HH, H_2), (HT, H), (HT, T), (TT, T_1), (TT, T_2)\},
\]

where \((HH, H_1)\) refers to the outcome of drawing the coin with two heads and of looking at the first side. Similarly, \((HH, H_2)\) refers to the outcome of choosing the coin with two heads and of looking at the second side, \((HT, T)\) refers to the outcome of choosing the coin with both heads and tails and of looking at the tails side, etc. Let \(A\) be the event that we choose the \(HH\) coin, and let \(B\) be the event that we see a heads when we put the coin down on the table. We wish to compute \(P(A|B)\). Since \(A = \{(HH, H_1), (HH, H_2)\}, \ B = \{(HH, H_1), (HH, H_2), (HT, H)\}, \) and \(A \cap B = A\), we see

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{|A|/|\Omega|}{|B|/|\Omega|} = \frac{2/6}{3/6} = \frac{2}{3}.
\]

**Comment:** This might be a bit surprising. For instance, it’s tempting to think: if I see a heads, then this means I drew either the \(HH\) or the \(HT\) coin, and both should be equally likely, so it’s 50-50 whether the other side is heads or tails. However, the calculation above shows that this argument is incorrect. The reason is that if we draw the \(HH\) coin, there are two ways to end up with it showing heads (according to which of the two sides is on top), whereas if we draw the \(HT\) coin, there is only one way to end up with it showing heads, so given that I see heads, it’s twice as likely that I drew the \(HH\) coin rather than the \(HT\) coin.

4. (12 pts.) **Learning in Monty Hall**

Consider the 4-door Monty Hall problem in HW1. Let us consider a model where the car is randomly behind one of the 4 doors but the contestant’s first choice is door 1 and the contestant follows the Stick-Stick strategy.
(a) \(2\) \text{ pts.} Compute and plot the probability distribution of the location of the car.

\textbf{Answer:} Let \(X \in \{1, 2, 3, 4\}\) be the location of the car, and \(Y_1, Y_2\) be Carol’s first and second door. By the assumption that the car is randomly behind one of the 4 doors, \(P(X = i) = 1/4\) for \(i = 1, \ldots, 4\).

(b) \(5\) \text{ pts.} Compute and plot the probability distribution of the location of the car, given that Carol’s first pick is door 2.

\textbf{Answer:} We would compute \(P(X = i|Y_1 = 2)\). Apparently \(P(X = 2|Y_1 = 2) = 0\).

\[
P(X = 1|Y_1 = 2) = \frac{P(X = 1, Y_1 = 2)}{P(Y_1 = 2)} = \frac{P(X = 1)P(Y_1 = 2|X = 1)}{P(Y_1 = 2)} = \frac{(1/4)(1/3)}{(1/3)} = 1/4.
\]

\[
P(X = 3|Y_1 = 2) = \frac{P(X = 3)P(Y_1 = 2|X = 3)}{P(Y_1 = 2)} = \frac{(1/4)(1/2)}{(1/3)} = 3/8.
\]

Similarly \(P(X = 4|Y_1 = 2) = 3/8\). 

\[\]
(c) **(5 pts.)** Compute and plot the probability distribution of the location of the car, given that Carol’s first pick is door 2 and her second pick is door 3.

**Answer:** We would compute \( P(X = i|Y_1 = 2, Y_2 = 3) \). Apparently \( P(X = 2|Y_1 = 2, Y_2 = 3) = P(X = 3|Y_1 = 2, Y_2 = 3) = 0 \).

\[
P(X = 1|Y_1 = 2, Y_2 = 3) = \frac{P(X = 1)P(Y_1 = 2, Y_2 = 3|X = 1)}{P(Y_1 = 2, Y_2 = 3)} = \frac{(1/4)(1/6)}{(1/6)} = 1/4.
\]

\[
P(X = 4|Y_1 = 2, Y_2 = 3) = 1 - P(X = 1|Y_1 = 2, Y_2 = 3) = 3/4.
\]

---

5. **(12 pts.) Getting FDA approval**

In the medical test example we discussed in lecture, we found that the false discovery rate of the medical is 0.81, unacceptably high for the FDA to approve. FDA wants a false discovery rate of at most 0.1. The pharmaceutical company can either invest in reducing the misdetection rate of the test, or reducing its false positive rate. (See definitions of these terms in Lecture Note 3.) What should it do? If the former, how much should the misdetection rate be reduced by? If the latter, how much should the false positive rate be reduced by?

**Answer:** Define the random variables \( H \) and \( T \) as in the lecture notes (\( H = 0 \) if the patient is healthy, \( H = 1 \) if affected, \( T = 0 \) if the test is negative, \( T = 1 \) if positive). As in the lecture notes, \( P(H = 1) = 1/20 \). Let the misdetection rate be \( P(T = 0|H = 1) = a \), and let the false positive rate be \( P(T = 1|H = 0) = b \).

To compute the false discovery rate, by Bayes rule,

\[
P(H = 0|T = 1) = \frac{P(H = 0)P(T = 1|H = 0)}{P(H = 0)P(T = 1|H = 0) + P(H = 1)P(T = 1|H = 1)} = \frac{(19/20)b}{(19/20)b + (1/20)(1 - a)} = \left(1 + \frac{1 - a}{19b}\right)^{-1}.
\]
Setting \((1 + \frac{1 - a}{10})^{-1} \leq 1/10\), we have \(1 - a \geq 171b\). As long as \(b = 0.2\), it is impossible to satisfy the constraint by decreasing \(a\). Hence we should decrease the false positive rate \(b\). When \(a = 0.1\), the constraint becomes \(1 - 0.1 \geq 171b\), and hence \(b \leq 0.00526\).

6. **(13 pts.) Being Discreet**
A non-profit wants to poll a sample of people to ask them whether they have ever had an extramarital affair. This being an extremely sensitive subject, one obvious problem is that if the surveyors ask this question straight-out, respondents may lie to avoid revealing personal information about their private lives.

The surveyers come up with the following clever scheme. They will ask the respondent to secretly roll a fair die. If the die comes up 1, 2, 3, or 4, the respondent is supposed to answer truthfully. If the die comes up 5 or 6, the respondent is supposed to answer the opposite of the truthful answer. The respondent is cautioned not to reveal what number came up on the die. Notice that if the respondent answers “Yes,” this answer is not necessarily incriminating: for all the surveyer knows, this particular respondent might have rolled a 5 or 6 and might have never had an affair in his/her life.

Let \(p\) be the probability that, if we select a person at random, then they will have had an extramarital affair. (Of course, the surveyers do not know \(p\); that is what they want to estimate.) Let \(q\) denote the probability that, if we select a person at random and have them follow the scheme above, then they will answer “Yes.”

*(a) (5 pts.)* Calculate a simple formula for \(q\), as a function of \(p\).

**Answer:** Let \(A\) be the event that the person had an affair, \(T\) the event that the person told the truth (i.e., that their die came up a 1, 2, 3, or 4), and \(S\) the event that the person says they had an affair.

We can use the Total Probability Rule to compute

\[
q = P(S) = P(S \cap T) + P(S \cap \overline{T})
\]

\[
= P(A \cap T) + P(A \cap \overline{T})
\]

\[
= P(A) \times P(T) + P(A) \times P(\overline{T})
\]

\[
= p \times \frac{4}{6} + (1 - p) \times \frac{2}{6}
\]

\[
= \frac{1}{3}p + \frac{1}{3}.
\]

In the second line we used the fact that \(P(E) = P(E \cap F) + P(E \cap \overline{F})\) holds for any pair of events \(E, F\) (this is the Total Probability Rule). In the third line we used the fact that \(S \cap T = A \cap T\) (i.e., a person tells the truth and says they have an affair if and only if they tell the truth and really had an affair) and \(S \cap \overline{T} = A \cap \overline{T}\) (i.e., a person lies and says they had an affair if and only if they lie and didn’t have an affair). In the fourth line we used the independence of the person’s fidelity and their dice roll.

Therefore, the formula is \(q = \frac{1}{3}p + \frac{1}{3}\).

*(b) (3 pts.)* Suppose the surveyors have surveyed 1000 people. What is a sensible way of estimating \(q\)?

**Answer:** We simply use the empirical probability as the estimate of \(q\).

\[
\hat{q} = \frac{\text{Number of respondents answering yes}}{1000}
\]

*(c) (5 pts.)* Next, suppose the surveyors have estimated \(q\). Now they want to solve for \(p\). Find a simple formula for \(p\), as a function of \(q\).
**Answer:** Solving for $p$ as a function of $q$ gives

\[
\frac{1}{3} + \frac{1}{3} = q \\
p + 1 = 3q \\
p = 3q - 1.
\]

Therefore, the formula is $p = 3q - 1$.

Note: This procedure of estimating $q$ first and then assuming the estimate is exact to solve for $p$ is not quite statistical sound. For one thing, the estimate of $q$ is not exactly $q$ (like in the simulation example of Lab 0), and if that estimate happens to be outside the range $[1/3, 2/3]$, then the resulting estimate of $p$ using the above formula will result in an estimate that is outside the $[0, 1]$ range: senseless! Later on in the class we will describe a statistically sound procedure called *maximum likelihood*, and we will return to this example.

7. (13 pts.) Communication Channel

(a) (6 pts.) Find the probabilities of the output symbols.

**Answer:** Let $X$ be a random variable denoting the value of the input and $Y$ be a random variable denoting the output. Using the law of total probability and conditional probability we get

\[
P(Y = 0) = P(Y = 0|X = 0)P(X = 0) + P(Y = 0|X = 2)P(X = 2) = \frac{1}{2}(1 - \epsilon) + \frac{1}{4}\epsilon = \frac{1}{2} - \frac{\epsilon}{4},
\]

\[
P(Y = 1) = \frac{1}{2}\epsilon + \frac{1}{4}(1 - \epsilon) = \frac{1}{4} + \frac{\epsilon}{4},
\]

\[
P(Y = 2) = \frac{1}{4}\epsilon + \frac{1}{4}(1 - \epsilon) = \frac{1}{4}.
\]

As expected, the sum of the probabilities of the events $B_i$ is 1.

(b) (7 pts.) Given that 1 is received, find the probabilities that the input was 0, 1, or 2

**Answer:** The conditional probabilities are

\[
P(X = 0|Y = 1) = \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{\frac{1}{2}\epsilon}{\frac{1}{4} + \frac{\epsilon}{4}} = \frac{2\epsilon}{1 + \epsilon},
\]

\[
P(X = 1|Y = 1) = \frac{\frac{1}{4}(1 - \epsilon)}{\frac{1}{4} + \frac{\epsilon}{4}} = \frac{1 - \epsilon}{1 + \epsilon},
\]

\[
P(X = 2|Y = 1) = 0.
\]

Again, note that the sum of the conditional probabilities is 1.
8. **Serve on the jury**

In the OJ Simpson murder trial, OJ Simpson was accused of murdering his ex-wife, Nicole Simpson. The prosecution introduced evidence showing that OJ had previously abused Nicole. One of Simpson’s defense lawyers, Alan Dershowitz, made the following argument in OJ Simpson’s defense. Dershowitz stated that 1 in 1,000 women abused by their husbands are later killed by their abuser, so the fact that OJ Simpson had previously abused his wife is not relevant and should be disregarded. Assume for this problem that Dershowitz’s 1 in 1,000 statistic is accurate.

(a) **(4 pts.)** Are we entitled to conclude that there is only a 1/1000 probability that OJ Simpson murdered Nicole? Why or why not?

**Answer 1:** No, we are not entitled to conclude that there is a 1/1000 probability that OJ Simpson murdered Nicole.

If we define events \( M \) and \( G \) as in part 2, then the probability that an abused woman gets murdered by her abuser is \( P(G) = \frac{1}{1000} \). However, this probability is diluted by the fact that many women are never murdered (by anyone).

In this case, we also know that Nicole was murdered. To take into account the fact that we know Nicole was murdered, we should be looking at \( P(G|M) \), the probability that a woman is murdered by her abuser given that she is murdered.

**Answer 2:** No, we are not entitled to conclude that there is a 1/1000 probability that OJ Simpson murdered Nicole. At the trial, the prosecutors produced all sorts of other evidence to support the contention that OJ murdered Nicole; that evidence presumably should increase our estimate of the likelihood that OJ murdered Nicole. Similarly, we should presumably also take into account any evidence introduced by the defense lawyers which indicates that OJ did not murder Nicole; such evidence would presumably reduce our estimate of the likelihood that OJ murdered Nicole.

(b) **(6 pts.)** Suppose we select at random a woman who has been abused by her husband. Define the following events: \( M \) is the event that the woman is murdered at some point in her life; \( G \) is the event that the woman is murdered by her abuser at some point in her life. A plausible estimate is that 0.2% of abused women will be murdered by someone other than their abuser at some point in their life. Calculate the probability that the selected woman is murdered by her abuser, given that she is murdered.

**Answer:**

![Venn diagram](image)
We know from the problem statement that:

\[ P(G) = \frac{1}{1000} \]
\[ P(G \cap M) = \frac{2}{1000}. \]

We need to determine \( P(G|M) \). We can calculate it, using the definition of conditional probability:

\[
P(G|M) = \frac{P(G \cap M)}{P(M)} = \frac{\frac{2}{1000}}{\frac{1}{1000} + \frac{2}{1000}} = \frac{\frac{2}{1000}}{\frac{3}{1000}} = \frac{2}{3}.
\]

Here we have used the fact that \( G \subseteq M \), so \( G \cap M = G \). In other words, if a woman is murdered by her abuser, she is certainly murdered, so \( G \) and \( M \) are both true if and only if \( G \) is true.

In summary, our final answer is \( P(G|M) = \frac{2}{3} \).

(c) (3 pts.) Based upon your answer to part (b), do you agree or disagree with Dershowitz’s argument? Based upon your calculation, would you consider it relevant that OJ Simpson previously abused Nicole? Would you judge it more accurate to use the 1 in 1,000 number or the number you calculated in part (b)? Why?

**Answer:** Dershowitz's argument is wrong. Based upon our calculations, it is relevant to know that OJ previously abused Nicole. We know that 1 in 1,000 abused women are later killed by their abusers. However, in this case we have the additional information that Nicole was murdered. Out of all murdered abused women, 1 in 3 women are murdered by their abusers. This number is more appropriate to use, because we know that Nicole was murdered. In other words, the conditional probability \( P(G|M) = \frac{2}{3} \) better takes into account all of the information available to us, and thus is more appropriate to use in this situation than the unconditional probability \( P(G) = \frac{1}{1000} \).

8. (13 pts.) M&M. (Alternate question) The blue M&M was introduced in 1995. Before then, the color mix in a bag of plain M&Ms was (30% Brown, 20% Yellow, 20% Red, 10% Green, 10% Orange, 10% Tan). Afterward it was (24% Blue, 20% Green, 16% Orange, 14% Yellow, 13% Red, 13% Brown).

(a) (6 pts.) A friend of mine has two bags of M&Ms, and he tells me that one is from 1994 and one from 1996. He won’t tell me which is which, but he gives me one randomly chosen M&M from a randomly chosen bag. It is yellow. What is the probability that the yellow M&M came from the 1994 bag?

**Answer:** Let \( X \) be the random variable denoting which bag is chosen (\( X = 0 \) for the 1994 bag, \( X = 1 \) for the 1996 bag). Let \( Y \) be the color of the randomly chosen M&M from the bag. By Bayes rule,

\[
P(X = 0|Y = \text{yellow}) = \frac{P(X = 0)P(Y = \text{yellow}|X = 0)}{P(X = 0)P(Y = \text{yellow}|X = 0) + P(X = 1)P(Y = \text{yellow}|X = 1)}
\]
\[
= \frac{(1/2) \cdot 0.2}{(1/2) \cdot 0.2 + (1/2) \cdot 0.14}
\]
\[
= 0.588.
\]
(b) (7 pts.) Suppose in addition the friend also gives me a randomly chosen M&M from the other bag and it is green. Given this observation, what is the probability that the yellow M&M came from the 1994 bag?

Answer: Let $Z$ be the color of the randomly chosen M&M from the other bag. By Bayes rule,

$$
P(X = 0|Y = \text{yellow}, Z = \text{green}) = \frac{P(X = 0)P(Y = \text{yellow}, Z = \text{green}|X = 0)}{P(X = 0)P(Y = \text{yellow}, Z = \text{green}|X = 0) + P(X = 1)P(Y = \text{yellow}, Z = \text{green}|X = 1)}$$

$$= \frac{\frac{1}{2} \cdot 0.2 \cdot 0.2}{(\frac{1}{2}) \cdot 0.2 \cdot 0.2 + \frac{1}{2} \cdot 0.14 \cdot 0.1}$$

$$= 0.741.$$