

# Conditional Probability - Problem Set

Declan Stacy

## 1 Question #1

Mario has two children. Assume that children are equally likely to be born as a boy or a girl and are equally likely to be born on any day of the week. I ask Mario if he has a daughter, and he says yes. What is the probability the other child is a son? What if instead I ask Jerry, who also has two kids, if he has a daughter that was born on a Tuesday, and he says yes; what is the probability the other child is a son in this scenario?

From the definition of conditional probability we can write,

$$P(\text{has son}|\text{has daughter}) = P(\text{has son} \cap \text{has daughter})/P(\text{has daughter})$$

$P(\text{has son} \cap \text{has daughter}) = \frac{1}{2}$  because he can either have two sons, two daughters, an older son and a younger daughter, or an older daughter and a younger son, all with equal probability. In two out of these four scenarios he has a son and a daughter, and  $\frac{2}{4} = \frac{1}{2}$ .

$P(\text{has daughter}) = \frac{3}{4}$  because in three out of these four scenarios he has at least one daughter.

So, the answer to the first question is  $\frac{1}{2} \div \frac{3}{4} = \boxed{\frac{2}{3}}$ .

For the second question, we are also concerned with what day of the week each child was born on, so there are  $(2 \cdot 7)^2$  scenarios (each child is either a boy or a girl, and is born on one of seven days), and all scenarios are equally likely.

We need to find the number of scenarios where Jerry has a daughter born on a Tuesday. This is  $14 + 14 - 1 = 27$  because there are 14 scenarios where the younger child is a daughter born on a Tuesday and 14 scenarios where the older child is a daughter born on a Tuesday because the other child can be of  $2 \cdot 7 = 14$  varieties. However, this counts the scenario where both children are daughters born on Tuesdays, so we must subtract one.

The number of scenarios where Jerry has a son and a daughter born on Tuesday is  $2 \cdot 7 = 14$  because the son can be the first or the second child and he can be born on any of the seven days of the week.

Remember that another interpretation of conditional probability is restricting the sample space; instead of having  $(2 \cdot 7)^2$  possible scenarios, with the information we are given we know that only 27 of these are possible, and are equally likely. Thus, the answer is  $\boxed{\frac{14}{27}}$  because out of the 27 scenarios, in 14 of them Jerry has a son.

(If you are confused, try the first method of using the definition of conditional probability and you will see that you get the same fraction but with the numerator and denominator divided by  $(2 \cdot 7)^2$ .)

## 2 Inference #1

**HMMT Feb 2019 Guts #13:** Reimu has 2019 coins  $C_0, C_1, \dots, C_{2018}$ , one of which is fake, though they look identical to each other (so each of them is equally likely to be fake). She has a machine that takes any two coins and picks one that is not fake. If both coins are not fake, the machine picks one uniformly at random. For each  $i = 1, 2, \dots, 1009$ , she puts  $C_0$  and  $C_i$  into the machine once, and the machine picks  $C_i$ . What is the probability that  $C_0$  is fake?

The first thing you should do when you see a probability problem is write out in words what you are trying to find. In this case, we want  $P(C_0 \text{ fake} \mid \text{never picked})$ . It can also be helpful to write this in a couple of different ways using the definition of conditional probability and Bayes' Theorem:

$$\begin{aligned} P(C_0 \text{ fake} \mid \text{never picked}) &= \frac{P(C_0 \text{ fake and never picked})}{P(C_0 \text{ never picked})} \\ &= P(C_0 \text{ never picked} \mid \text{fake}) \cdot \frac{P(C_0 \text{ fake})}{P(C_0 \text{ never picked})} \end{aligned}$$

In this case, the second option (applying Bayes' Theorem) looks like the easiest to compute.

1.  $P(C_0 \text{ never picked} \mid \text{fake}) = 1$

The question says that the machine will never pick the fake coin, so this is clearly 1.

2.  $P(C_0 \text{ fake}) = \frac{1}{2019}$

In the question it says that one of the 2019 coins is fake, and they are equally likely to be fake. So, the probability of an individual coin being fake is simply  $\frac{1}{2019}$ .

3.  $P(C_0 \text{ never picked}) = \frac{2^{1009} + 1009}{2019 \cdot 2^{1009}}$

Looking at the question, the answer to this is not explicitly given like with the other two probabilities. When you don't know what to do, casework is often a good option. So what should we do casework on? In this case, knowing which coin was fake would make computing probabilities a lot easier, so we should do casework on that. Let the fake coin be  $C_x$ . We will consider three cases:  $x = 0$  ( $C_0$  fake),  $x \in [1, 1009]$ , and  $x \in [1010, 2018]$  and use the Total Probability Theorem:

$$\begin{aligned} P(C_0 \text{ never picked}) &= P(C_0 \text{ never picked} \mid x = 0) \cdot P(x = 0) \\ &\quad + P(C_0 \text{ never picked} \mid x \in [1, 1009]) \cdot P(x \in [1, 1009]) \\ &\quad + P(C_0 \text{ never picked} \mid x \in [1010, 2018]) \cdot P(x \in [1010, 2018]) \end{aligned}$$

- (a) Case 1:  $x = 0$  ( $C_0$  fake)

We already know  $P(C_0 \text{ never picked} \mid \text{fake}) = 1$  and  $P(C_0 \text{ fake}) = \frac{1}{2019}$  from above, so this case is already done.

- (b) Case 2:  $x \in [1, 1009]$

In this case,  $C_0$  is a real coin and is put into the machine 1008 times with another real coin, and 1 time with a fake coin. When it is put in with the fake coin, the machine must pick  $C_0$  because the machine never picks the fake coin. Thus,  $P(C_0 \text{ never picked} \mid x \in [1, 1009]) = 0$  since  $C_0$  must be picked at least once.

- (c) Case 3:  $x \in [1010, 2018]$

In this case,  $C_0$  is a real coin and is put into the machine 1009 times with another real coin. Each time, there is a  $\frac{1}{2}$  chance that  $C_0$  will be picked (if both coins are real the machine picks one at random). The decision the machine makes each time is independent of the decisions it made before. The probability of multiple independent events occurring is the product of their individual probabilities, so we can write  $P(C_0 \text{ never picked} \mid x \in [1010, 2018]) = \frac{1}{2}^{1009}$ .

Now all we need to do is compute  $P(x \in [1010, 2018])$ . Since the events  $x = 1010, x = 1011, \dots, x = 2019$  are disjoint, we can write  $P(x \in [1010, 2018]) = P(x = 1010) + P(x = 1011) + \dots + P(x = 2019)$ . The probability of any individual coin being fake is  $\frac{1}{2019}$ , so this is simply  $1009 \cdot \frac{1}{2019} = \frac{1009}{2019}$ . (Quick tip: multiply when computing the probability of multiple independent events occurring, add when computing the probability of at least one of many disjoint events occurring.)

Now we can plug in all of the stuff we just computed into our equation for  $P(C_0 \text{ never picked})$  to get:

$$\begin{aligned} P(C_0 \text{ never picked}) &= 1 \cdot \frac{1}{2019} + 0 \cdot P(x \in [1, 1009]) + \frac{1}{2^{1009}} \cdot \frac{1009}{2019} \\ &= \frac{2^{1009} + 1009}{2019 \cdot 2^{1009}} \end{aligned}$$

Finally, we have all the parts to compute

$$\begin{aligned} &P(C_0 \text{ fake} \mid \text{never picked}) \\ &= P(C_0 \text{ never picked} \mid \text{fake}) \cdot \frac{P(C_0 \text{ fake})}{P(C_0 \text{ never picked})} \\ &= 1 \cdot \frac{\frac{1}{2019}}{\frac{2^{1009} + 1009}{2019 \cdot 2^{1009}}} \\ &= \boxed{\frac{2^{1009}}{2^{1009} + 1009}} \end{aligned}$$

This is a great example of how Bayes' Theorem and the Total Probability Theorem can be used to break up a problem into simple pieces that are easy to compute.

### 3 Prediction #1

**HMMT Feb 2018 Combo #5:** A bag contains nine blue marbles, ten ugly marbles, and one special marble. Ryan picks marbles randomly from this bag with replacement until he draws the special marble. He notices that none of the marbles he drew were ugly. Given this information, what is the expected value of the number of total marbles he drew?

Like the last problem, we should write out what we are looking for:

$E(\text{marbles} \mid \text{not ugly})$

When a problem involves a process that could go on forever, you will often have to compute an infinite sum. If we use the definitions of expected value and conditional probability, we can write:

$$\begin{aligned} E(\text{marbles} \mid \text{not ugly}) &= \sum_{i=1}^{\infty} i \cdot P(i \text{ marbles} \mid \text{not ugly}) \\ &= \sum_{i=1}^{\infty} i \cdot \frac{P(i \text{ marbles} \cap \text{not ugly})}{P(\text{not ugly})} \end{aligned}$$

$$1. P(i \text{ marbles} \cap \text{not ugly}) = \frac{9}{20}^{i-1} \cdot \frac{1}{20}$$

The probability of drawing a blue marble is  $\frac{9}{20}$ , and the probability of drawing the special marble is  $\frac{1}{20}$ . Also, each draw is independent of the last. Thus, the probability of drawing  $i$  marbles and not drawing any uglies is  $\frac{9}{20}^{i-1} \cdot \frac{1}{20}$  because we drew  $i - 1$  blue marbles followed by 1 special marble.

$$2. P(\text{not ugly}) = \frac{1}{11}$$

To compute  $P(\text{not ugly})$ , we will use the Total Probability Theorem:

$$P(\text{not ugly}) = \sum_{i=1}^{\infty} P(\text{not ugly} \mid i \text{ marbles}) \cdot P(i \text{ marbles})$$

This is essentially casework on the number of marbles that were drawn.

$P(\text{not ugly} \mid i \text{ marbles}) = \frac{9}{19}^{i-1}$  since the first  $i - 1$  marbles are either blue or ugly, and

$$\frac{P(\text{marble is ugly})}{P(\text{marble is ugly or blue})} = \frac{\frac{9}{20}}{\frac{19}{20}} = \frac{9}{19}.$$

$P(i \text{ marbles}) = \frac{19}{20}^{i-1} \cdot \frac{1}{20}$  since the first  $i - 1$  marbles must be ugly or red and the last marble must be special.

Thus,

$$\begin{aligned} P(\text{not ugly}) &= \sum_{i=1}^{\infty} \frac{9}{19}^{i-1} \cdot \frac{19}{20}^{i-1} \cdot \frac{1}{20} \\ &= \sum_{i=1}^{\infty} \frac{9}{20}^{i-1} \cdot \frac{1}{20} \\ &= \frac{\frac{1}{20}}{1 - \frac{9}{20}} \\ &= \frac{1}{11} \end{aligned}$$

where we recognized the sum as an infinite geometric series with a common ratio between  $-1$  and  $1$ .

Now we have all the parts to compute our original sum:

$$\begin{aligned} E(\text{marbles} \mid \text{not ugly}) &= \sum_{i=1}^{\infty} i \cdot \frac{\frac{9}{20}^{i-1} \cdot \frac{1}{20}}{\frac{1}{11}} \\ &= \frac{11}{20} \cdot \sum_{i=1}^{\infty} i \cdot \frac{9}{20}^{i-1} \\ &= \frac{11}{20} \cdot \frac{20}{9} \cdot \sum_{i=1}^{\infty} (i-1) \cdot \frac{9}{20}^{i-1} \end{aligned}$$

Multiplying the third line of the equation by  $\frac{9}{20}$  and subtracting it from the second line yields:

$$\begin{aligned} \left(1 - \frac{9}{20}\right)(\text{marbles} \mid \text{not ugly}) &= \frac{11}{20} \cdot \sum_{i=1}^{\infty} \frac{9}{20}^{i-1} \\ E(\text{marbles} \mid \text{not ugly}) &= \sum_{i=1}^{\infty} \frac{9}{20}^{i-1} \\ &= \frac{1}{1 - \frac{9}{20}} \\ &= \boxed{\frac{20}{11}} \end{aligned}$$

Overall, this question utilized the same concepts as the last question, except we had to do casework on infinitely many events and compute infinite sums.

## 4 Prediction #2

**PUMaC 2018 Live Round Calculus #1:** Noted magician Casimir the Conjurer has an infinite chest full of weighted coins. For each  $p \in [0, 1]$ , there is exactly one coin with probability  $p$  of turning up heads. Kapil the Kingly draws a coin at random from Casimir the Conjurer's chest, and flips it 10 times. To the amazement of both, the coin lands heads up each time! On his next flip, if the expected probability that Kapil the Kingly flips a head is written in simplest form as  $\frac{p}{q}$ , then compute  $p + q$ .  
(Note: Uses Calculus)

Sometimes when dealing with a question like this where there are infinitely many options, it is useful to solve the problem as if there were a finite number of choices. What I mean is, instead of Casimir picking a coin with a probability  $p \in [0, 1]$ , let's say he only has 5 options for the probability of the coin. For example, we could choose a  $p \in \{0, .2, .4, .6, .8\}$ . If you did not solve the actual problem, try solving that simpler problem (with  $p \in \{0, .2, .4, .6, .8\}$ ) and then come back and read this solution.

To make this more general, let there be  $n$  options for the probability of the coin, and let the set of options be  $S = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\}$ .

We want  $P(11\text{th heads} \mid 10 \text{ heads})$ . Using the Total Probability Theorem, we can write:

$$P(11\text{th heads} \mid 10 \text{ heads}) = \sum_{i=0}^{n-1} P(11\text{th heads} \mid 10 \text{ heads} \cap p = \frac{i}{n}) (p = \frac{i}{n} \mid 10 \text{ heads})$$

If we know  $p$ , we know the probability of flipping another heads, because  $p$  is the probability of the coin flipping heads. Thus,  $P(11\text{th heads} \mid 10 \text{ heads} \cap p = \frac{i}{n}) = \frac{i}{n}$ .

To evaluate  $P(p = \frac{i}{n} \mid 10 \text{ heads})$ , we can use Bayes' Theorem:

$$P(p = \frac{i}{n} \mid 10 \text{ heads}) = P(10 \text{ heads} \mid p = \frac{i}{n}) \cdot \frac{P(p = \frac{i}{n})}{P(10 \text{ heads})}$$

$P(10 \text{ heads} \mid p = \frac{i}{n}) = \frac{i}{n}^{10}$  because each flip is independent of the last and there are 10 of them.

$P(p = \frac{i}{n}) = \frac{1}{n}$  because we are picking  $p$  randomly from  $S$  and  $|S| = n$ .

For  $P(10 \text{ heads})$ , we will have to use the Total Probability Theorem again.

$$\begin{aligned} P(10 \text{ heads}) &= \sum_{i=0}^{n-1} P(10 \text{ heads} \mid p = \frac{i}{n}) (p = \frac{i}{n}) \\ &= \sum_{i=0}^{n-1} \frac{i^{10}}{n} \cdot \frac{1}{n} \end{aligned}$$

So, if we chose what  $n$  was, then we would have all the parts to solve the problem. Looking back at the way we defined  $S$ , as  $n$  becomes very large, the set  $S$  becomes more and more similar to  $[0, 1]$ . Thus, if we take the limit as  $n$  approaches  $\infty$ , we will get the answer.

Note that our expression for  $P(10 \text{ heads})$  is a Riemann Sum, so:

$$\begin{aligned} P(10 \text{ heads}) &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{i^{10}}{n} \cdot \frac{1}{n} \\ &= \int_0^1 x^{10} dx \\ &= \frac{1}{11} \end{aligned}$$

Now we can substitute in all of the stuff we have computed to get the answer:

$$\begin{aligned}
P(11\text{th heads} \mid 10 \text{ heads}) &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} P(11\text{th heads} \mid 10 \text{ heads} \cap p = \frac{i}{n}) (p = \frac{i}{n} \mid 10 \text{ heads}) \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{i}{n} \cdot \frac{i^{10}}{n^{10}} \cdot \frac{1}{n} \\
&= 11 \cdot \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{i^{11}}{n^{11}} \cdot \frac{1}{n} \\
&= 11 \cdot \int_0^1 x^{11} dx \\
&= \frac{11}{12}
\end{aligned}$$

The question asks for the sum of the numerator and the denominator of the fraction, so the answer is  $11 + 12 = \boxed{23}$ .

It makes sense that the answer should be close to 1 because if a coin lands heads 10 times in a row, it is more likely to be weighted heavily towards heads. So, as we observe more and more heads, our estimate of what the true bias of the coin is should go up. You can also think of it as if we took a bunch of coins, flipped them 10 times, and then threw out the ones that did not flip 10 heads. The average bias of the remaining coins would most likely be heavily in favor of heads.

You did not have to do all this defining our own set and doing Riemann Sums business if you saw immediately what the integrals were going to be. However, if you did not know what to do, then imagining the problem dealt with a finite set instead of an infinite set and thinking about how you would solve it can help.

## 5 Inference #2

Johnny has a deck of 100 cards, all of which are initially black. An integer  $n$  is picked at random between 0 and 100, inclusive, and Johnny paints  $n$  of the 100 cards red. Johnny shuffles the cards and starts drawing them from the deck. What is the least number of red cards Johnny has to draw before the probability that all the remaining cards are red is greater than .5?

If all the cards are red, that means that  $n = 100$ , so what we want is the least whole number  $k$  that satisfies the inequality  $P(n = 100 \mid \text{first } k \text{ are red}) > .5$ . To get an expression for  $P(n = 100 \mid \text{first } k \text{ are red})$ , let's start with Bayes' Theorem:

$$P(n = 100 \mid \text{first } k \text{ are red}) = P(\text{first } k \text{ are red} \mid n = 100) \cdot \frac{P(n = 100)}{P(\text{first } k \text{ are red})}$$

$P(\text{first } k \text{ are red} \mid n = 100) = 1$  because  $n = 100$  means that all the cards are red, so the first  $k$  will always be red.

$P(n = 100) = \frac{1}{101}$  since  $n$  is chosen uniformly at random from 101 options.

To calculate  $P(\text{first } k \text{ are red})$ , we will use the Total Probability Theorem:

$$P(\text{first } k \text{ are red}) = \sum_{i=k}^{100} P(n = i) P(\text{first } k \text{ are red} \mid n = i)$$

$P(n = i) = \frac{1}{101}$  (same explanation as for  $P(n = 100)$ ).

The tricky part is  $P(\text{first } k \text{ are red} \mid n = i)$ . What that is saying is that you draw  $k$  cards out of 100, and they are each one of the  $i$  red cards in the deck. There are  $\binom{100}{k}$  sets of cards that could be the top  $k$

cards. Out of those  $\binom{100}{k}$  sets,  $\binom{i}{k}$  of them are all red cards because there are  $i$  red cards to choose from. Thus,  $P(\text{first } k \text{ are red} | n = i) = \frac{\binom{i}{k}}{\binom{100}{k}}$

(Quick Tip: here we are considering all the cards to be unique because it makes the calculations easier, but some questions are more easily answered when you consider all objects of the same type to be the same. No matter which way you do it, remember to be consistent. Don't do one calculation where you consider all the red cards to be the same and then another where you consider them to be different.)

Now let's substitute that back into our expression for  $P(\text{first } k \text{ are red})$ :

$$\begin{aligned} P(\text{first } k \text{ are red}) &= \sum_{i=k}^{100} P(n = i) P(\text{first } k \text{ are red} | n = i) \\ &= \sum_{i=k}^{100} \frac{1}{101} \cdot \frac{\binom{i}{k}}{\binom{100}{k}} \\ &= \frac{1}{101} \cdot \frac{1}{\binom{100}{k}} \cdot \sum_{i=k}^{100} \binom{i}{k} \end{aligned}$$

Now we can use a theorem not mentioned in the article, the Baseball Theorem, which says that  $\sum_{i=a}^b \binom{i}{a} = \binom{b+1}{a+1}$ , to write:

$$\begin{aligned} P(\text{first } k \text{ are red}) &= \frac{1}{101} \cdot \frac{1}{\binom{100}{k}} \cdot \sum_{i=k}^{100} \binom{i}{k} \\ &= \frac{1}{101} \cdot \frac{\binom{101}{k+1}}{\binom{100}{k}} \\ &= \frac{1}{101} \cdot \frac{\frac{101!}{(k+1)!(100-k)!}}{\frac{100!}{k!(100-k)!}} \\ &= \frac{1}{101} \cdot \frac{101!!}{100! \cdot (k+1)!} \\ &= \frac{1}{101} \cdot \frac{101}{k+1} \\ &= \frac{1}{k+1} \end{aligned}$$

(Alternatively, you can notice that  $\sum_{i=k}^{100} \binom{i}{k}$  is the same as choosing  $k+1$  of the first 101 natural numbers, doing casework on the largest number you choose. For example, if the largest number you choose is 50, then there are  $\binom{49}{k}$  options for other  $k$  numbers since there are 49 natural numbers less than 50 we are choosing  $k+1$  numbers total. The largest number you choose can range from  $k+1$  to 101, so  $\binom{101}{k+1} = \sum_{i=k+1}^{101} \binom{i-1}{k} = \sum_{i=k}^{100} \binom{i}{k}$ .)

Finally, we can plug everything into our equation for  $P(n = 100 | \text{first } k \text{ are red})$  to obtain:

$$\begin{aligned} P(n = 100 | \text{first } k \text{ are red}) &= P(\text{first } k \text{ are red} | n = 100) \cdot \frac{P(n = 100)}{P(\text{first } k \text{ are red})} \\ &= 1 \cdot \frac{\frac{1}{101}}{\frac{1}{k+1}} \\ &= \frac{k+1}{101} \end{aligned}$$

The question asks for the least whole number  $k$  that satisfies the inequality  $P(n = 100 | \text{first } k \text{ are red}) = \frac{k+1}{101} > .5$ , so the answer is clearly  $\boxed{50}$ .

When I first solved the problem, I was surprised by this result; I thought the answer was going to be much larger. Part of this came from flawed logic that should be avoided when dealing with conditional

probability questions. For example, the argument "the answer is 99 because after drawing 99 red cards the remaining card can either be red or black, so there is a 50% chance of drawing a red card" could make sense initially, but it assumes that the last card is red or black with equal probability. In fact, there is a much higher chance of drawing red because if we started with 99 red cards and 1 black card, it would be very unlikely that the first 99 cards were all red (1% chance). There is an equal probability of starting with 99 red cards as starting with 100 red cards, but after observing the 99 red cards in a row we should come to the conclusion that 100 red cards is much more likely.

## 6 Inference #3

**HMMT Feb 2019 Guts #29:** Yannick picks a number  $N$  randomly from the set of positive integers such that the probability that  $n$  is selected is  $2^{-n}$  for each positive integer  $n$ . He then puts  $N$  identical slips of paper numbered 1 through  $N$  into a hat and gives the hat to Annie. Annie does not know the value of  $N$ , but she draws one of the slips uniformly at random and discovers that it is the number 2. What is the expected value of  $N$  given Annie's information? (Note: Uses Calculus)

We want to find  $E(N|\text{draws } 2)$ . Using the definition of expected value and Bayes' Theorem, we can write

$$\begin{aligned} E(N|\text{draws } 2) &= \sum_{i=2}^{\infty} i(N = i | \text{draws } 2) \\ &= \sum_{i=2}^{\infty} i(\text{draws } 2 | N = i) \cdot \frac{P(N = i)}{P(\text{draws } 2)} \end{aligned}$$

$P(\text{draws } 2 | N = i) = \frac{1}{i}$  since there are  $i$  slips and Annie chooses one of them at random (technically if  $i < 2$  than it would be 0 but in our sum  $i$  starts at 2).

$P(N = i) = 2^{-i}$  because that is given in the problem.

Can you guess what we are going to do to compute  $P(\text{draws } 2)$ ? Yes, for the fifth time we are going to use the Total Probability Theorem.

$$\begin{aligned} P(\text{draws } 2) &= \sum_{i=2}^{\infty} P(\text{draws } 2 | N = i)(N = i) \\ &= \sum_{i=2}^{\infty} \frac{1}{i} \cdot 2^{-i} \end{aligned}$$

This sum looks like a geometric series with a common ratio of  $\frac{1}{2}$ , but unfortunately we have that  $\frac{1}{i}$  messing it up. In order to get rid of the  $\frac{1}{i}$ , we need to do something that will result in each term being multiplied by  $i$ . It will be easier to see if we replace  $\frac{1}{2}$  with  $x$  and then plug in  $\frac{1}{2}$  for  $x$  later:

$$f(x) = \sum_{i=2}^{\infty} \frac{x^i}{i}$$

and  $P(\text{draws } 2) = f(\frac{1}{2})$ .

Now it is clear that we have to differentiate:



$$\begin{aligned}
f'(x) &= \sum_{i=2}^{\infty} \frac{i \cdot x^{i-1}}{i} \\
&= \sum_{i=2}^{\infty} x^{i-1} \\
&= \sum_{i=1}^{\infty} x^i \\
&= \frac{1}{1-x} - 1 \quad \text{provided that } |x| < 1, \text{ which it is.}
\end{aligned}$$

To solve for  $f(x)$ , integrate:

$$\begin{aligned}
\int f'(x) dx &= \int \left( \frac{1}{1-x} - 1 \right) dx \\
f(x) &= -\ln(1-x) - x + C
\end{aligned}$$

To solve for  $C$ , plug in 0 for  $x$ :

$$f(0) = -\ln(1-0) - 0 + C = \frac{1}{1-0} - 1 = 0$$

Thus,  $C = 0$ ,  $f(x) = -\ln(1-x) - x$ , and  $P(\text{draws } 2) = f(\frac{1}{2}) = \ln(2) - \frac{1}{2}$   
Now we can go back to our equation to compute  $E(N | \text{draws } 2)$

$$\begin{aligned}
E(N | \text{draws } 2) &= \sum_{i=2}^{\infty} i(\text{draws } 2 | N = i) \cdot \frac{P(N = i)}{P(\text{draws } 2)} \\
&= \sum_{i=2}^{\infty} i \cdot \frac{1}{i} \cdot \frac{2^{-i}}{\ln(2) - \frac{1}{2}} \\
&= \frac{1}{\ln(2) - \frac{1}{2}} \cdot \sum_{i=2}^{\infty} \frac{1}{2^i} \\
&= \frac{1}{\ln(2) - \frac{1}{2}} \cdot \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{2}} \quad (\text{sum of geometric series}) \\
&= \frac{1}{\ln(2) - \frac{1}{2}} \cdot \frac{1}{4} \cdot 2 \\
&= \boxed{\frac{1}{2 \ln(2) - 1}}
\end{aligned}$$