Math 431: HW2 Solutions

1. **Exercise 1.3**

   (a) This is a Cartesian product where the first factor covers the outcome of the coin flip ($\{H, T\}$ or $\{0, 1\}$, depending on how you want to encode heads and tails) and the second factor represents the outcome of the die. Hence

   $$\Omega = \{0,1\} \times \{1,2,\dots,6\} = \{(i,j) : i = 0 \text{ or } 1 \text{ and } j \in \{1,2,\dots,6\}\}.$$

   (b) Now we need a larger Cartesian product space because the outcome has to contain the coin flip and die roll of each person. Let c_i be the outcome of the coin flip of person i, and let d_i be the outcome of the die roll of person i. Index i runs from 1 to 10 (one index value for each person). Each $c_i \in \{0,1\}$ and each $d_i \in \{1,2,\dots,6\}$. Here are various ways of writing down the sample space:

   $$\begin{aligned}\Omega &= (\{0,1\} \times \{1,2,\dots,6\})^{10} \\ &= \{(c_1, d_1, c_2, d_2, \dots, c_{10}, d_{10}) : \text{each } c_i \in \{0,1\} \text{ and each } d_i \in \{1,2,\dots,6\}\} \\ &= \{(c_i, d_i)_{1\le i\le 10} : \text{each } c_i \in \{0,1\} \text{ and each } d_i \in \{1,2,\dots,6\}\}\end{aligned}$$

   The last formula illustrates the use of indexing to shorten the writing of the 20-tuple of all outcomes. The number of elements is $\#\Omega = 2^{10} \cdot 6^{10} = 12^{10} = 61,917,364,224$.

   (c) If nobody rolled a five, then each die outcome d_i comes from the set $\{1,2,3,4,6\}$ that has 5 elements. Hence the number of these outcomes is $2^{10} \cdot 5^{10} = 10^{10}$. To get the number of outcomes where at least 1 person rolls a five, subtract the number of outcomes where no one rolls a 5 from the total: $12^{10} - 10^{10} = 51,917,364,224$. (The power representations of the answers are sufficient.)

2. **Exercise 1.7**

   (a) This is an ordered sample without replacement. The population size is 7. You can consider the balls as labelled from 1 to 7, with 1, 2, and 3 being green balls. 4, 5, 6, and 7 are yellow balls. The sample size is 3. So we have that $\#\Omega = (7)_3$.

   Now we must count the number of ways the outcome can be Green, Yellow, Green.

   First, you must choose one of the 3 green balls. There are 3 ways to do this.

   Second, you must choose one of the 4 yellow balls. There are 4 ways to do this.

   Third, you must choose one of the 2 remaining green balls. There are 2 ways to do this.

   So there are $3 \cdot 4 \cdot 2$ ways for the event to occur.

   Putting these pieces together, we have the probability

   $$P(E) = \frac{3 \cdot 4 \cdot 2}{(7)_3} = \frac{4}{35}.$$

(b) Now we must compute the number of ways our sample can contain 2 green balls and 1 yellow ball. This means the sample must come out as GGY, GYG, or YGG.

$$\begin{aligned}
&P(\text{2 greens and 1 yellow}) \\
&\quad = P(\text{Green, Green, Yellow}) + P(\text{Green, Yellow, Green}) + P(\text{Yellow, Green, Green}) \\
&\quad = \frac{3 \cdot 2 \cdot 4}{(7)_3} + \frac{3 \cdot 4 \cdot 2}{(7)_3} + \frac{4 \cdot 3 \cdot 2}{(7)_3} \\
&\quad = \frac{12}{35}.
\end{aligned}$$

We have used the same logic as in part (a) to arrive at the 3 probabilities.

3. **Exercise 1.8**

(a) Label the letters from 1 to 14 so that the first 5 are `E`s, the next 4 are `A`s, the next 3 are `N`s and the last 2 are `B`s. This is the population set $S = \{1, 2, \ldots, 14\}$. Our Ω consists of (ordered) sequences of four distinct elements:

$$\Omega = \{(a_1, a_2, a_3, a_4) : a_i \neq a_j, a_i \in S\}.$$

The size of Ω is $14 \cdot 13 \cdot 12 \cdot 11 = 24024$. (Because we can choose a 1 14 different ways, then a 2 13 different ways and so on.)

The event C consists of sequences (a_1, a_2, a_3, a_4) consisting of two numbers between 1 and 5, one between 6 and 9 and one between 10 and 12. We can count these by constructing such a sequence step-by-step: we first choose the positions of the two `E`s: we can do that $\binom{4}{2} = 6$ ways. Then we choose a first `E` out of the 5 choices and place it to the first chosen position. Then we choose the second `E` out of the remaining 4 and place it to the second (remaining) chosen position. Then we choose the `A` out of the 4 choices, and its position (there are 2 possibilities left). Finally we choose the letter `N` out of the 3 choices and place it in the remaining position (we only have one possibility here). In each step the number of choices did not depend on the previous choices so we can just multiply the numbers together to get $6 \cdot 5 \cdot 4 \cdot 4 \cdot 2 \cdot 3 \cdot 1 = 2880$.

Alternately, first you can choose your 4 tiles. 2 must be `E`, so there are $\binom{5}{2} = 10$ ways to do this. Then choose `A`. There are $\binom{4}{1} = 4$ ways to do this. Then choose `N`. There are $\binom{3}{1} = 3$ ways to do this. Once you have your four tiles, you put them in order. There are 4! ways to do this. So in total, there are $10 \cdot 4 \cdot 3 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 2880$ ways to draw the tiles.

In either case, we have

$$P(C) = \frac{2880}{24024} = \frac{120}{1001}.$$

(b) We have an unordered sample without replacement. We can represent the population as $S = \{1, 2, \ldots, 14\}$, where 1 through 5 are `E`, 6 through 9 are `A`, 10 through 12 are `N`, and 13 through 14 are `S`. So the sample space is

$$\Omega = \{\omega \subset S \, : \, \#\omega = 4\}.$$

Thus we have $\#\Omega = \binom{14}{4}$.

The event C is that $\{a_1, a_2, a_3, a_4\}$ has two numbers between 1 and 5, one between 6 and 9 and one between 10 and 12. The number of ways we can choose such a set is $\binom{5}{2}\binom{4}{1}\binom{3}{1} = 120$. (Because we can choose the two `E`s out of 5 possibilities, the single `A` out of 4 possibilities and the single `N` out of 3 possibilities.)
So the conclusion is that

$$P(C) = \frac{\binom{5}{2}\binom{4}{1}\binom{3}{1}}{\binom{14}{4}} = \frac{120}{1001}.$$

4. **Exercise 1.9**

Assume the length of the chalk is one unit. Then we set $\Omega = [0, 1]$, to represent the location where the piece of chalk breaks. Then we have

$$\begin{aligned} E &= \text{Event that shorter piece is less than } 1/5 \text{ of original} \\ &= [0, 1/5] \cup [4/5, 1]. \end{aligned}$$

So

$$\begin{aligned} P(E) &= P([0, 1/5]) + P([4/5, 1]) \\ &= \frac{1/5 - 0}{1 - 0} + \frac{1 - 4/5}{1 - 0} \\ &= \frac{2}{5}. \end{aligned}$$

5. **Exercise 1.10**

(a) Modeling the outcome of the experiment as the number of times we rolled the die (as we did in Example 1.16), we take

$$\Omega = \{\infty, 1, 2, 3, \dots\}.$$

For $k \geq 1$, we have

$$P(k) = P\{\text{needed } k \text{ rolls}\} = P\{\text{no fours in the first } k-1 \text{ rolls, then a 4}\}.$$

Counting the number of ways of not rolling a four on the $k-1$ first rolls, we have

$$P(k) = P\{\text{no fours in the first } k-1 \text{ rolls, then a 4}\} = \frac{5^{k-1} \cdot 1}{6^k} = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}.$$

(b) Since the outcomes are mutually exclusive,

$$\begin{aligned} 1 = P(\Omega) &= P(\infty) + \sum_{k=1}^{\infty} P(k) \\ &= P(\infty) + \sum_{k=1}^{\infty} \left(\frac{5}{6}\right)^{k-1} \frac{1}{6} \\ &= P(\infty) + \frac{1}{6} \sum_{j=0}^{\infty} \left(\frac{5}{6}\right)^{j} && \text{(reindex)} \\ &= P(\infty) + \frac{1}{6} \frac{1}{1 - 5/6} && \text{(geometric series)} \\ &= P(\infty) + 1. \end{aligned}$$

Thus, $P(\infty) = 0$.

Another possible solution: You can use the monotonicity of probability, though this is not proved until the subsequent textbook section. As the textbook is organized, this approach would not be (proven) valid in Section 1.3. Define the events

$$\begin{aligned} E_n &= \text{No fours in the first } n \text{ rolls} \\ \{\infty\} &= \text{No four ever appears} \\ \Rightarrow \{\infty\} &\subset E_n \quad \text{for all } n. \end{aligned}$$

Thus, monotonicity gives us

$$\begin{aligned} P(\infty) \leq P(E_n) &= \left(\frac{5}{6}\right)^n \qquad \text{for all } n \\ \Rightarrow P(\infty) \leq \lim_{n\to\infty} \left(\frac{5}{6}\right)^n &= 0 \end{aligned}$$

The probability on the left must be zero.

6. **Exercise 1.28**

Assume that both m and n are at least 1 so the problem is not trivial.

(a) Without replacement. We can compute the answer using either an ordered or an unordered sample, the result will be the same. It helps to assume that the balls are labeled (e.g. by numbering them from 1 to $n+m$), although the actual labeling will not play a role in the computation.

With an ordered sample we have $(m+n)(m+n-1)$ outcomes (we have $m+n$ choices for the first pick and $m+n-1$ choices for the second). The favorable outcomes can be counted by considering green-green and yellow-yellow pairs separately: this is $m(m-1)+n(n-1)$. The answer is the ratio of the number of favorable outcomes and the total number of outcomes,

$$P\{(\text{g,g}) \text{ or } (\text{y,y})\} = \frac{m(m-1)+n(n-1)}{(m+n)(m+n-1)}.$$

The unordered sample calculation gives the same answer:

$$P\{\text{a set of two greens or a set of two yellows}\} = \frac{\binom{m}{2}+\binom{n}{2}}{\binom{m+n}{2}} = \frac{m(m-1)+n(n-1)}{(m+n)(m+n-1)}.$$

Note: for integers $0 \leq k < \ell$, the convention is $\binom{k}{\ell} = 0$. This makes the answers above correct even if m or n or both are 1.

(b) With replacement. Now the sample has to be ordered (there is a first pick and a second pick). The total number of outcomes is $(m+n)^2$, and the number of favorable outcomes (again counting the green-green and yellow-yellow pairs separately) is m^2+n^2. This gives

$$P\{(\text{g,g}) \text{ or } (\text{y,y})\} = \frac{m^2+n^2}{(m+n)^2}.$$

(c) The question means: under what conditions on m and n is the probability larger in part (a). We simplify the inequality through a sequence of equivalences, by cancelling factors, multiplying away the denominators, and then cancelling some more.

$$\begin{aligned}
& \text{answer to (a)} \ > \ \text{answer to (b)} \\
\iff \quad & \frac{m(m-1)+n(n-1)}{(m+n)(m+n-1)} > \frac{m^2+n^2}{(m+n)^2} \\
\iff \quad & \frac{m(m-1)+n(n-1)}{m+n-1} > \frac{m^2+n^2}{m+n} \\
\iff \quad & (m(m-1)+n(n-1))(m+n) > (m^2+n^2)(m+n-1) \\
\iff \quad & (m^2-m+n^2-n)(m+n) > (m^2+n^2)(m+n) - m^2 - n^2 \\
\iff \quad & (-m-n)(m+n) > -m^2-n^2 \\
\iff \quad & (m+n)^2 < m^2+n^2 \\
\iff \quad & 2mn < 0.
\end{aligned}$$

The last inequality is *always false* for positive m or n. Since the last inequality is equivalent to the first one, the first one is also *always false*.

The conclusion we take from this is that if you want to maximize your chances of getting two of the same color, you want to sample with replacement rather than without replacement. Intuitively this should be obvious: once you remove a ball, you have diminished the chances of drawing another one of the same color.