Worksheet 03 (Solutions)

1. Assume that you have a cup with m black marbles and m white marbles. Consider selecting a marble, putting it back, and selecting another marble. What is the probability that both marbles are the same color? The answer might be obvious from last class. Think about treating this as a proof that proves your intuition based only on the naïve definition of probability and basic rule of counting, rather than focusing on the numerical answer.

Solution: There are $2m$ possible outcomes when selecting one marble. Using the basic rule of counting, then, there are $2m \times 2m = 4m^2$ total ways of selecting two marbles with replacement. For the marbles to match, we can either select two white marbles or two black marbles. Consider selecting two white marbles: there are m possible selections for the first and m possible selections for the second, so $m \times m = m^2$ total choices. The same holds for the black balls, so there are $m^2 + m^2 = 2m^2$ ways of having matching marbles. This gives a probability of:

$$\frac{2m^2}{4m^2} = \frac{1}{2} = 0.5.$$

2. Assume that you have a cup with m black marbles and m white marbles for some $m > 1$. Consider selecting one marble and then another marble, without putting the first one back. What is the probability that the two marbles are the same color? Again, treat this as a proof rather than a calculation. What happens when m becomes large?

Solution: The total number of possible selections is a two-stage experiment, with $2m$ options in the first round and $(2m - 1)$ options in the second round. So, there are a total of $2m \cdot (2m - 1)$ total possible ways to select two marbles. For the marbles to match, we can start with any marble we want $(2m)$, but in the second round can only select the remaining $m - 1$ marbles of the same color. So the probability is:

$$\frac{2m \cdot (m-1)}{2m \cdot (2m-1)} = \frac{m-1}{2m-1}.$$

As $m \to \infty$, we have $m - 1 \approx m$ and $2m - 1 \approx 2m$, so that the ratio becomes $\frac{m}{2m} = \frac{1}{2}$. Formally, we can use L'Hôpital's rule: the top and bottom both limit to $+\infty$ as $m \to \infty$. The derivative of the top is 1 and of the bottom is 2.

The intuition that you should be building is that it makes sense that the difference between replacing the first marble or not decreases (to 0 in the limit) as the total number of marbles increases.

3. Ignore the issue with leap years and assume that birthdays are evently distributed throughout the year. What is the probability that 23 randomly selected people will have no shared birthdays? Do not try to simplify the result yet; just write it in terms of factorials, powers, and products.[1]

[1] This question is often called the Birthday Problem, the first of many famous probability questions we will study this semester.

Solution: The tricky part here is seeing that the two thing we need to count come straight from the worksheet. We can view someone's birthday as being a selection from a set of size 365. We are assigning birthdays to 23 people, this is an ordered selection done with replacement, since different people could share the same birthday. The event we want to compute the probability of is when nobody shares a birthday. How many ways are there to do this? It is just the number of ways to selecting 23 things from a set of size 365 without replacement. So:

$$\begin{aligned}\mathbb{P}(\text{no match}) &= \frac{(365)!/(365-23)!}{365^{23}} \\ &= \frac{(365)!}{365^{23}\cdot 342!}\end{aligned}$$

Plugging this directly into a calculator will almost certainly not work as the number (365)! is far too large to compute directly. The number 365^{23} (just) works in most programming languages, but will almost certainly exceed the sizes for a standard calculator as well.

4. Take the logarithm of the previous result and simplify as much as possible, writing the result in terms of $lf(\cdot)$, the logarithm of the factorial function. Most programming languages have an quick function to compute the log factorial. For example, $lf(365) = 1792.33$ and $lf(342) = 1657.34$. Using your result, now calculate the decimal version of the result from the previous question. Does the result seem surprising to you?

Solution: Using the rules of logarithms, we have:

$$\begin{aligned}\log\left(\frac{(365)!}{365^{23}\cdot 342!}\right) &= \log(365!) - \log(365^{23}) - \log(342!) \\ &= lf(365) - 23\cdot\log(365) - lf(342)\end{aligned}$$

Plugging in the values, we have:

$$\begin{aligned}lf(365) - 23\cdot\log(365) - lf(342) &= 1792.33 - 23\cdot 5.899897 - 1657.34 \\ &= -0.708\end{aligned}$$

Remember that this is the log probability. To get the actual probability, we take the value as a power of e and get $e^{-0.708} = 0.492 \approx 0.50$. So, there is about a 50/50 chance that a set of 23 people will not share a brithday. And, by symmetry, a 50/50 chance that someone will share a birthday.

The proportion here may seem surprisingly high since 23 is many times less than 365. The next question helps explain what is going on a bit better.

5. How many ways can you pick a pair of people from a set of 23 total people? ($\star$) How does this help explain the solution to the previous question?

Solution: If we cared about ordering, there are just $23 \cdot 22 = 506$. However, this double counts the pairs, since it would count the pairing of person 1 and person 2 as distinct from the pairing of person 2 and person 1. Dividing by 2 gives the number of pairs: $506/2 = 253$.

We could view the Birthday Problem in terms of considering pairs of individuals. While you might have thought we would need a group of around $365/2 = 182.5$ people to have a probability of 0.5 of sharing birthday, it is really more accurate to count the pairs of students, which is what we are considering in the problem. Since 253 is is greater than 182.5 if anything we might actually expect the probability to be a bit larger than it is. To understand why it is not around $253/365 \approx 0.7$, we will need much more probability theory than we have so far produced. We will get there soon!

6. Three students get on a bus to downtown at the same time. The bus makes three stops once it arrives in downtown Richmond. If each student randomly decides which stop to disembark, what is the probability that everyone gets off at the same stop?

Solution: Using our naïve counting definition we get:

$$\begin{aligned}\mathbb{P}(event) &= \frac{\#\{\text{ways get off same stop}\}}{\#\{\text{ways of getting off bus}\}} \\ &= \frac{3}{3^3} \\ &\approx 0.111.\end{aligned}$$

7. It is a well known result from calculus that the limit of $(1 - 1/n)^n$ as n goes to infinity is e^{-1}. Consider a set of N^2 items for which there are N black items and $N^2 - N$ red items. If we sample N items with replacement from this set, what is the probability that all of them are are red? Assuming that N is sufficently large, write this in terms of e using the formula given above.

Now, take a deep breath. There are approximately 10^{22} air molecules in a breath of air and approximately 10^{44} air molecules in the atmosphere. Assuming air has had plenty of time to mix around the world in the past two millennia, what is the probability that the breath of air you just took contains a molecule of air that Caesar exhaled in his dying breath?

Solution: For the first part, we can use the naïve definition of counting and our theorem for ordered sampling from a set with replacement. The total number of options has $n = N^2$ possible items and we are sampling $k = N$ things, so the total number is $(N^2)^N = N^{2N}$. In the numerator, we have $N^2 - N$ red items. How many ways can we select $k = N$ of them? That's just $n^k = (N^2 - N)^N$. So, the probability is:

$$\begin{aligned}\mathbb{P}[\text{all red}] &= \frac{(N^2 - N)^N}{N^{2N}} \\ &= \left(\frac{N^2 - N}{N^2}\right)^N \\ &= \left(1 - \frac{1}{N}\right)^N\end{aligned}$$

And from the handout, we see that in the limit:

$$\lim_{N\to\infty} \mathbb{P}[\text{all red}] = \lim_{N\to\infty} \left(1 - \frac{1}{N}\right)^N = e^{-1}$$

For the second part, we see that this is just the above problem where $N = 10^{22}$, 'red' items are molecules not in Caesar's dying breath and 'black' ones are those that are in being in it. The probability of not having any of the molecules is e^{-1}, so the probability of having at least one is $1 - e^{-1} \approx 0.632$, so greater than 50%! Note that it would be more appropriate to model this as sampling without replacement, but given the large amount of molecules the result would not be noticably different.

ASIDE: A few students asked for a proof of the limit result on the notes. Here is my best attempt in terms of the fact that e is the only base b such that $f(x) = b^x$ is equal to its own derivative. Define $a_n = [1 - 1/n]^n$ and let:

$$f_n(x) = a_n^x = [1 - 1/n]^{n\cdot x}$$

For simplicity, substitute $m = nx$. Then:

$$f_n(x) = [1 - x/m]^m .$$

Now, the trick is to take the derivative of $f_n(x)$ as a function of x:

$$\begin{aligned}\frac{d}{dx}\left[f_n(x)\right] &= \frac{d}{dx}\left[[1-x/m]^m\right] \\ &= m\cdot[1-x/m]^{m-1}\cdot\frac{-1}{m} \\ &= -1\cdot[1-x/m]^{m-1} \\ &= -1\cdot f_n(x)\cdot\left[1-\frac{x}{m}\right]^{-1} \\ &= -1\cdot f_n(x)\cdot[1-n]^{-1}\end{aligned}$$

Notice that the limit of the derivative becomes:

$$\begin{aligned}\lim_{n\to\infty}\frac{d}{dx}\left[f_n(x)\right] &= \lim_{n\to\infty} -1\cdot f_n(x)\cdot[1-n]^{-1} \\ &= -1\cdot f_n(x)\end{aligned}$$

So, $f_n(x)$ is a function such that, for large n, it's derivative is equal to -1 times itself. The only base such that b^x is equal to its own derivative is e^x and therefore the only function that has a derivative equal to the negative of it self is (from the chain rule) e^{-x}. And if $f_n(x) \approx e^{-x}$ then $a_n \approx e^{-1}$.

8. ($\star$) Write a formal proof of Theorem 3.1 using the basic rule of counting and proof by induction. For a proof by induction, you show that something is true for $k = 1$ (trival in this case) and then show that if it is true for k it must be true for $k + 1$.

Solution: The base case is essentially given by the definition: if we have n things there are n ways to pick one ($k = 1$) thing. Now, assume that there are n^k ways to pick k ordered things. We can view picking $k + 1$ things as a two-stage experiment: the first stage is selecting the first k things and the second stage is picking the $(k + 1)$th thing. By assumption, the first stage has n^k options. We've already shown that selecting one thing from a set of n things has n options. By the basic rule of counting, the number of ways of picking $k + 1$ things is then $n^k \cdot n = n^{k+1}$, which finishes the result.

While we will often skip going through all of the formal steps of induction in this course, this technique is essential to providing a proper proof of many probability results.