Fun Questions about Conditional Probability

Jeffrey S. Rosenthal, University of Toronto, July 2026

One of the most intriguing and challenging topics in elementary probability theory is *conditional* probability. The conditional probability of an event A given another event B, written $\mathbf{P}(A|B)$, represents the "updated" probability of A, conditional on *knowing* that some other event B has occurred. That is, $\mathbf{P}(A|B)$ is the *fraction* of those times when B occurs, that A also occurs. Formally, it is defined as $\mathbf{P}(A|B) = \frac{\mathbf{P}(A \text{ AND } B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$. (We assume that $\mathbf{P}(B) > 0$, though see the Continuous Case section below.) However, even in very simple-seeming examples, the correct value of $\mathbf{P}(A|B)$ can sometimes be unexpected and surprisingly elusive, as we shall see herein.

In this paper, we present a bunch of elementary conditional probability questions, which illustrate the subtleties of this topic in fun and educational ways. Furthermore, the questions are all designed so that their answer is one of the following six multiple-choice answers:

(a) 1/4. (b) 1/3. (c) 1/2. (d) 2/3. (e) 3/4. (f) 4/5.

Hence, these questions are well suited for use with class polling software such as iClicker or PollEverywhere or Google Forms, using these same multiple-choice answer choices over and over again. To allow readers to ponder the questions themselves, we defer presenting the answers until this paper's final section. Readers are urged to try to solve the questions themselves, before checking the answer section; you might be surprised!

Colleague's Two Children

For our first questions, assume that your colleague has two children, one Older and one Younger, each independently equally likely (probability 1/2 each) to be a Boy or a Girl. An easy warm-up question (not even conditional probability!) is:

Q1: What is $\mathbf{P}$(Both Girls), the probability that both children are Girls?

We now introduce conditional probability. Suppose first that you are told that at least one of the two children is a Girl, i.e. they are not both Boys. How does this affect your assessment of the probability that they are both Girls? That is:

Q2: What is $\mathbf{P}$(Both Girls | At Least One Girl)?

We next consider a variant in which you are told not just that at least one child is a Girl, but specifically that the *Older* child is a Girl. How does that affect your assessment of the probability that both children are Girls? Is it the same answer as before, or is it different now that your information is specifically about the Older child? That is:

Q3: What is $\mathbf{P}$(Both Girls | Older Child is Girl)?

For another variation, suppose you learn that a *randomly* chosen child is a Girl. (For example, perhaps you just happen to spot one of their children.) Assume you were equally likely to see either one of their children. Then how does that observation affect the probability that both children are Girls? Is it the same as the previous answer? That is:

Q4: What is $\mathbf{P}$(Both Girls | Randomly-Chosen Child Is Girl)?

Comparing the two previous questions raises the issue of, once you've seen that a randomly-chosen child is a Girl, does it then matter if someone tells you it was the Older child? Does that change the conditional probability? That is, do the two previous questions have different answers, or the same answer? I once got into a heated argument about this with a math professor (who, yes, had a PhD in mathematics). I was right, and he was wrong, but I sure had a tough time convincing him.

Drawers Of Riches

Suppose now that a drawer contains 3 Gold pieces, and 1 Silver piece. You reach in (in the dark) and take out two pieces, uniformly at random. We can then ask:

Q5: What is $\mathbf{P}$(Both pieces are Gold)?

To make it more interesting, we now suppose that there are *two* drawers, each with Gold and Silver Pieces. Specifically, let's assume that Drawer #1 consists of Gold-Silver, i.e. 1 Gold and 1 Silver, while Drawer #2 consists of Gold-Silver-Silver-Silver-Silver-Silver, i.e. 1 Gold and 5 Silver. Suppose we first choose one of the two drawers, uniform at random. Then, we choose one Piece within that drawer, uniformly at random. For that set-up, we can ask about that chosen Piece:

Q6: What is $\mathbf{P}$(Piece is Gold)?

We can then look at the conditional probabilities for the chosen drawer. Of course, originally we have $\mathbf{P}(\text{Drawer \#1}) = 1/2$. But suppose we see (after moving into the light) that the Piece we have chosen is actually Gold. How does that affect our assessment of which drawer we chose? Is it still 1/2? That is:

Q7: What is $\mathbf{P}$(Drawer #1 | Piece is Gold)?

Three Card Thriller

Suppose there are three Cards, each with two sides. Card #1 is Blue-Blue, i.e. Blue on both sides. Card #2 is Yellow-Yellow, i.e. Yellow on both sides. Card #3 is Blue-Yellow, i.e. Blue on one side and Yellow on the other. We choose a Card uniformly at random. Then we show one Side of that Card, uniformly at random. A first question is:

Q8: What is $\mathbf{P}$(Side is Blue)?

We can then investigate how observing a Blue side affects our probabilities for which card was chosen. Originally we had $\mathbf{P}(\text{chose Card \#1}) = 1/3$. We can now ask:

Q9: What is $\mathbf{P}(\text{chose Card \#1} \mid \text{Side is Blue})$?

Two-Headed Coins and Bayesian Statistics

It is a true fact that I am the proud owner of a two-headed coin, which I purchased from a magic shop. Now, suppose I choose *either* my two-headed coin, *or* a regular coin, each with probability 1/2. So, $\mathbf{P}(\text{Coin is Two-Headed}) = 1/2$. I do not show you which coin I have chosen. I do, however, *flip* my chosen coin once. For that random flip, we can ask:

Q10: What is $\mathbf{P}(\text{Heads})$, the probability that the coin flip shows Heads?

Suppose now that the flip does indeed come up *Heads*. Then how does that affect the probability that I chose the two-headed coin? That is:

Q11: What is $\mathbf{P}(\text{Two-Headed} \mid \text{Heads})$, the *conditional* probability that my chosen coin is two-headed, given that I flipped it once and got Heads?

In the context of Bayesian statistics, we can think of the original probability 1/2 for the two-headed coin as the "prior" probability, and the Heads result from the flip as the "data", and the conditional probability $\mathbf{P}(\text{Two-Headed} \mid \text{Shows Heads})$ as the "posterior" probability. So, the above question can be rephrased as asking for the posterior probability of a two-headed coin, given the uniform prior and the single Heads data.

Now, what if instead I get *two* Heads in a row? That is, I choose my coin at random just like before, but now I flip it *twice*, and get Heads both times. In that case:

Q12: What is $\mathbf{P}(\text{Two-Headed} \mid \text{Get Two Heads})$?

(And while you're at it, you could also consider what happens if you flip the coin n times, and get Heads every single time.)

Disease Assessment

Another interesting application of conditional probability is to disease testing. Typical medical tests are not perfect, but rather have certain *false positive* and *false negative* rates. Although these rates are usually small, they can still have a significant effect on conditional probabilities.

To be specific, suppose a certain disease affects 1/10 of the population, i.e. $\mathbf{P}(\text{Disease}) = 1/10$. Suppose there is a certain test for this disease. The test has *zero false negative rate*, i.e. if a patient is diseased then the test always says "Yes". However, the test has *1/3 false positive rate*, i.e. if a patient is healthy then the test still has probability 1/3 of saying "Yes". Suppose a random person tests positive ("Yes") for the disease. Then we can ask:

Q13: What is $\mathbf{P}(\text{Disease} \mid \text{Yes})$, the conditional probability that they have the disease given that the test was positive?

(And while you're at it, also consider a more extreme case, where say $\mathbf{P}(\text{Disease})=1/10{,}000$, with 1% false positive rate.)

Monty Hall Problem

The Monty Hall Problem was published in 1990, and quickly became the most famous probability puzzle of all time. So, you have probably heard of it. But do you understand it as well as you think you do?

To recap, there are three Doors. One of them (chosen uniformly at random) has a Car behind it, while the other two have Goats. The player initially guesses one door (say, Door #1). The host then opens a *different* door (say, Door #3), revealing a Goat. The player is then given the option to stick with their original door (#1), or switch to the other closed door (#2). Should they switch? That is, what is the *conditional* probability that the other closed door (#2) has the car, conditional on the host's actions?

My answer to this, is that it *depends* precisely what we are "conditioning" on. To be specific, assume you initially guessed Door #1, and then the Host revealed that Door #3 has a Goat. What is the "this" that you are conditioning on? One interpretation is simply "conditional on Door #3 having a Goat", i.e. the Host just *fell* into Door #3 which just *happened* to have a Goat. (This has been called the "Monty Fall Problem".) In that case, the question becomes:

Q14: What is $\mathbf{P}(D2 = \text{Car} \mid D3 = \text{Goat})$, the conditional probability that Door #2 has a Car, given that Door #3 has a Goat?

However, another interpretation (which was apparently intended in the original problem) is that we condition on the Host *choosing* to open Door #3, which he *knows* has a Goat. That is, the Host will always open one of the doors that you did not choose (i.e. Door #2 or #3), and will always make sure to choose a door which has a Goat. To be precise, assume that if the Host has a choice of doors (i.e., if Door #1 has a Car, so Doors #2 and #3 both have Goats and the Host could choose either one of them), then he will choose uniformly, i.e. make either choice with probability 1/2 each. Under this interpretation:

Q15: What is $\mathbf{P}(D2 = \text{Car} \mid H = 3)$, the conditional probability that Door #2 has a Car given that the Host *chose* to open Door #3 (which had a Goat)?

Is the answer the same under these two different interpretations? Perhaps not! Indeed, confusion between these two interpretations seems to be what caused this question to become so controversial. For more discussion, see my article at: probability.ca/montyfall

Sleeping Beauty Problem

A different but related question comes from philosophers, who posited a *Sleeping Beauty Problem* defined as follows. A fair coin (say a Nickel) is flipped. If the Nickel is Heads, Beauty is interviewed on Monday only. If the Nickel is *Tails*, Beauty is interviewed on both Monday *and* Tuesday. She is given an "amnesia" (forgetting) drug between interviews, so she cannot remember any previous interviews, and furthermore she is not told what day it is. In each interview, she is asked for *her probability* that the Nickel was Heads. Then a big and controversial epistemological question is:

Q16: In this scenario, what should Beauty say is $\mathbf{P}$(Nickel is Heads)?

Unlike all of the other questions in this paper, this one does not have a definitive answer. Some philosophers ("halfers") argue that it should be 1/2, since originally $\mathbf{P}$(Nickel is Heads) = 1/2, and Beauty already knew she was going to be interviewed at least once, so the probability should still be 1/2. Others ("thirders") argue that it should be 1/3, because on average Beauty is interviewed twice as often when the coin is Tails as when it is Heads, so the probability of Tails should be twice as large. Many philosophical papers and much philosophical debate have ensued. The problem seems to be that we want to condition on the event that "Beauty is currently being interviewed", but it is not clear how to formulate such conditioning properly.

When I heard about this problem, I initially struggled. Then I came up with what I thought was an elegant solution, as follows. Suppose that before the interviews, in addition to the Nickel, we also flip a (fair) *Dime*. Then we turn the Dime *over* on Monday night. In this way, Beauty is only interviewed at most *once* while the Dime is Heads (and similarly for Tails). Assume Beauty is allowed to see the Dime (though not the Nickel of course). Then it seems more clear what to condition on. Namely, if Beauty sees that the Dime is Heads, then she can do her conditioning on the event that "Beauty was interviewed while the Dime was Heads". Since she would never be interviewed more than once while the Dime is Heads, this seems like a clear event for conditioning. Then we can ask:

Q17: What is $\mathbf{P}$(Nickel=Heads | Dime=Heads), the conditional probability that the Nickel was Heads, given that Beauty can see that the Dime is Heads during her interview?

This last question provides a clear mathematical calculation for the probability that the Nickel was Heads, given that the Dime was Heads during the interview. By symmetry, the probability that the Nickel was Heads, given that the Dime was Tails during the interview is exactly the same. So, this tells Beauty what she should say is $\mathbf{P}$(Nickel is Heads) whether the Dime shows Heads or Tails. Hence, it stands to reason that she should still say the same thing, even if she cannot see the Dime or it does not exist.

I felt that this argument provided a clear and unambiguous answer to the Sleeping Beauty problem. However, the philosophers were not as convinced as I'd hoped. Oh well. For more details, see my article at: probability.ca/beauty

Continuous Case

Finally, we say a few words about the continuous case. All of our conditional probability so far has been of the form $\mathbf{P}(A \,|\, B)$ where $\mathbf{P}(B) > 0$. But what if $\mathbf{P}(B) = 0$? What if we want to condition on continuous variables, for which individual probabilities are all zero. Can we do so at all? The answer is, "sort of", at least in a certain context.

To be specific, suppose X is a real-valued absolutely continuous random variable, so it has a *density* function $f(x)$, meaning that $\mathbf{P}(a < X < b) = \int_a^b f(x)\,dx$ for all real numbers $a < b$. Suppose we are interested in some other event, like say the probability that some other random variable Y is bigger than 5, and we want to *condition* on a particular value of X like $X = 2$. Does an expression like $\mathbf{P}(Y > 5 \,|\, X = 2)$ make any sense? Not obviously! Since X is continuous, $\mathbf{P}(X = 2) = \int_2^2 f(x)\,dx = 0$. So the expression $\mathbf{P}(Y > 5 \,|\, X = 2)$ is like $\mathbf{P}(A|B)$, but in the case where $\mathbf{P}(B) = 0$. This is undefined! What can be done?

Some hope comes from the law of total probability. We know that we always have e.g. $\mathbf{P}(A \cap B) + \mathbf{P}(A \cap B^C) = \mathbf{P}(A)$, and hence $\mathbf{P}(B)\ \mathbf{P}(A|B) + \mathbf{P}(B^C)\ \mathbf{P}(A|B^C) = \mathbf{P}(A)$. More generally, if $\{B_i\}_{i=1}^n$ form a *partition*, meaning they are disjoint events whose union is the entire space, then we should have $\sum_{i=1}^n \mathbf{P}(B_i)\ \mathbf{P}(A|B_i) = \mathbf{P}(A)$.

Generalising this principle to the continuous case, if a quantity like $\mathbf{P}(Y > 5 \,|\, X = x)$ has any meaning, then we should have:

$$\int_{-\infty}^{\infty} f(x)\ \mathbf{P}(Y > 5 \,|\, X = x)\ dx \ = \ \mathbf{P}(Y > 5)\,.$$

And more generally, for any $r < s$, we should have

$$\int_r^s f(x)\ \mathbf{P}(Y > 5 \,|\, X = x)\,dx \ = \ \mathbf{P}(Y > 5 \text{ AND } \ r < X < s)\,.$$

So far, this doesn't tell us what $\mathbf{P}(Y > 5 \,|\, X = x)$ should equal, just that it should satisfy certain properties. But it does specify what $\int_r^s f(x)\ \mathbf{P}(Y > 5 \,|\, X = x)\ dx$ should equal, for all $r < s$. Hence, this actually does specify the value of $\mathbf{P}(Y > 5 \,|\, X = x)$, not for *all* x, but for *almost all* x. Which isn't perfect, but it's much better than nothing.

This is the essence of the "sub-σ-algebra" approach to continuous conditional probability. We cannot directly condition on a probability-zero event like $\{X = 2\}$. However, we can condition on the σ-algebra of all events which depend only on the variable X. We can do this in a consistent way, to get quantities which are defined up to sets of measure zero, i.e. for almost all (but not all) values of x. To define such continuous conditioning properly requires technical measure theory arguments. But the gist is that it is, kinda, sorta, possible to condition on continuous variables like X, and hence on probability-zero events like $\{X = 2\}$. You just have to be very careful, and accept that the conditioning is only defined for almost all values. Otherwise, it's best to stick to conditioning for discrete models where $\mathbf{P}(B) > 0$.

Answers to all of the Questions

We now finally come to the answers to all of the above questions. Remember that each answer is one of the following six choices:

(a) 1/4. (b) 1/3. (c) 1/2. (d) 2/3. (e) 3/4. (f) 4/5.

The individual answers (with justification) are as follows:

A1: Since the two children are assumed to be independent, of course $\mathbf{P}(\text{Both Girls}) = \mathbf{P}(\text{Older is Girl}) \times \mathbf{P}(\text{Younger is Girl}) = (1/2)(1/2) = 1/4$. (a)

A2: Setting $A = \{\text{Both Girls}\}$ and $B = \{\text{At least One Girl}\}$, we see that $A \cap B = A = \{GG\}$ is simply the event that the Older and Younger children are both Girls. And, $B = \{GG, GB, BG\}$ is the event that either they are both Girls, or the Older is a Girl while the Younger is a Boy, or the Older is a Boy while the Younger is a Girl. Hence, $\mathbf{P}(\text{Both Girls} \mid \text{At Least One Girl}) = \mathbf{P}(A \mid B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}\{GG\}}{\mathbf{P}\{GG,GB,BG\}} = \frac{1/4}{3/4} = 1/3$. (b)

A3: Setting $A = \{\text{Both Girls}\}$ and $B = \{\text{Older is Girl}\}$, we again have $A \cap B = A = \{GG\}$, but here $B == \{GG, GB\}$. Hence, $\mathbf{P}(\text{Both Girls} \mid \text{Older is Girl}) = \mathbf{P}(A \mid B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}\{GG\}}{\mathbf{P}\{GG,GB\}} = \frac{1/4}{2/4} = 1/2$. (c)

A4: Setting $A = \{\text{Both Girls}\}$ and $B = \{\text{Randomly-Chosen is Girl}\}$, we again have $A \cap B = A = \{GG\}$, but here B is more complicated. It seems that by *symmetry*, a randomly-chosen child should be equally likely to be a Girl or Boy, so we must have $\mathbf{P}(B) = 1/2$. To reason more formally, we can let R be the randomly-chosen child, and then use the *law of total probability*, i.e. add up the different ways that we could have $R = \text{Girl}$: $\mathbf{P}(B) = \mathbf{P}(R{=}\text{Girl}) = \mathbf{P}(GG \text{ AND } R{=}\text{Girl}) + \mathbf{P}(GB \text{ AND } R{=}\text{Girl}) + \mathbf{P}(BG \text{ AND } R{=}\text{Girl}) + \mathbf{P}(BB \text{ AND } R{=}\text{Girl}) = (1/4)(1) + (1/4)(1/2) + (1/4)(1/2) + 0 = 1/2$. We then compute that $\mathbf{P}(\text{Both Girls} \mid \text{R=Girl}) = \mathbf{P}(A \mid B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(GG,\ R=G)}{\mathbf{P}(R=G)} = \frac{(1/4)(1)}{1/2} = 1/2$. (c)

This is indeed the *same* answer as the previous question. So, it doesn't matter if the randomly-seen child was older! That may be surprising, since the previous question eliminated $\{BB, BG\}$, leaving $\{GG, GB\}$, while this question eliminates just $\{BB\}$, leaving $\{GG, GB, BG\}$. However, seeing that a randomly-chosen Child is a Girl also *changes* some of the other probabilities, making $\{GG\}$ more likely but $\{GB\}$ and $\{BG\}$ less likely, which all cancels out in the end.

A5: Although this question does not directly involve conditional probability, one possible solution does. We can use the *joint probability formula*, which notes that since $\mathbf{P}(A|B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)}$, therefore $\mathbf{P}(A \cap B) = \mathbf{P}(B)\, \mathbf{P}(A|B)$, i.e. the joint probability equals the first probability times the conditional probability of the second given the first. So, to compute the joint probability of two events, we can first compute one probability, and then compute the other probability *conditionally* on the first one. In this case, we therefore compute that $\mathbf{P}(\text{Both Gold}) = \mathbf{P}(\text{First Gold}) \times \mathbf{P}(\text{Second Gold} \mid \text{First Gold}) = (3/4) \times (2/3) = 2/4 = 1/2$. (c)

(As an aside, a similar approach is useful for e.g. computing probabilities when dealing multiple cards from a deck, etc. And, the joint probability formula is also the basis of the

Gibbs Sampler (Glauber Dynamics) Monte Carlo algorithm, which first samples $X|Y$, then $Y|X$, etc. That algorithm preserves the joint stationarity distribution of the two variables, precisely because of the joint probability formula.)

For an alternative solution, note that there were $\binom{4}{2}$ ways to select two pieces out of four, of which $\binom{3}{2}$ are selected just from the three Gold pieces. Hence, the probability that the two pieces are both Gold must be equal to $\binom{3}{2}/\binom{4}{2} = (3 \cdot 2/2)/(4 \cdot 3/2) = 2/4 = 1/2$ (c).

A6: We can again use the law of total probability, i.e. add up all the different possible ways: $\mathbf{P}$(Piece is Gold) = $\mathbf{P}$(Drawer #1 AND Piece is Gold) + $\mathbf{P}$(Drawer #2 AND Piece is Gold) $= (1/2)(1/2) + (1/2)(1/6) = (1/4) + (1/12) = 4/12 = 1/3$. (b)

A7: Using the previous answer, we compute that $\mathbf{P}$(Drawer #1 | Piece is Gold)

$$= \frac{\mathbf{P}(\text{Drawer \#1 AND Piece is Gold})}{\mathbf{P}(\text{Piece is Gold})} = \frac{(1/2)(1/2)}{1/3} = 3/4. \text{ (e)}$$

Hence, seeing that the Piece is Gold makes it more likely that we chose Door #1.

A8: By symmetry, we must have $\mathbf{P}$(Side is Blue) $= 1/2$. Or, to calculate this more systematically (which would be required in a less symmetric case), we again use the law of total probability: $\mathbf{P}$(Side is Blue) = $\mathbf{P}$(Card #1 AND Side is Blue) + $\mathbf{P}$(Card #2 AND Side is Blue) + $\mathbf{P}$(Card #3 AND Side is Blue) $= (1/3)(1) + (1/3)(0) + (1/3)(1/2) = 1/2$. (c)

A9: Here $\mathbf{P}(\text{chose Card \#1} \mid \text{Side is Blue}) = \dfrac{\mathbf{P}(\text{chose Card \#1 AND Side is Blue})}{\mathbf{P}(\text{Side is Blue})}$

$$= \frac{\mathbf{P}(\text{chose Card \#1})\ \mathbf{P}(\text{Side is Blue} \mid \text{chose Card \#1}))}{\mathbf{P}(\text{Side is Blue})} = \frac{(1/3)(1)}{1/2} = 2/3. \text{ (d)}$$

This answer may seem surprising, since seeing that the Side is Blue eliminates Card #3 but still leaves Cards #1 and #2 as possibilities. However, it also makes Card #2 less likely, since that card could have shown Blue but could have shown Yellow. So, in the end, Card #1 is twice as likely. Another way to think about this is that the visible Side could have been any of the three different blue *sides* among all the cards, and they are then all equally likely, with two of them (2/3) coming from Card #1.

A10: We again use the law of total probability, to compute that $\mathbf{P}$(Heads) = $\mathbf{P}$(Two-Headed coin AND Heads) + $\mathbf{P}$(Regular coin AND Heads) $= (1/2)(1) + (1/2)(1/2) = 3/4$. (e)

A11: Here we compute that $\mathbf{P}$(Two-Headed | Shows Heads)

$$= \frac{\mathbf{P}(\text{Two-Headed AND Shows Heads})}{\mathbf{P}(\text{Shows Heads})} = \frac{(1/2)(1)}{(1/2)(1)+(1/2)(1/2)} = \frac{1/2}{3/4} = 2/3. \text{ (d)}$$

A12: The calculation is very similar to the above, except this time we have to *square* the probabilities of getting Heads for each coin. Hence, $\mathbf{P}$(Two-Headed | Get Two Heads) =

$$\frac{\mathbf{P}(\text{Two-Headed AND Two Heads})}{\mathbf{P}(\text{Two Heads})} = \frac{(1/2)(1)^2}{(1/2)(1)^2+(1/2)(1/2)^2} = \frac{1/2}{5/8} = 4/5. \text{ (f)}$$

We could continue this further, and imagine we repeat and get n Heads in a row. Then $\mathbf{P}(\text{Two-Headed} \mid n \text{ Heads}) = \dfrac{\mathbf{P}(\text{Two-Headed AND } n \text{ Heads})}{\mathbf{P}(n \text{ Heads})} = \frac{(1/2)(1)^n}{(1/2)(1)^n+(1/2)(1/2)^n}$ $= \frac{1/2}{1/2+(1/2)^{n+1}} = \frac{1}{1+(1/2)^n} = \frac{2^n}{2^n+1} \to 1$ as $n \to \infty$.

That is, as we flip more and more coins, and get Heads every time, it becomes more and more likely that we have the Two-Headed coin, and less and less likely that it is an ordinary (fair) coin. In the Bayesian context, this illustrates the principle that "the data swamps the prior", i.e. the more data we get, the more certain an answer we get, and the less we care that the original probability of the Two-Headed coin was 1/2.

A13: Here $\mathbf{P}(\text{Disease} \,|\, \text{Yes}) = \dfrac{\mathbf{P}(\text{Disease AND Yes})}{\mathbf{P}(\text{Yes})} = \frac{(1/10)(1)}{(1/10)(1)+(9/10)(1/3)} = \frac{1/10}{4/10} = 1/4$. (a)

Or, in the more extreme case with $\mathbf{P}(\text{Disease})$=1/10,000 and 1% false positive rate, $\mathbf{P}(\text{Disease} \mid \text{Yes}) = \frac{(1/10000)(1)}{(1/10000)(1)+(9999/10000)(1/100)} \doteq 1/101$. This is quite small! So, even a positive test result with a very small false positive rate does not necessarily indicate that you are unwell – which is good to know.

A14: We compute here that $\mathbf{P}(D2 = \text{Car} \,|\, D3 = \text{Goat}) = \frac{\mathbf{P}(D2=\text{Car AND } D3=\text{Goat})}{\mathbf{P}(D3=\text{Goat})} = \frac{(1/3)(1)}{2/3} = 1/2$. (c)

Similarly, $\mathbf{P}(D1 = \text{Car} \,|\, D3 = \text{Goat}) = 1/2$, too. Each of the unopened doors has equal chance of having the Car in this case.

A15: Under this interpretation, $\mathbf{P}(H = 3 \,|\, D2 = \text{Car}) = 1$ since the Host has no choice, but $\mathbf{P}(H = 3 \,|\, D1 = \text{Car}) = 1/2$ since the Host could choose either door. So, $\mathbf{P}(D2 = \text{Car} \,|\, H = 3) = \frac{\mathbf{P}(D2=\text{Car AND } H=3)}{\mathbf{P}(H=3)} = \frac{(1/3)(1)}{1/2} = 2/3$. (d)

Similarly, $\mathbf{P}(D1 = \text{Car} \,|\, H = 3) = \frac{\mathbf{P}(D1=\text{Car AND } H=3)}{\mathbf{P}(H=3)} = \frac{(1/3)(1/2)}{1/2} = 1/3$.

So, under this interpretation, you should *switch* from #1 to #2! Intuitively, the Host opening Door #3 *tells you* that it is (conditionally) more likely that Door #2 has the Car.

A16: Some say 1/2 (c), some say 1/3 (b), and there is much debate!

A17: Let $A = \{\text{Nickel=Heads}\}$, and $B = \{\text{Interviewed while Dime=Heads}\}$. Then $\mathbf{P}(A|B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \frac{(1/2)(1/2)}{(1/2)(1/2)+(1/2)(1)} = 1/3$. (b)

As a further check, $\mathbf{P}(A^C|B) = \frac{\mathbf{P}(A^C\cap B)}{\mathbf{P}(B)} = \frac{(1/2)(1)}{(1/2)(1/2)+(1/2)(1)} = 2/3$, which is consistent.

So, if Beauty sees that the Dime is Heads, then she should conclude that there is a 1/3 chance that the Nickel was Heads, and a 2/3 chance that the Nickel was Tails. I believe that this provides convincing support for the "thirder" position – though not everyone agrees.

Acknowledgements: I thank the Canadian Undergraduate Math Conference 2026 at McMaster University for inviting me to give a talk which led to this paper.