Section 3 – Solutions

Review

- **Conditional Probability.** $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) =$ __________

  $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) = \frac{\mathbb{P}(\mathcal{B} \cap \mathcal{A})}{\mathbb{P}(\mathcal{A})}$

- **Independent Events.** Two events $\mathcal{A}, \mathcal{B}$ are **independent** if $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) =$ __________.

  $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\,\mathbb{P}(\mathcal{B})$

  If $\mathbb{P}(\mathcal{A}) \neq 0$, this is equivalent to $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) =$ __________.

  $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) = \mathbb{P}(\mathcal{B})$

  If $\mathbb{P}(\mathcal{B}) \neq 0$, this is equivalent to $\mathbb{P}(\mathcal{A} \mid \mathcal{B}) =$ __________.

  $\mathbb{P}(\mathcal{A} \mid \mathcal{B}) = \mathbb{P}(\mathcal{A})$

- **Partition.** Nonempty events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ partition the sample space Ω iff

  (1) ____________________ (2) ____________________

  (1) $\bigcup_{i=1}^{n} \mathcal{E}_i = \Omega$ (2) $\forall i \forall i \neq j, \mathcal{E}_i \cap \mathcal{E}_j = \varnothing$

- **Bayes Rule.** For any events $\mathcal{A}$ and $\mathcal{B}$, $\mathbb{P}(\mathcal{A} \mid \mathcal{B}) =$ ____________ .

  $\mathbb{P}(\mathcal{A} \mid \mathcal{B}) = \frac{\mathbb{P}(\mathcal{B}|\mathcal{A})\mathbb{P}(\mathcal{A})}{\mathbb{P}(\mathcal{B})}$

- **Chain Rule:** Suppose $\mathcal{A}_1, ..., \mathcal{A}_n$ are events. Then,

$$\mathbb{P}(\mathcal{A}_1 \cap ... \cap \mathcal{A}_n) = \_\_\_\_\_\_\_\_\_\_$$

  $\mathbb{P}(\mathcal{A}_1 \cap ... \cap \mathcal{A}_n) = \prod_{i=1}^{n} \mathbb{P}(\mathcal{A}_i \mid \mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_{i-1})$

- **Law of Total Probability (LTP)**: Suppose $\mathcal{E}_1, \ldots, \mathcal{E}_n$ is a partition of Ω and let $\mathcal{B}$ be any event. Then $\mathbb{P}(\mathcal{B}) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \cap \mathcal{E}_i) =$ __________

  $\mathbb{P}(\mathcal{B}) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \cap \mathcal{E}_i) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \mid \mathcal{E}_i)\,\mathbb{P}(\mathcal{E}_i)$

- **Bayes Theorem with LTP:** $\mathcal{E}_1, \ldots, \mathcal{E}_n$ is a partition of Ω and let $\mathcal{B}$ be any event. Then

$$\mathbb{P}(\mathcal{E}_1 \mid \mathcal{B}) = \frac{\_\_\_\_\_\_\_\_\_\_}{\sum_{i=1}^{n} \_\_\_\_\_\_\_\_\_\_}.$$

  $\mathbb{P}(\mathcal{E}_1 \mid \mathcal{B}) = \frac{\mathbb{P}(\mathcal{B} \mid \mathcal{E}_1)\mathbb{P}(\mathcal{E}_1)}{\sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \mid \mathcal{E}_i)\mathbb{P}(\mathcal{E}_i)}$.

The following will be covered in lecture on Friday.

- **Random Variable (rv)**: A numeric function $X : \Omega \to \mathbb{R}$ of the outcome.

- **Range/Support:** The support/range of a random variable X, denoted Ω_X, is the set of all possible values that X can take on.

- **Discrete Random Variable (drv)**: A random variable taking on a countable (either finite or countably infinite) number of possible values.
- **Probability Mass Function (pmf) for a discrete random variable $\mathbf{X}$**: a function $p_X : \Omega_X \rightarrow [0, 1]$ with $p_X(x) = \mathbb{P}(() X = x)$ that maps possible values of a discrete random variable to the probability of that value happening, such that $\sum_x p_X(x) = 1$.
- **Cumulative Distribution Function (CDF) for a random variable $\mathbf{X}$**: a function $F_X : R \rightarrow R$ with $F_X(x) = \mathbb{P}(() X \leqslant x)$
- **Expectation (expected value, mean, or average)**: The expectation of a discrete random variable is defined to be $\mathbb{E}[X] = \sum_x x p_X(x) = \sum_x x \mathbb{P}(() X = x)$. The expectation of a function of a discrete random variable $g(X)$ is $\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$.

Task 1 – Naive Bayes

This Ed Lesson (https://edstem.org/us/courses/50547/lessons/87272/slides/479664) is an introduction to using Bayes theorem to classify spam emails. This will be what you implement in the coding portion of PSet 3!

Task 2 – Flipping Coins

We consider two independent tosses of the same coin. The coin is "heads" one quarter of the time.

a) What is the probability that the second toss is "heads" given that the first toss is "tails"?

Consider the probability space with sample space $\Omega = \{HH, TT, HT, TH\}$. Because heads come $1/4$ of the time, and tails $3/4$, we have

$$\begin{aligned} \mathbb{P}(HH) &= 1/4 \times 1/4 = 1/16 \\ \mathbb{P}(HT) &= 1/4 \times 3/4 = 3/16 \\ \mathbb{P}(TH) &= 3/4 \times 1/4 = 3/16 \\ \mathbb{P}(TT) &= 3/4 \times 3/4 = 9/16 \, . \end{aligned}$$

Let $\mathcal{A}$ be the event that the first toss is tails, and let $\mathcal{B}$ be the event that the second toss is heads. By Bayes Rule,

$$\mathbb{P}(\mathcal{B} \mid \mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} \, .$$

Note that $\mathcal{A} = \{TT, TH\}$ and $\mathcal{B} = \{HH, TH\}$, and thus

$$\begin{aligned} \mathbb{P}(\mathcal{A}) &= \mathbb{P}(TT) + \mathbb{P}(TH) = 9/16 + 3/16 = 12/16 = 3/4 \\ \mathbb{P}(\mathcal{A} \cap \mathcal{B}) &= \mathbb{P}(TH) = 3/16 \, . \end{aligned}$$

Therefore,

$$\mathbb{P}(\mathcal{B} \mid \mathcal{A}) = \frac{3/16}{3/4} = \frac{1}{4} \, .$$

We see here that the probability the second toss is heads is "independent" of the first toss being tails. This is exactly what we would have expected – indeed, we model the coins to be independent.

b) What is the probability that the second toss is "heads" given that at least one of the tosses is "tails"?

We are looking to find the probability that

$$\underbrace{\textbf{the second toss is ``heads''}}_{\mathcal{B}} \text{ given that } \underbrace{\textbf{at least one of the tosses is ``tails''}}_{\mathcal{C}} \, .$$

To do this, let us define $\mathcal{B} = \{HH, TH\}$ to the event that the second toss is heads (notice that this is the same as in the previous part). We also need to define $\mathcal{C}$, the event that we have at least one of the tosses is "tails" – this is $\mathcal{C} = \{TH, HT, TT\}$. As we noted above, we want $\mathbb{P}(\mathcal{B} \mid \mathcal{C})$. Note that

$$\begin{aligned} \mathbb{P}(\mathcal{C}) &= 1 - \mathbb{P}(HH) = 15/16 \\ \mathbb{P}(\mathcal{B} \cap \mathcal{C}) &= \mathbb{P}(TH) = 3/16 \, . \end{aligned}$$

We can also calculate $\mathbb{P}(\mathcal{C})$ directly:

$$\mathbb{P}(\mathcal{C}) = \mathbb{P}(TH) + \mathbb{P}(HT) + \mathbb{P}(TT) = 3/16 + 3/16 + 9/16 = 15/16 \, .$$

Therefore,

$$\mathbb{P}(\mathcal{B} \mid \mathcal{C}) = \frac{\mathbb{P}(\mathcal{B} \cap \mathcal{C})}{\mathbb{P}(\mathcal{C})} = \frac{3/16}{15/16} = \frac{3}{15} = \frac{1}{5} \, .$$

c) In the probability space of this task, give an example of two events that are disjoint but not independent.

Let $\mathcal{A}$ be the event that we get only heads, and let $\mathcal{B}$ be the event that we get only tails. Formally, $\mathcal{A} = \{HH\}$ and $\mathcal{B} = \{TT\}$.

The two events are disjoint, but not independent. To see that they are disjoint, we note that $\mathcal{A} \cap \mathcal{B} = \varnothing$, and so $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = 0$. However,

$$\mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}) = 9/16 \cdot 1/16 = 9/256 \neq 0 = \mathbb{P}(\mathcal{A} \cap \mathcal{B}) \ .$$

Remember that two events are independent if and only if the probability they both occur is equal to their individual probabilities multiplied together – hence the name :).

d) In the probability space of this task, give an example of two events that are independent but not disjoint.

Let $\mathcal{A}$ be the event that we get heads in our second toss, and let $\mathcal{B}$ be the event that we get tails in our first toss. Formally, $\mathcal{A} = \{TH, HH\}$ and $\mathcal{B} = \{TH, TT\}$.

The two events are not disjoint, but are independent! To see that they are not disjoint, we note that TH is in both of $\mathcal{A}$ and $\mathcal{B}$, i.e., $\mathcal{A} \cap \mathcal{B} = \{TH\} \neq \varnothing$. Additionally,

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(TH) = \frac{3}{16} \ .$$

We also have

$$\mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}) = \left(\frac{3}{16} + \frac{1}{16}\right) \cdot \left(\frac{3}{16} + \frac{9}{16}\right) = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16} = \mathbb{P}(\mathcal{A} \cap \mathcal{B}) \ ,$$

so the two events must be independent!

Task 3 – Balls from an Urn – Take 2

Say an urn contains three red balls and four blue balls. Imagine we draw three balls without replacement. (You can assume every ball is uniformly selected among those remaining in the urn.)

a) What is the probability that all three balls are all of the same color?

The experiment is modeled with $\Omega = \{r, b\}^3$, i.e., every possible arrangement of r and b in three balls. Probabilities are assigned as we have seen in class, by assuming every draw is uniform among the remaining balls. Then, note that

$$\mathbb{P}(rrr) = \frac{3}{7} \cdot \frac{2}{6} \cdot \frac{1}{5} = \frac{1}{35} \quad \text{and} \quad \mathbb{P}(bbb) = \frac{4}{7} \cdot \frac{3}{6} \cdot \frac{2}{5} = \frac{4}{35} \; .$$

Therefore, the probability that they all have the same color is $1/35 + 4/35 = 1/7$.

b) What is the probability that we get more than one red ball given the first ball is red?

Let $\mathcal{R}$ be the event that the first ball is red. It is clear from the problem statement that $\mathbb{P}(\mathcal{R}) = \frac{3}{7}$. We also consider the event $\mathcal{M}$ that we have more than one red ball. Let $\mathcal{M}$ be the event that more than one ball is red. We need to now compute the probability $\mathbb{P}(\mathcal{M} \cap \mathcal{R})$. Note that by the law of total probability,

$$\mathbb{P}(\mathcal{R}) = \mathbb{P}(\mathcal{R} \cap \mathcal{M}) + \mathbb{P}(\mathcal{R} \cap \mathcal{M}^c) \; ,$$

and so

$$\mathbb{P}(\mathcal{M} \cap \mathcal{R}) = \mathbb{P}(\mathcal{R}) - \mathbb{P}(\mathcal{R} \cap \mathcal{M}^c) = 3/7 - \mathbb{P}(\mathcal{R} \cap \mathcal{M}^c) \; .$$

Notice that $\mathcal{M}^c$ is the event that there is not more than one red ball, i.e., there is at most one red ball. Then $\mathcal{R} \cap \mathcal{M}^c$ is the event that the first ball is red, and both remaining balls are blue (we cannot have any more red balls). This gives us

$$\mathbb{P}(\mathcal{M}^c \cap \mathcal{R}) = \frac{3}{7} \cdot \frac{4}{6} \cdot \frac{3}{5} = \frac{6}{35} \; .$$

Thus,

$$\mathbb{P}(\mathcal{M} \cap \mathcal{R}) = 3/7 - 6/35 = 9/35 \; .$$

We apply the definition of conditional probability to get

$$\mathbb{P}(\mathcal{M} \mid \mathcal{R}) = \frac{\mathbb{P}(\mathcal{M} \cap \mathcal{R})}{\mathbb{P}(\mathcal{R})} = \frac{9/35}{3/7} = \frac{3}{5} \; .$$

Task 4 – Game Show

Corrupted by their power, the judges running the popular game show America's Next Top Mathematician have been taking bribes from many of the contestants. During each of two episodes, a given contestant is either allowed to stay on the show or is kicked off. If the contestant has been bribing the judges, they will be allowed to stay with probability 1. If the contestant has not been bribing the judges, they will be allowed to stay with probability 1/3, independent of what happens in earlier episodes. Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds.

a) If you pick a random contestant, what is the probability that they are allowed to stay during the first episode?

Let S_i be the event they stayed during the i-th episode. We define B to be the event that a contestant has bribed the judges. As stated in the problem,

$$\begin{aligned}
\mathbb{P}(B) &= 1/4 \\
\mathbb{P}(\overline{B}) &= 1 - 1/4 = 3/4 \\
\mathbb{P}(S_1 \mid B) &= 1 \\
\mathbb{P}(S_1 \mid \overline{B}) &= 1/3 \,.
\end{aligned}$$

By the Law of Total Probability – conditioning on whether the contestant bribed the judges – we get

$$\mathbb{P}(S_1) = \mathbb{P}(S_1 \mid B)\,\mathbb{P}(B) + \mathbb{P}(S_1 \mid \overline{B})\,\mathbb{P}(\overline{B}) = 1 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} = \boxed{\frac{1}{2}}\,.$$

b) If you pick a random contestant, what is the probability they are allowed to stay during both episodes?

Let S_i be defined as before. Staying during both episodes is equivalent to the contestant staying in episodes 1 and 2, so the event is $S_1 \cap S_2$. By the Law of Total Probability, conditioned on whether the contestant bribed the judges, we get

$$\mathbb{P}(S_1 \cap S_2) = \mathbb{P}(S_1 \cap S_2 \mid B)\,\mathbb{P}(B) + \mathbb{P}(S_1 \cap S_2 \mid \overline{B})\,\mathbb{P}(\overline{B}) \tag{1}$$

We know a contestant is guaranteed to stay on the show, given that they are bribing the judges, hence:

$$\mathbb{P}(S_1 \cap S_2 \mid B) = 1$$

On the other hand, if they have not been bribing judges, then the probability they stay on the show is $1/3$, independent of what happens on earlier episodes. By conditional independence, we have:

$$\mathbb{P}(S_1 \cap S_2 \mid \overline{B}) = \mathbb{P}(S_1 \mid \overline{B})\,\mathbb{P}(S_2 \mid \overline{B}) = \frac{1}{3} \cdot \frac{1}{3}$$

Plugging our results above into equation (1) gives us:

$$\mathbb{P}(S_1 \cap S_2) = 1 \cdot \frac{1}{4} + \left(\frac{1}{3} \cdot \frac{1}{3}\right) \cdot \frac{3}{4} = \boxed{\frac{1}{3}}$$

c) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they get kicked off during the second episode?

By the definition of conditional probability and the Law of Total Probability,

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{\mathbb{P}(S_1 \cap \overline{S_2})}{\mathbb{P}(S_1)} = \frac{\mathbb{P}(S_1 \cap \overline{S_2} \mid B)\, \mathbb{P}(B) + \mathbb{P}(S_1 \cap \overline{S_2} \mid \overline{B})\, \mathbb{P}(\overline{B})}{\mathbb{P}(S_1)} .$$

We have already computed $\mathbb{P}(S_1)$ in part (a). We compute the numerator term by term. Given that a contestant is bribing the judges, they are guaranteed to stay on the show. As such:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid B) = \mathbb{P}(S_1 \mid B) \cdot \mathbb{P}(\overline{S_2} \mid B) = 1 \cdot 0 = 0 .$$

On the other hand, if they have not been bribing judges, the probability they leave the show is $2/3$ (by complementing). We can then write:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid \overline{B}) = \mathbb{P}(S_1 \mid \overline{B}) \cdot \mathbb{P}(\overline{S_2} \mid \overline{B}) = \frac{1}{3} \cdot \frac{2}{3} .$$

We can now evaluate our initial expression:

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{0 \cdot \frac{1}{4} + \left(\frac{1}{3} \cdot \frac{2}{3}\right) \cdot \frac{3}{4}}{\frac{1}{2}} = \frac{1/6}{1/2} = \boxed{\frac{1}{3}} .$$

d) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they were bribing the judges?

By Bayes' Theorem,

$$\mathbb{P}(B \mid S_1) = \frac{\mathbb{P}(S_1 \mid B)\, \mathbb{P}(B)}{\mathbb{P}(S_1)} = \frac{1 \cdot \frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}} .$$

Task 5 – Allergy Season

In a certain population, everyone is equally susceptible to colds. Each person, in particular, catches a cold with probability 0.2.

The number of colds suffered by each person during each winter season ranges from 0 to 4, with probability 0.2 for each value (see table below). A new cold prevention drug is introduced that, for people for whom the drug is effective, changes the probabilities as shown in the table. Unfortunately, the effects of the drug last only the duration of one winter season, and the drug is only effective in 20% of people, independently.

number of colds	no drug or ineffective	drug effective
0	0.2	0.4
1	0.2	0.3
2	0.2	0.2
3	0.2	0.1
4	0.2	0.0

a) Sneezy decides to take the drug. Given that he gets 1 cold that winter, what is the probability that the drug is effective for Sneezy?

Let E be the event that the drug is effective for Sneezy, and C_i be the event that he gets i colds the first winter. By Bayes' Theorem,

$$\mathbb{P}(E \mid C_1) = \frac{\mathbb{P}(C_1 \mid E)\, \mathbb{P}(E)}{\mathbb{P}(C_1 \mid E)\, \mathbb{P}(E) + \mathbb{P}(C_1 \mid \overline{E})\, \mathbb{P}(\overline{E})} = \frac{0.3 \cdot 0.2}{0.3 \cdot 0.2 + 0.2 \cdot 0.8} = \frac{3}{11}$$

b) The next year he takes the drug again. Given that he gets 2 colds in this winter, what is the updated probability that the drug is effective for Sneezy?

Here, we need to consider the idea that Sneezy's drug effectiveness is no longer 20% like the general population. We will need to reduce our sample to the event of one occurrence of cold in previous season, and effectiveness of the drug given the one occurrence. Let the reduced sample space for part (b) be C_1 from part (a), so that $\mathbb{P}_{C_1}(E) = \mathbb{P}_{\Omega}(E \mid C_1)$. Let D_i be the event that he gets i colds the second winter. By Bayes' Theorem,

$$\mathbb{P}(E \mid D_2) = \frac{\mathbb{P}(D_2 \mid E)\,\mathbb{P}(E)}{\mathbb{P}(D_2 \mid E)\,\mathbb{P}(E) + \mathbb{P}(D_2 \mid \overline{E})\,\mathbb{P}(\overline{E})} = \frac{0.2 \cdot \frac{3}{11}}{0.2 \cdot \frac{3}{11} + 0.2 \cdot \frac{8}{11}} = \frac{3}{11}$$

c) Why is the answer to (b) the same as the answer to (a)?

The probability of two colds whether or not the drug was effective is the same. Hence knowing that Sneezy got two colds does not change the probability of the drug's effectiveness.

Task 6 – Coins

There are three coins, C_1, C_2, and C_3. The probability of "heads" is 1 for C_1, 0 for C_2, and p for C_3. A coin is picked among these three uniformly at random, and then flipped a certain number of times.

a) What is the probability that the first n flips are tails?

By the Law of Total Probability, partitioning on which coin was chosen, we have

$$\frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot (1-p)^n = \frac{1}{3} + \frac{1}{3}(1-p)^n \, .$$

b) Given that the first n flips were tails, what is the probability that C_1 was flipped? The probability that C_2 was flipped? The probability that C_3 was flipped?

We use Bayes Rule:

$$\mathbb{P}(C_i \mid n \text{ tails}) = \frac{\mathbb{P}(n \text{ tails} \mid C_i)\,\mathbb{P}(C_i)}{\mathbb{P}(n \text{ tails})}$$

Applying this, we get

$$\begin{aligned}
\mathbb{P}(C_1 \mid n \text{ tails}) &= \frac{1/3 \cdot 0}{1/3 + 1/3(1-p)^n} = 0 \\
\mathbb{P}(C_2 \mid n \text{ tails}) &= \frac{1/3 \cdot 1}{1/3 + 1/3(1-p)^n} = \frac{1}{1+(1-p)^n} \\
\mathbb{P}(C_3 \mid n \text{ tails}) &= \frac{1/3 \cdot (1-p)^n}{1/3 + 1/3(1-p)^n} = \frac{(1-p)^n}{1+(1-p)^n}
\end{aligned}$$

Task 7 – Parallel Systems

A parallel system functions whenever at least one of its components works. Consider a parallel system of n components and suppose that each component works with probability p independently.

a) What is the probability the system is functioning?

Let C_i be the event component i is working, and F be the event that the system is functioning. For the system to function, it is sufficient for any component to be working. This means that the only case in which the system does not function is when none of the components work. We can then use complementing to compute $\mathbb{P}(F)$, knowing that $\mathbb{P}(C_i) = p$. That is,

$$\begin{aligned}
\mathbb{P}(F) &= 1 - \mathbb{P}(F^c) = 1 - \mathbb{P}\left(\bigcap_{i=1}^{n} C_i^c\right) = 1 - \prod_{i=1}^{n} \mathbb{P}(C_i^c) = 1 - \prod_{i=1}^{n} \big(1 - \mathbb{P}(C_i)\big) \\
&= 1 - \prod_{i=1}^{n} (1-p) = \boxed{1-(1-p)^n} \, .
\end{aligned}$$

Note that $\mathbb{P}\big(\bigcap_{i=1}^n C_i^c\big) = \prod_{i=1}^n \mathbb{P}(C_i^c)$ due to independence of C_i (components working <u>independently</u> of each other). Note also that $\prod_{i=1}^n a = a^n$ for any constant a.

b) If the system is functioning, what is the probability that component 1 is working?

We know that for the system to function only one component needs to be working, so for all i, we have $\mathbb{P}(F \mid C_i) = 1$. Using Bayes Theorem, we get:

$$\mathbb{P}(C_1 | F) = \frac{\mathbb{P}(F|C_1)\,\mathbb{P}(C_1)}{\mathbb{P}(F)} = \frac{1 \cdot p}{1-(1-p)^n} = \boxed{\frac{p}{1-(1-p)^n}}$$

c) If the system is functioning and component 2 is working, what is the probability that component 1 is working?

We have

$$\mathbb{P}(C_1 \mid C_2 \cap F) = \mathbb{P}(C_1 \mid C_2) = \mathbb{P}(C_1) = p \,,$$

where the first equality holds because knowing C_2 and F is just as good as knowing C_2 (since if C_2 happens, F does too), and the second equality holds because the components working are independent of each other.

More formally, we can use the definition of conditional probability along with a careful application of the chain rule to get the same result. We start with the following expression:

$$\mathbb{P}(C_1 \mid C_2 \cap F) = \frac{\mathbb{P}(C_1 \cap C_2 \cap F)}{\mathbb{P}(C_2 \cap F)} = \frac{\mathbb{P}(F \mid C_1 \cap C_2) \cdot P(C_1 \mid C_2) \cdot \mathbb{P}(C_2)}{\mathbb{P}(F \mid C_2) \cdot \mathbb{P}(C_2)}$$

We note that the system is guaranteed to work if any one component is working, so $\mathbb{P}(F \mid C_1 \cap C_2) = \mathbb{P}(F \mid C_2) = 1$. We also note that components work independently of each other, hence $\mathbb{P}(C_1 \mid C_2) = \mathbb{P}(C_1)$. With that in mind, we can rewrite our expression so that:

$$\mathbb{P}(C_1 \mid C_2 \cap F) = \frac{1 \cdot \mathbb{P}(C_1) \cdot \mathbb{P}(C_2)}{1 \cdot \mathbb{P}(C_2)} = \mathbb{P}(C_1) = \boxed{p}$$

Task 8 – Marbles in Pockets

A girl has 5 blue and 3 white marbles in her left pocket, and 4 blue and 4 white marbles in her right pocket. If she transfers a randomly chosen marble from left pocket to right pocket without looking, and then draws a randomly chosen marble from her right pocket, what is the probability that it is blue?

Let $W_\_, B_\_$ denote the event that we choose a white marble or a blue marble respectively, with subscripts L, R indicating from which pocket we are picking – left and right, respectively.
We know that we will pick from the left pocket first, and right pocket second. We can then use the Law of Total Probability conditioning on the color of the transferred marble so that:

$$\mathbb{P}(B_R) = \mathbb{P}(W_L) \cdot \mathbb{P}(B_R \mid W_L) + \mathbb{P}(B_L) \cdot \mathbb{P}(B_R \mid B_L) = \frac{3}{8} \cdot \frac{4}{9} + \frac{5}{8} \cdot \frac{5}{9} = \boxed{\frac{37}{72}}$$

Task 9 – A game

Pemi and Shreya are playing the following game: A 6-sided die is thrown and each time it's thrown, regardless of the history, it is equally likely to show any of the six numbers.

- If it shows 5, Pemi wins.
- If it shows 1, 2, or 6, Shreya wins.
- Otherwise, they play a second round and so on.

a) What is the probability that Shreya wins on the 4th round?

Let S_i be the event that Shreya wins on the i-th round and let N_i be the event that nobody wins on the i-th round. Then we are interested in the event

$$N_1 \cap N_2 \cap N_3 \cap S_4 \,.$$

Using the chain rule, we have

$$\begin{aligned}\mathbb{P}(N_1 \cap N_2 \cap N_3 \cap S_4) &= \mathbb{P}(N_1) \cdot \mathbb{P}(N_2 \mid N_1) \cdot \mathbb{P}(N_3 \mid N_1 \cap N_2) \cdot \mathbb{P}(S_4 \mid N_1 \cap N_2 \cap N_3) \\ &= \frac{2}{6} \cdot \frac{2}{6} \cdot \frac{2}{6} \cdot \frac{3}{6} = \frac{1}{54}.\end{aligned}$$

We used the fact that the probability that the game continues for another round (or, equivalently, nobody wins that round) is the probability of rolling a 3 or 4, which is $1/3$.

b) What is the probability that Shreya wins on the ith round?

Using the same event definitions as the previous part, we are now interested in the event

$$N_1 \cap N_2 \cap N_3 \cap ... \cap N_{i-1} \cap S_i \ .$$

Using the chain rule, we have $\mathbb{P}(N_1 \cap N_2 \cap N_3 \cap ... \cap N_{i-1} \cap S_i) =$
$\mathbb{P}(N_1) \cdot \mathbb{P}(N_2 \mid N_1) \cdot \mathbb{P}(N_3 \mid N_1 \cap N_2) \cdot ... \cdot \mathbb{P}(N_{i-1} \mid N_1 \cap N_2 \cap ... \cap N_{i-2}) \cdot \mathbb{P}(S_i \mid N_1 \cap N_2 \cap ... \cap N_{i-1})$

The probability of nobody winning a particular round is always $2/6 = 1/3$ independent of the history. The probability of Shreya winning a particular round is always $3/6 = 1/2$ independent of previous rounds. Thus, the final probability is

$$\left(\frac{1}{3}\right)^{i-1} \cdot \frac{1}{2}$$

c) What is the probability that Shreya wins? Use the fact that if $|x| < 1$, $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$

From part b, we know that the probability that Shreya wins on the ith round is $\mathbb{P}(S_i) = \frac{1}{2} \cdot \left(\frac{1}{3}\right)^{i-1}$

Let S be the even that Shreya wins. Notice that this means she count win on round 1 OR round 2 OR ... OR after an infinite number of rounds! Then,

$$\mathbb{P}(S) = \mathbb{P}(S_1 \cup S_2 \cup ...) = \sum_{i=1}^{\infty} \mathbb{P}(S_i)$$

since all S_i's are disjoint.

Thus,

$$\begin{aligned}\mathbb{P}(S) &= \sum_{i=1}^{\infty} \frac{1}{2} \cdot \left(\frac{1}{3}\right)^{i-1} \\ &= \frac{1}{2} \cdot \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-1} \\ &= \frac{1}{2} \cdot \frac{3}{2} && \left[\text{if } |x| < 1, \sum_{i=0}^{\infty} x^i = \frac{1}{1-x}\right] \\ \mathbb{P}(S) &= \frac{3}{4}\end{aligned}$$

Task 10 – Another game

Leiyi and Luxi are playing a tournament in which they stop as soon as one of them wins n games. Luxi wins each game with probability p and Leiyi wins with probability $1-p$, independently of other games. What is the probability that Luxi wins and that when the match is over, Leiyi has won k games?

Since the match is over when someone wins the n^{th} game, and Luxi won the match, Luxi won the last game.

Before this, Luxi must've won $n-1$ games and Leiyi must've won k games. Therefore, the probability that we reach a point in time when Luxi has won $n-1$ games and Leiyi has won k games is: $p^{n-1} \cdot (1-p)^k \cdot \binom{n-1+k}{k}$. The binomial coefficient counts the number of ways of picking the k games that Leiyi has won out of $n-1+k$ games.

At that point in time, we want Luxi to win the next game so that she has won n games. This happens with probability p, independent of previous outcomes. Therefore, our final probability is:

$$p^{n-1} \cdot (1-p)^k \cdot \binom{n-1+k}{k} \cdot p = p^n \cdot (1-p)^k \cdot \binom{n-1+k}{k}.$$

Task 11 – Balls from an Urn – Take 3

An urn contains 3 red and 3 blue balls with probability 3/5, 3 red and 1 blue balls with probability 1/10, and 5 red and 7 blue balls with probability 3/10. We draw a ball at random from the urn. Let R be the event that we draw a red ball. Let $xRyB$ be the event that the urn contains x red balls and y blue balls.
Are the events R and $3R3B$ independent?

By the definition of independence, we know that R are 3R3B are independent if and only if $\mathbb{P}(R \mid 3R3B) = \mathbb{P}(R)$. Let's calculate these.

$\mathbb{P}(R \mid 3R3B)$ is the probability of choosing a red ball given that the urn has 3 red and 3 blue balls. Thus,

$$\mathbb{P}(R \mid 3R3B) = \frac{3}{6} = \frac{1}{2}$$

$\mathbb{P}(R)$ is the probability of choosing a red ball in general. Since the urn can have a different split of red vs blue balls, we use the possible splits to partition the sample space and apply the law of total probability:

$$\begin{aligned}
\mathbb{P}(R) &= \mathbb{P}(R \mid 3R3B) \cdot \mathbb{P}(3R3B) + \mathbb{P}(R \mid 3R1B) \cdot \mathbb{P}(3R1B) + \mathbb{P}(R \mid 5R7B) \cdot \mathbb{P}(5R7B) \\
&= \frac{1}{2} \cdot \frac{3}{5} + \frac{3}{4} \cdot \frac{1}{10} + \frac{5}{12} \cdot \frac{3}{10} \\
\mathbb{P}(R) &= \frac{1}{2}
\end{aligned}$$

Since $\mathbb{P}(R \mid 3R3B) = \mathbb{P}(R)$, we conclude that these two events are independent (by definition of independence)!

Task 12 – Random Variables

(The material for this problem will be covered on Friday.)

Assume that we roll a fair 3-sided die three times. Here, the sides have values $1, 2, 3$.

a) Describe the PMF of the random variable X giving the sum of the first two rolls.

We have $\mathbb{P}(X = 2) = 1/9$, $\mathbb{P}(X = 3) = 2/9$, $\mathbb{P}(X = 4) = 3/9$, $\mathbb{P}(X = 5) = 2/9$, and $\mathbb{P}(X = 6) = 1/9$.

b) Give the expectation $\mathbb{E}[X]$.

We give a direct proof here, and note that

$$\mathbb{E}[X] = 1/9 \cdot (2 + 6) + 2/9 \cdot (3 + 5) + 3/9 \cdot 4 = (8 + 16 + 12)/9 = 4 .$$

c) Compute $\mathbb{P}(X > 3)$.

$\mathbb{P}(X > 3) = 3/9 + 2/9 + 1/9 = 6/9 = 2/3$.

d) Let Y be the random variable describing the sum of the three rolls. Compute $\mathbb{P}(X = 5 \mid Y = 7)$.

First, $\mathbb{P}(X = 5 \mid Y = 7) = \mathbb{P}(X = 5, Y = 7) / \mathbb{P}(Y = 7)$. Then, $\mathbb{P}(X = 5, Y = 7) = 2/27$, whereas

$$\mathbb{P}(Y = 7) = 3/27 + 2/27 + 1/27 = 2/9 .$$

Task 13 – Hungry Washing Machine

You have 10 pairs of socks (so 20 socks in total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks chosen at random. Every subset of 4 socks is equally probable to be the subset that gets eaten. Let X be the number of complete pairs of socks that you have left.

a) What is the range of X, Ω_X (the set of possible values it can take on)? What is the probability mass function of X?

The washing machine eats 4 socks every time. It can either eat a single sock from 4 pairs of socks, leaving us with 6 complete pairs, or a single sock from 2 pairs and a matching pair, leaving us with 7 complete pairs, or 2 pairs of matching socks, leaving us with 8 complete pairs. That is,

$$\Omega_X = \{6, 7, 8\} .$$

We are dealing with a sample space with equally likely outcomes. As such, we can use the formula $\mathbb{P}(E) = \frac{|E|}{|\Omega|}$. We know that $|\Omega| = \binom{20}{4}$ because the washing machine picks a set of 4 socks out of 20 possible socks. To define the pmf of X, we consider each value in the range of X.

- For $k = 6$, we first pick 4 out of 10 pairs of socks from which we will eat a single sock ($\binom{10}{4}$ ways), and for each of these 4 pairs we have two socks to pick from ($\binom{2}{1}^4$ ways). Using the product rule, we get $|X = 6| = \binom{10}{4}2^4$.
- For $k = 7$, we first pick 1 out of 10 pairs of socks to eat in its entirety ($\binom{10}{1}$ ways), and then 2 out of the 9 remaining pairs from which we will eat a single sock ($\binom{9}{2}$ ways), and for each of these 2 pairs we have two socks to pick from ($\binom{2}{1}^2$ ways). Using the product rule, we get $|X = 7| = 10\binom{9}{2}2^2$.
- For $k = 8$, we pick 2 out of 10 pairs of socks to eat ($\binom{10}{2}$ ways). We get $|X = 8| = \binom{10}{2}$.

Thus,

$$p_X(k) = \begin{cases} \frac{\binom{10}{4}2^4}{\binom{20}{4}} & k = 6 \\ \frac{10\binom{9}{2}2^2}{\binom{20}{4}} & k = 7 \\ \frac{\binom{10}{2}}{\binom{20}{4}} & k = 8 \\ 0 & \text{otherwise} \end{cases}$$

b) Find $\mathbb{E}[X]$ from the definition of expectation.

We calculate directly from the formula for expectation:

$$\mathbb{E}[X] = \sum_{k \in \Omega_X} k \cdot p_X(k) = 6 \cdot \frac{\binom{10}{4}2^4}{\binom{20}{4}} + 7 \cdot \frac{10\binom{9}{2}2^2}{\binom{20}{4}} + 8 \cdot \frac{\binom{10}{2}}{\binom{20}{4}} = \boxed{\frac{120}{19}}.$$