1 Boy or Girl Paradox

You know Mr. Smith has two children, at least one of whom is a boy. Assume that gender is independent and uniformly distributed, so for any child, the probability that they are a boy is the same as the probability they are a girl, which is $1/2$.

(a) What is the probability that both children are boys?

(b) Now suppose you knock on Mr. Smith's front door and you are greeted by a boy who you correctly deduce to be Mr. Smith's older child. What is the probability that he has two boys? Compare your answer to the answer in part (a).

Solution:

(a) Let B_1 be the event that the first child is a boy, and B_2 be the event that the second child is a boy. We are asked to find $\mathbb{P}[(B_1 \cap B_2) \mid (B_1 \cup B_2)]$:

$$\begin{aligned}
\mathbb{P}[(B_1 \cap B_2) \mid (B_1 \cup B_2)] &= \frac{\mathbb{P}[(B_1 \cap B_2) \cap (B_1 \cup B_2)]}{\mathbb{P}[B_1 \cup B_2]} \\
&= \frac{\mathbb{P}[B_1 \cap B_2]}{\mathbb{P}[B_1] + \mathbb{P}[B_2] - \mathbb{P}[B_1 \cap B_2]} \\
&= \frac{(1/2)(1/2)}{1/2 + 1/2 - 1/4} = \frac{1/4}{3/4} \\
&= \frac{1}{3}
\end{aligned}$$

Note: It is tempting to think that because the children's genders are independent, the probability of the second child being a boy given that the first is a boy is simply $1/2$. While this is true, when we write it out in terms of events, we can see that this is not the quantity that we want. See part (b) for more details.

(b) In this part, we want to find $\mathbb{P}[(B_1 \cap B_2) \mid B_1]$:

$$\begin{aligned}
\mathbb{P}[(B_1 \cap B_2) \mid B_1] &= \frac{\mathbb{P}[(B_1 \cap B_2) \cap B_1]}{\mathbb{P}[B_1]} \\
&= \frac{\mathbb{P}[B_1 \cap B_2]}{\mathbb{P}[B_1]} \qquad (1) \\
&= \frac{1/4}{1/2} = \frac{1}{2}
\end{aligned}$$

Note the distinction between this part and part (a), and that a common mistake in determining the answer to part (a) is solving part (b) instead.

2 Box of Marbles

You are given two boxes: one of them containing 900 red marbles and 100 blue marbles, the other one contains 500 red marbles and 500 blue marbles.

(a) If we pick one of the boxes randomly, and pick a marble what is the probability that it is blue?

(b) If we see that the marble is blue, what is the probability that it is chosen from box 1?

(c) Suppose we pick one marble from box 1 and without looking at its color we put it aside. Then we pick another marble from box 1. What is the probability that the second marble is blue?

Solution:

(a) Let B be the event that the picked marble is blue, R be the event that it is red, A_1 be the event that the marble is picked from box 1, and A_2 be the event that the marble is picked from box 2. Therefore we want to calculate $\mathbb{P}(B)$. By total probability,

$$\mathbb{P}(B) = \mathbb{P}(B \mid A_1)P(A_1) + \mathbb{P}(B \mid A_2)\mathbb{P}(A_2) = 0.5 \times 0.1 + 0.5 \times 0.5 = 0.3.$$

(b) In this part, we want to find $\mathbb{P}(A_1 \mid B)$. By Bayes rule,

$$\mathbb{P}(A_1 \mid B) = \frac{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1)}{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(B \mid A_2)\mathbb{P}(A_2)} = \frac{0.1 \times 0.5}{0.5 \times 0.1 + 0.5 \times 0.5} = \frac{1}{6}.$$

(c) Let B_1 be the event that first marble is blue, R_1 be the event that the first marble is red, and B_2 be the event that second marble is blue without looking at the color of first marble. We want to find $\mathbb{P}(B_2)$. By total probability,

$$\mathbb{P}(B_2) = \mathbb{P}(B_2 \mid B_1)\mathbb{P}(B_1) + \mathbb{P}(B_2 \mid R_1)\mathbb{P}(R_1) = \frac{99}{999} \times 0.1 + \frac{100}{999} \times 0.9 = 0.1.$$

More generally, one can see that the probability that the n-th marble picked from box 1 is blue with probability 0.1. This is clear by symmetry: all the permutations of the 1000 marbles have the same probability, so the probability that the n-th marble is blue is the same as the probability that the first marble is blue.

3 Lie Detector

A lie detector is known to be $4/5$ reliable when the person is guilty and $9/10$ reliable when the person is innocent. If a suspect is chosen from a group of suspects of which only $1/100$ have ever committed a crime, and the test indicates that the person is guilty, what is the probability that he is innocent?

Solution:

Let A denote the event that the test indicates that the person is guilty, and B the event that the person is innocent. Note that

$$\mathbb{P}[B] = \frac{99}{100}, \quad \mathbb{P}[\overline{B}] = \frac{1}{100}, \quad \mathbb{P}[A \mid B] = \frac{1}{10}, \quad \mathbb{P}[A \mid \overline{B}] = \frac{4}{5}.$$

Using the Bayesian Inference Rule, we can compute the desired probability as follows:

$$\mathbb{P}[B \mid A] = \frac{\mathbb{P}[B]\mathbb{P}[A \mid B]}{\mathbb{P}[B]\mathbb{P}[A \mid B] + \mathbb{P}[\overline{B}]\mathbb{P}[A \mid \overline{B}]} = \frac{(99/100)(1/10)}{(99/100)(1/10) + (1/100)(4/5)} = \frac{99}{107}$$

4 Bayes Rule – Man Speaks Truth

(a) A man speaks the truth 3 out of 4 times. He flips a biased coin that comes up Heads $1/3$ of the time and reports that it is Heads. What is the probability it is Heads?

(b) A man speaks the truth 3 out of 4 times. He rolls a fair 6-sided die. When you ask him if the die came up with a 6, he answers "yes". What is the probability it is really 6?

Solution:

(a) Let E denote the event the man reports heads, S_1 be the event that the coin comes up heads, and S_2 be the event that the coin comes up tails.

We have:

$$\mathbb{P}(E \mid S_1) = \frac{3}{4}, \quad \mathbb{P}(E \mid S_2) = \frac{1}{4}, \quad \mathbb{P}(S_1) = \frac{1}{3}, \quad \mathbb{P}(S_2) = \frac{2}{3}.$$

We want to compute $\mathbb{P}(S_1 \mid E)$, and let's do so by applying Bayes Rule.

$$\mathbb{P}(S_1 \mid E) = \frac{\mathbb{P}(S_1 \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E \mid S_1)\mathbb{P}(S_1)}{\mathbb{P}(E \mid S_1)\mathbb{P}(S_1) + \mathbb{P}(E \mid S_2)\mathbb{P}(S_2)} = \frac{(3/4)\cdot(1/3)}{(3/4)\cdot(1/3) + (1/4)\cdot(2/3)}$$
$$= \frac{3}{5}.$$

(b) Let E denote the event that the man says "yes" to your question, S_1 be the event that the dice comes up 6, and S_2 be the event that the dice comes up not 6.

We have:

$$\mathbb{P}(E \mid S_1) = \frac{3}{4}, \quad \mathbb{P}(E \mid S_2) = \frac{1}{4}, \quad \mathbb{P}(S_1) = \frac{1}{6}, \quad \mathbb{P}(S_2) = \frac{5}{6}.$$

We want to compute $\mathbb{P}(S_1 \mid E)$, and let's do so by applying Bayes Rule.

$$\begin{aligned}\mathbb{P}(S_1 \mid E) &= \frac{\mathbb{P}(S_1 \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E \mid S_1)\mathbb{P}(S_1)}{\mathbb{P}(E \mid S_1)\mathbb{P}(S_1) + \mathbb{P}(E \mid S_2)\mathbb{P}(S_2)} = \frac{(3/4)\cdot(1/6)}{(3/4)\cdot(1/6)+(1/4)\cdot(5/6)} \\ &= \frac{3}{8}.\end{aligned}$$

5 Weathermen

Tom is a weatherman in New York. On days when it snows, Tom correctly predicts the snow 70% of the time. When it doesn't snow, he correctly predicts no snow 95% of the time. In New York, it snows on 10% of all days.

(a) If Tom says that it is going to snow, what is the probability it will actually snow?

(b) What is Tom's overall accuracy?

(c) Tom's friend Jerry is a weatherman in Alaska. Jerry claims that she is a better weatherman than Tom even though her overall accuracy is lower. After looking at their records, you determine that Jerry is indeed better than Tom at predicting snow on snowy days and sun on sunny day. Give an instance of the situation described above. *Hint: what is the weather like in Alaska?*

Solution:

(a) Let S be the event that it snows and T be the event that Tom predicts snow.

$$\begin{aligned}P(S|T) &= \frac{P(S\cap T)}{P(T)} \\ &= \frac{P(S\cap T)}{P(S\cap T)+P(\overline{S}\cap T)} \\ &= \frac{.1\times .7}{.1\times .7 + .9\times .05}\end{aligned}$$

(b)

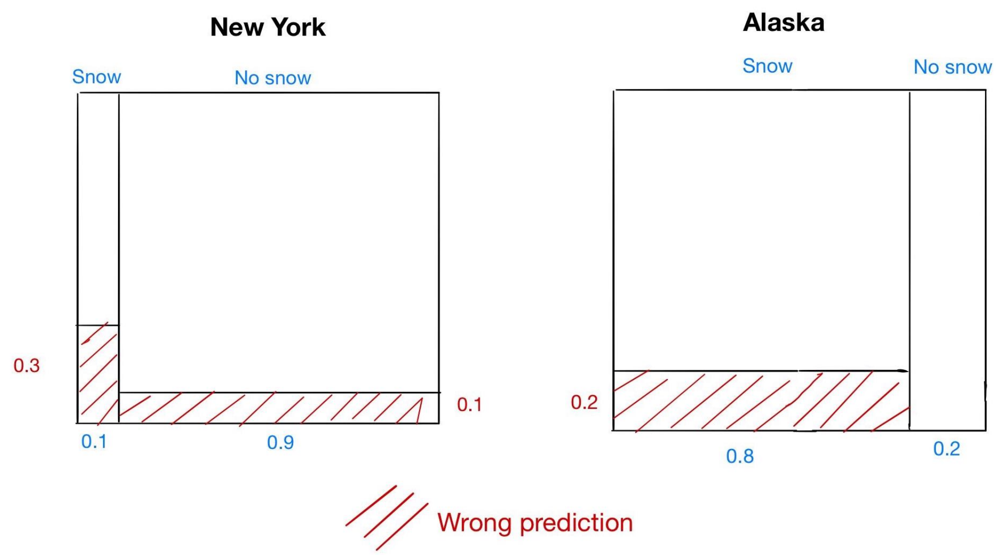
$$\begin{aligned}P(\text{Tom is correct}) &= P(S\cap T) + P(\overline{S}\cap\overline{T}) \\ &= .1\times .7 + .9\times .95\end{aligned}$$

(c) Even though Jerry's overall accuracy is lower, it is still possible that she is a better weatherman if the weather is different.

For example, let's assume that it snows 80% of days in Alaska.

- When it snows, Jerry correctly predicts snow 80% of the time.
- When it doesn't snow, Jerry correctly predicts no snow 100% of the time.

Jerry's overall accuracy turns out to be less than Tom's even though she is better at predicting both categories! The following diagram gives an illustration of the situation. The intuition is that Jerry's error gets penalized more heavily than Tom because it snows more often in Alaska.



For more info on this kind of phenomena, check out Simpson's Paradox!