

## Section #1 Solutions

---

### 1 Generative Processes: The Birthday Problem

**Preamble:** When solving a counting problem, it can often be useful to come up with a generative process, a series of steps that “generates” examples. A correct generative process to count the elements of set  $A$  will (1) *generate every element of  $A$*  and (2) *not generate any element of  $A$  more than once*. If our process has the added property that (3) *any given step always has the same number of possible outcomes*, then we can use the product rule of counting.

**Example:** Say we want to count the number of ways to roll two (distinct) dice where one die is even and one die is odd. Our process could be: (1) choose a number for the first die, (2) choose a number of opposite parity for the second die. Since the first step has 6 options and the second step has 3 options regardless of the outcome of the first step, the number of possibilities is  $6 * 3 = 18$ .

**Problem:** Assume that birthdays happen on any of the 365 days of the year with equal likelihood (we’ll ignore leap years).

- a. What is the probability that of the  $n$  people in your section, at least two people share the same birthday?

It is much easier to calculate the probability that no one shares a birthday. Let our sample space,  $S$  be the set of all possible assignments of birthdays to the students in section. By the assumptions of this problem, each of those assignments is equally likely, so this is a good choice of sample space. We can use the product rule of counting to calculate  $|S|$ :

$$|S| = (365)^n$$

Our event  $E$  will be the set of assignments in which there are no matches (i.e. everyone has a different birthday). We can think of this as a generative process where there are 365 choices of birthdays for the first student, 364 for the second (since it can’t be the same birthday as the first student), and so on. Verify for yourself that this process satisfies the three conditions listed above. We can then use the product rule of counting:

$$|E| = (365) \cdot (364) \cdot \dots \cdot (365 - n + 1)$$

$$\begin{aligned} P(\text{birthday match}) &= 1 - P(\text{no matches}) \\ &= 1 - \frac{|E|}{|S|} \\ &= 1 - \frac{(365) \cdot (364) \dots (365 - n + 1)}{(365)^n} \end{aligned}$$

An common misconception is that the size of the event  $E$  can be computed as  $|E| = \binom{365}{n}$  by choosing  $n$  distinct birthdates from 365 options. However, outcomes in this event ( $n$  unordered distinct dates) cannot recreate any outcomes in the sample space  $|S| = 365^n$  ( $n$  distinct dates, one for each distinct person). However if we compute the size of event  $E$  as  $|E| = \binom{365}{n}n!$  (equivalent to the number above), then we can assign the  $n$  birthdays to each person in a way consistent with the sample space.

Interesting values: ( $n = 13 : p \approx 0.19$ ), ( $n = 23 : p \approx 0.5$ ), ( $n = 70 : p \geq 0.99$ )

- b. What is the probability that your section contains exactly one pair of people who share a birthday?

We can use the same sample space, but our event is a little bit trickier. Now  $E$  is the set of birthday assignments in which exactly two students share a birthday and the rest have different birthdays. One generative process that works for this is (1) choose the two students who share a birthday, (2) choose  $n - 1$  birthdays in the same manner as in part a (i.e. one for the pair of students and one for each of the remaining students). We then have:

$$P(\text{exactly one match}) = \frac{|E|}{|S|} = \frac{\binom{n}{2}(365) \cdot (364) \cdot \dots \cdot (365 - n + 2)}{(365)^n}$$

Many other generative processes work for this problem. Try to think of some other ones and make sure you get the same answer!

## 2 Permutations and Combinations: Flipping Coins

**Preamble:** One thing that students often find tricky when learning combinatorics is how to figure out when a problem involves permutations and when it involves combinations. Naturally, we will look at a problem that can be solved with both approaches. Pay attention to what parts of your solution represent distinct objects and what parts represent indistinct objects.

**Problem:** We flip a fair coin  $n$  times, hoping (for some reason) to get  $k$  heads.

- a. How many ways are there to get exactly  $k$  heads? Characterize your answer as a *permutation* of H's and T's.

We want to know the number of sequences of  $n$  H's and T's such that there are  $k$  H's and  $n - k$  T's. This is the same as permuting  $n$  objects of which one set of  $k$  is indistinguishable and one set of  $n - k$  is indistinguishable. Using our formula for the permutation of indistinguishable objects, we get  $\frac{n!}{k!(n-k)!}$

- b. For what  $x$  and  $y$  is your answer to part a equal to  $\binom{x}{y}$ ? Why does this *combination* make sense as an answer?

Our answer to part a is equal to  $\binom{n}{k}$ . This makes sense because we can come up with a valid sequence by *choosing*  $k$  flips to come out to heads (and implicitly define the other  $n - k$  to be tails). The answer is also equivalent to  $\binom{n}{n-k}$  for which the same logic applies except with choosing flips to be tails.

- c. What is the probability that we get exactly  $k$  heads?

If we define our sample space to be all possible sequences of flips, then our event is the number of sequences where we get exactly  $k$  heads, meaning that  $|E|$  is (conveniently) the answer to the previous two parts. Our probability is then  $\frac{|E|}{|S|} = \frac{\binom{n}{k}}{2^n}$ .

### 3 Bayes Rule: Song Identification

**Preamble:** In this class, seeing a problem written in English can often throw you off of its scent. In this problem, we will practice translating a problem from English to equations and then applying Bayes Rule, which you learned this week.

**Problem:** Shazam is an application which can predict what song is playing. Based on the frequency of requests it's been getting these days, Shazam has found that:

- 80% of songs are Hold Up by Beyonce
- 20% of songs are Can't Get Used to Losing You by Andy Williams

When a request is made, Shazam receives an audio sample that it uses to update its belief. From one particular audio sample,  $S$ , Shazam estimates that:

- $S$  would have a 50% chance of appearing if Hold Up were playing.
- $S$  would have a 90% chance of appearing if Can't Get Used to Losing You were playing.

What is the updated probability that the song is Hold Up given the audio sample heard? HINT: Define variables and write all of the information we have given to you in terms of those variables.

Let  $X_1$  be the event that the song is Hold Up and let  $X_2$  be the event that the song is Can't Get Used to Losing You. We can write the information from the problem as:  $P(X_1) = 0.8$ ,  $P(X_2) = 0.2$ ,  $P(S|X_1) = 0.5$ ,  $P(S|X_2) = 0.9$ . We are looking for  $P(X_1|S)$ , which we can tackle with Bayes theorem:

$$P(X_1|S) = \frac{P(S|X_1)P(X_1)}{P(S)}$$

Since there are only two songs, we can expand the denominator using the law of total probability:

$$\begin{aligned} P(X_1|S) &= \frac{P(S|X_1)P(X_1)}{P(S|X_1)P(X_1) + P(S|X_2)P(X_2)} \\ &= \frac{0.50 \cdot 0.80}{0.50 \cdot 0.80 + 0.90 \cdot 0.20} \\ &\approx 0.69 \end{aligned}$$

#### 4 Probability Misunderstood: The Sally Clark Case

**Preamble:** Conditional probabilities are hard to interpret, especially if they are extremely close to zero or one. You should be careful about how you convey meaning to your audience, whether they are on a jury (as below) or are users of software that you have written.

**Problem:** Sally Clark was a British lawyer who was wrongly sentenced to life in prison in 1999 for the deaths of her two infant children. Her elder son Christopher died at age 11 weeks in December 1996 and her younger son Harry at 8 weeks in January 1998. At her trial, the defence argued that the deaths were due to sudden infant death syndrome (SIDS). Clark was convicted on the basis of testimony by pediatrician Sir Roy Meadow, who made the following argument:

- Hospital records show that the ratio of SIDS deaths to live births in affluent non-smoking families is about  $\frac{1}{8500}$ . (A live birth is a birth in which a child is born alive; not a still birth.)
- The chance of two SIDS deaths occurring in the same family is about  $\frac{1}{8500^2} \approx \frac{1}{73000000}$ .
- It is therefore extremely unlikely that Clark is innocent.

As a result of this prosecution, Clark spent more than 3 years in prison and was finally exonerated in 2003 after it was determined that Meadow's expert testimony was flawed. Two other women against whom Meadow provided expert testimony had their convictions overturned as well.

- a. Identify a flaw in Meadow's  $\frac{1}{73000000}$  figure.

The probability of two SIDS deaths occurring in the same family is actually likely to be much higher than  $\frac{1}{73000000}$ . The flaw in Meadow's reasoning is that SIDS deaths within the same family are unlikely to occur as independent events. Today, we believe that there are underlying genetic and environmental risk factors that predispose certain families to SIDS, making a second SIDS death more likely for those families.

Let's understand these ideas in the notation of probability. Consider an affluent non-smoking family with two children. Let  $C_1$  and  $C_2$  be the events that each child dies of SIDS, respectively. According to Meadow's data,

$$P(C_1) = P(C_2) \approx \frac{1}{8500}.$$

But SIDS deaths within a family are not independent, so the probability of both children dying of SIDS is not the product of the probabilities of each of them dying of SIDS:

$$P(C_1 \cap C_2) \neq P(C_1)P(C_2) \approx \frac{1}{73000000}.$$

Since there are underlying factors that elevate the risk for certain families, it is likely that

$$P(C_1 \cap C_2) > \frac{1}{73000000}.$$

We can restate this comparison using conditional probability:

$$P(C_2|C_1) = \frac{P(C_1 \cap C_2)}{P(C_1)} > \frac{1/73000000}{1/8500} \approx \frac{1}{8500}.$$

In words, the probability that the second child dies of SIDS given that the first child has died of SIDS is greater than  $\frac{1}{8500}$ .

- b. Even if we accept Meadow's  $\frac{1}{73000000}$  calculation as valid, what is wrong with a juror interpreting it as the probability of Clark's innocence?

Let's be generous to Meadow and accept his claim that the probability of two SIDS deaths in the same family is approximately  $\frac{1}{73000000}$ . It is still wrong for the jury to interpret this value as the probability that Clark is innocent. Let's understand the difference by defining some events.

Let  $D$  be the event that two children in family die (of any cause). Let  $I$  be the event that their mother is innocent of murder. According to Meadow, the probability of the two deaths given that the mother is innocent is miniscule:

$$P(D|I) \approx \frac{1}{73000000}.$$

But what Clark's jury should consider instead is  $P(I|D)$ , the probability that the mother is innocent given that her children have died. These two quantities are related by Bayes rule:

$$P(I|D) = \frac{P(D|I)P(I)}{P(D)}.$$

Notice that  $P(I)$ , the prior probability of the mother's innocence, should be very close to 1; the vast majority of mothers do not murder their children. Likewise,  $P(D)$ , the prior probability of two children in a family dying, should be very close to 0. Therefore,

$$P(I|D) \gg P(D|I).$$

In the context of criminal trials, mixing up conditional probabilities in this way is known as the prosecutor's fallacy. The jury should have disregarded Meadow's argument. It says almost nothing about Clark's innocence.