1.6 The Law of Total Probability

In the previous section we considered an urn and looked at the probability that both balls we picked were red. Furthermore we may ask how to find $\mathbb{P}(R2)$?, that is, the probability that the second ball drawn is red. In this situation we need to take into account the different possibilities that occur within the first draw, and show how this affects the outcome of the second. To make this more precise we consider the following lemma:

Lemma 1.6.1 (Law of Total Probability). Suppose A and B are any events in a sample space. Then we have the following:

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c).$$

Essentially this lemma states that the probability an event occurs can be broken up into smaller conditional events. In the case of urn in Example 1.5.1, the lemma says that the probability that the second ball is red can happen in two distinct ways. Either the first ball is red ($R1$) and the second ball is red, or the first ball is blue ($R1^c$) and the second ball is red. In each case, we can compute the conditional probability of the second ball being red given that either $R1$ or $R1^c$ has occurred. The Law of Total Probability tells us how to combine these two scenarios together to get the total probability of $R2$ occurring.

Example 1.6.1. Consider the urn scenario in Example 1.5.3, what is the probability that the second ball drawn is red?

We recall that $R1$ is the event that the first ball drawn is red, and that $R2$ is the event that the second ball drawn is red. Then the Law of Total Probability states that,

$$\mathbb{P}(R2) = \mathbb{P}(R2|R1)\mathbb{P}(R1) + \mathbb{P}(R2|R1^c)\mathbb{P}(R1^c).$$

We have already computed the terms involving $R1$ in the previous example. So we first consider $\mathbb{P}(R1^c)$. This is precisely $1-\mathbb{P}(R1)$, by definition of complements, therefore

$$\mathbb{P}(R1^c) = 1 - 7/12 = 5/12.$$

We now consider $\mathbb{P}(R2|R1^c)$. In this setting we have that the first ball picked was blue. Therefore when we pick the second ball, we have that urn contains 7 red balls and 6 blue balls. Therefore the probability that the second ball picked is red is,

$$\frac{7}{7+4} = \frac{7}{11}.$$

Hence combining the above, we have that:

$$\mathbb{P}(R2) = \frac{42}{132} + \frac{5}{12}\cdot\frac{7}{11} = \frac{7}{12}.$$

We look at one last slightly more involved example, which applies all of the ideas that we have seen previously. We provide two solutions, it is important to understand how both of these approaches work.

Example 1.6.2. Suppose a raffle takes place. The participants involve a pair of identical twins called Sam and Dan, along with 18 other people. Exactly two people are chosen at random to receive a prize, each person being equally likely to be picked. What is the probability that at least one of the twins receives a prize?

Firstly we observe that this problem is not too dissimilar to the urn example stated above. We can think of the raffle running in two rounds. It first picks a person at random from the twenty participants as the first winner. In the second round the winner is removed, and the lottery selects a second winner from the remaining 19 participants. You should think why this equivalent to picking a pair of people at random.

We let T_1 be the event that one of the twins is picked in round one, while we let T_2 be the event that one of the twins is picked in round two. Therefore we are looking to compute $\mathbb{P}(T_1 \cup T_2)$.

Solution 1

Rather than directly computing $\mathbb{P}(T_1 \cup T_2)$ we look at the complement $\mathbb{P}((T_1 \cup T_2)^c)$. By definition of complements we have that:

$$\mathbb{P}(T_1 \cup T_2) + \mathbb{P}((T_1 \cup T_2)^c) = 1.$$

As $T_1 \cup T_2$ describes the event that at least one twin is picked, the complement event must be the case where neither twin is chosen. Therefore neither twin is chosen in round one or chosen in round two, this event is precisely ${T_1}^c \cap {T_2}^c$, (this is precisely an application of De Morgan's Law from day two of the exercises). Hence it follows that:

$$\mathbb{P}((T_1 \cup T_2)^c) = \mathbb{P}({T_1}^c \cap {T_2}^c).$$

We now apply the definition of conditional probability, hence we have that:

$$\mathbb{P}({T_1}^c \cap {T_2}^c) = \mathbb{P}({T_2}^c | {T_1}^c)\mathbb{P}({T_1}^c).$$

We now work term by term. For $\mathbb{P}({T_1}^c)$ there are a total of 20 participants, and two of them are the twins. Therefore the probability a twin is not chosen is $18/20 = 9/10$. For $\mathbb{P}({T_2}^c | {T_1}^c)$ we note a twin was not chosen in the first round, so there are 19 participants in the second round and two of them are twins. Therefore we have that $\mathbb{P}({T_2}^c | {T_1}^c) = 17/19$. Putting this all together we have that:

$$\mathbb{P}(T_1 \cup T_2) = 1 - \frac{17}{19} \times \frac{9}{10} = \frac{37}{190}.$$

Solution 2

Firstly we note that T_1 and T_2 are **not** mutually exclusive. If one of the twins is selected in round one, and the other in round two then T_1 and T_2 both occur. Therefore we apply the inclusion-exclusion principle:

$$\mathbb{P}(T_1 \cup T_2) = \mathbb{P}(T_1) + \mathbb{P}(T_2) - \mathbb{P}(T_1 \cap T_2). \tag{1.1}$$

We now look at $\mathbb{P}(T_2)$, we note that the chance the events occurs depends on whether a twin is picked in round one, this is same a reasoning as how to the second ball is picked in the urn example. Therefore, we apply the law of total probability:

$$\begin{aligned}\mathbb{P}(T_2) &= \mathbb{P}(T_2|T_1)\mathbb{P}(T_1) + \mathbb{P}(T_2|T_1{}^c)\mathbb{P}(T_1{}^c)\\ &= \mathbb{P}(T_2 \cap T_1) + \mathbb{P}(T_2|T_1{}^c)\mathbb{P}(T_1{}^c)\end{aligned}$$

The second equality follows from applying the definition of conditional probability, $\mathbb{P}(T_2|T_1)\mathbb{P}(T_1) = \mathbb{P}(T_2 \cap T_1)$. By substituting into Equation 1.1 and cancelling like terms we have:

$$\begin{aligned}\mathbb{P}(T_1 \cup T_2) &= \mathbb{P}(T_1) + \cancel{\mathbb{P}(T_2 \cap T_1)} + \mathbb{P}(T_2|T_1{}^c)\mathbb{P}(T_1{}^c) - \cancel{\mathbb{P}(T_1 \cap T_2)}\\ &= \mathbb{P}(T_1) + \mathbb{P}(T_2|T_1{}^c)\mathbb{P}(T_1{}^c).\end{aligned}$$

We now work term by term, as we have in previous examples. For $\mathbb{P}(T_1)$ we note that are 20 total are participants and 2 of them are the twins. Therefore $\mathbb{P}(T_1) = 2/20 = 1/10$. Consequently this implies that $\mathbb{P}(T_1{}^c) = 9/10$. We now consider $\mathbb{P}(T_2|T_1{}^c)$. In this setting we assume that $T_1{}^c$ has occurred, therefore one of the twins was not chosen in the first round. Hence in the second round there is a total of 19 participants, of these, there are Sam and Dan. Therefore we have that $\mathbb{P}(T_2|T_1{}^c) = 2/19$. Thus putting this together, we have that:

$$\mathbb{P}(T_1 \cup T_2) = \frac{1}{10} + \frac{2}{19} \times \frac{9}{10} = \frac{37}{190}.$$