Problem 1.

[The Four-Door Deal]

Let’s see what happens when *Let’s Make a Deal* is played with four doors. A prize is hidden behind one of the four doors. Then the contestant picks a door. Next, the host opens an unpicked door that has no prize behind it. The contestant is allowed to stick with their original door or to switch to one of the two unopened, unpicked doors. The contestant wins if their final choice is the door hiding the prize.

Let’s make the same assumptions as in the original problem:

1. The prize is equally likely to be behind each door.
2. The contestant is equally likely to pick each door initially, regardless of the prize’s location.
3. The host is equally likely to reveal each door that does not conceal the prize and was not selected by the player.

Use The Four Step Method to find the following probabilities. The tree diagram may become awkwardly large, in which case just draw enough of it to make its structure clear. Also, indicate the set of outcomes in each of the events below. A numerical probability without a demonstration of the Method is not a satisfactory answer.

(a) Contestant Stu, a sanitation engineer from Trenton, New Jersey, stays with his original door. What is the probability that Stu wins the prize?

Solution. A partial tree diagram is shown below. The remaining subtrees are symmetric to the fully-expanded subtree. The probability that Stu wins the prize is:

\[
\Pr[\text{Stu wins}] = 4 \cdot \left( \frac{1}{48} + \frac{1}{48} + \frac{1}{48} \right) = \frac{1}{4}
\]

We multiply by 4 to account for the four subtrees, of which we’ve only drawn one.

Notice that we expanded the tree out to the third (“door revealed”) level to spell out the outcomes, but in this case we could, in fact, have stopped at the second level (“player’s initial guess”). This follows because the win/lose outcome is determined by the prize location and Stu’s selected door, regardless of what happens after that.

(b) Contestant Zelda, an alien abduction researcher from Helena, Montana, switches to one of the remaining two doors with equal probability. What is the probability that Zelda wins the prize?

Solution. A partial tree diagram is worked out below. The probability that Zelda wins the prize is:
Problem 2.
Suppose there is a system, built by Caltech graduates, with \( n \) components. We know from past experience that any particular component will fail in a given year with probability \( p \). That is, letting \( F_i \) be the event that the \( i \)th component fails within one year, we have

\[
\Pr[F_i] = p
\]

for \( 1 \leq i \leq n \). The system will fail if any one of its components fails. What can we say about the probability that the system will fail within one year?

Let \( F \) be the event that the system fails within one year. Without any additional assumptions, we can’t get an exact answer for \( \Pr[F] \). However, we can give useful upper and lower bounds, namely,

\[
p \leq \Pr[F] \leq np.
\] (1)

We may as well assume \( p < 1/n \), since the upper bound is trivial otherwise. For example, if \( n = 100 \) and \( p = 10^{-5} \), we conclude that there is at most one chance in 1000 of system failure within a year and at least one chance in 100,000.
Let’s model this situation with the sample space $S := \text{pow}([1, n])$ whose outcomes are subsets of positive integers $\leq n$, where $s \in S$ corresponds to the indices of exactly those components that fail within one year. For example, $\{2, 5\}$ is the outcome that the second and fifth components failed within a year and none of the other components failed. So the outcome that the system did not fail corresponds to the empty set, $\emptyset$.

**STAFF NOTE:** Encourage students to begin by stating explicitly what outcomes (subsets of $[1, n]$) are in the event $F_i$ and $F$.

**(a)** Show that the probability that the system fails could be as small as $p$ by describing appropriate probabilities for the outcomes. Make sure to verify that the sum of your outcome probabilities is 1.

**Solution.** According to the description,  

$$F_i = \{ s \in S \mid i \in s \}$$

$$F = \bigcup_{i=1}^{n} F_i = \{ s \in S \mid s \neq \emptyset \}$$

There could be a probability $p$ of system failure if the individual failures always occur together. That is, if

$$\Pr[[1, \ldots, n]] = p, \quad \Pr[\emptyset] = 1 - p.$$ 

and the probability of all other outcomes is zero. Then

$$\Pr[F_i] = \Pr[[1, \ldots, n]] + 0 + 0 + \cdots + 0 = \Pr[[1, \ldots, n]] = p.$$ 

Also, the only outcome with positive probability in $F$ is $\{1, \ldots, n\}$, so also $\Pr[F] = p$, as required.
(b) Show that the probability that the system fails could actually be as large as $np$ by describing appropriate probabilities for the outcomes. Make sure to verify that the sum of your outcome probabilities is 1.

**Solution.** Suppose at most one component ever fails at a time. That is, $\Pr\{i\} = p$ for $1 \leq i \leq n$, $\Pr\{\emptyset\} = 1 - np$, and probability of all other outcomes is zero. The sum of the probabilities of all the outcomes is one, so this is a well-defined probability space. Also, the only outcome in $F_i$ with positive probability is $\{i\}$, so $\Pr[F_i] = \Pr\{\{i\}\} = p$ as required. Finally, $\Pr[F] = np$ because $F$ contains all the $n$ outcomes of the form $\{i\}$.

(c) Prove inequality (1).

**Solution.** $F = \bigcup_{i=1}^{n} F_i$ so

\[
p = \Pr[F_1] \leq \Pr[F] = \Pr\left(\bigcup_{i=1}^{n} F_i\right) \leq \sum_{i=1}^{n} \Pr[F_i] = np.
\]

**Problem 3.**

To determine which of two people gets a prize, a coin is flipped twice. If the flips are a Head and then a Tail, the first player wins. If the flips are a Tail and then a Head, the second player wins. However, if both coins land the same way, the flips don’t count and the whole process starts over.

Assume that on each flip, a Head comes up with probability $p$, regardless of what happened on other flips. Use the four step method to find a simple formula for the probability that the first player wins. What is the probability that neither player wins?

*Hint:* The tree diagram and sample space are infinite, so you’re not going to finish drawing the tree. Try drawing only enough to see a pattern. Summing all the winning outcome probabilities directly is cumbersome. However, a neat trick solves this problem—and many others. Let $s$ be the sum of all winning outcome probabilities in the whole tree. Notice that you can write the sum of all the winning probabilities in certain subtrees as a function of $s$. Use this observation to write an equation in $s$ and then solve.

**Solution.** In the tree diagram below, the small triangles represent subtrees that are themselves complete copies of the whole tree.

Let $s$ equal the sum of all winning probabilities in the whole tree. There are two extra edges with probability $p$ on the path to each outcome in the top subtree. Therefore, the sum of winning probabilities in the upper tree is $p^2 s$. Similarly, the sum of winning probabilities in the lower subtree is $(1 - p)^2 s$. This gives the equation:

\[
s = p^2 s + (1 - p)^2 s + p(1 - p)
\]

The solution to this equation is $s = 1/2$, for all $p$ between 0 and 1.

By symmetry, the probability that the first player loses is 1/2. This means that the event, if any, of flipping forever can only have probability zero.
Formally, the sample space is the (infinite) set of leaves of the tree, namely,

\[ S := \{TT, HH\}^* \cdot \{HT, TH\} \]

where \( \{TT, HH\}^* \) denotes the set of strings formed by concatenating a sequence of HH’s and TT’s. For example,

\[ TTTTHHHHT, HHTTTH, HHHHHHHHHT, HT \in S. \]

For any string \( s \in S \),

\[ \Pr[s] := p^{\#H’s \ in \ s}(1-p)^{\#T’s \ in \ s}. \]

To verify that this defines a probability space, we must show that \( \sum_{s \in S} \Pr[s] = 1 \). The probability that two tosses match is \( p^2 + (1-p)^2 \), and that they don’t match is \( 2p(1-p) \). Thus, the probability of getting matching tosses for the first 2n coins and finally getting differing tosses on the last 2 coins is

\[ \Pr[\text{Succeeding after } 2n + 2 \text{ tosses}] = (p^2 + (1-p)^2)^n(2p(1-p)) \]

Summing over all \( n \), we get

\[
\sum_{s \in S} \Pr[s] = \sum_{n \geq 0} \sum_{|s| = 2n+2} \Pr[s] \\
= \sum_{n \geq 0} (p^2 + (1-p)^2)^n(2p(1-p)) \\
= 2p(1-p) \sum_{n \geq 0} (p^2 + (1-p)^2)^n \\
= 2p(1-p) \frac{1 - (p^2 + (1-p)^2)}{2p - 2p^2} \\
= \frac{2p(1-p)}{2p - 2p^2} = 1.
\]
Problem 4.
Prove the following probabilistic inequality, referred to as the Union Bound.

Let \( A_1, A_2, \ldots, A_n, \ldots \) be events. Then

\[
\Pr \left[ \bigcup_{n \in \mathbb{N}} A_n \right] \leq \sum_{n \in \mathbb{N}} \Pr[A_n].
\]

Hint: Replace the \( A_n \)'s by pairwise disjoint events and use the Sum Rule.

Solution. The trick is to convert the union of the \( A_i \)'s into a union of disjoint events \( E_i \). This is easy to do: just define \( E_i \) to be the new elements that \( A_i \) adds to the union of the earlier \( A_j \)'s. That is, define

\[
E_0 := A_0
\]
\[
E_{n+1} := A_{n+1} - \bigcup_{i=0}^{n} A_i.
\]

So by definition,

\[
E_n \subseteq A_n \tag{2}
\]
\[
\bigcup_{i=0}^{n} A_i = \bigcup_{i=0}^{n} E_i \tag{3}
\]

and \( E_0, E_1, \ldots, E_n, \ldots \) are pairwise disjoint events.

But if all the finite unions are equal, then so is the infinite union, namely,

STAFF NOTE: Discuss: properties of finite unions don’t always hold for infinite unions. For example, any finite union of finite sets is finite, but infinite unions of finite sets certainly need not be finite. So why in this case does the equality of the finite unions imply equality of the infinite unions?

The answer is that if there was an element in one infinite union and not the other, that element must not be in some finite union, so the finite unions would differ.

\[
\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} E_n \tag{4}
\]

Now we have

\[
\Pr \left[ \bigcup_{n \in \mathbb{N}} A_n \right] = \Pr \left[ \bigcup_{n \in \mathbb{N}} E_n \right] \quad \text{(by (4))}
\]
\[
= \sum_{n \in \mathbb{N}} \Pr[E_n] \quad \text{(Sum Rule)}
\]
\[
\leq \sum_{n \in \mathbb{N}} \Pr[A_n] \quad \text{by (2)}.
\]
Solution. Any set \( A \) is the disjoint union of \( A - B \) and \( A \cap B \), so

\[
\Pr[A] = \Pr[A - B] + \Pr[A \cap B]
\]

by the Disjoint Sum Rule.

\[
\Pr[\overline{A}] = 1 - \Pr[A]
\]

(Complement Rule)

Solution. \( \overline{A} := S - A \), so by the Difference Rule

\[
\Pr[\overline{A}] = \Pr[S] - \Pr[A] = 1 - \Pr[A].
\]

Solution. \( A \cup B \) is the disjoint union of \( A \) and \( B - A \) so

\[
\Pr[A \cup B] = \Pr[A] + \Pr[B - A] \quad \text{(Disjoint Sum Rule)}
\]

\[
= \Pr[A] + (\Pr[B] - \Pr[A \cap B]) \quad \text{(Difference Rule)}
\]

\[
\Pr[A \cup B] \leq \Pr[A] + \Pr[B] \quad \text{(2-event Union Bound)}
\]

Solution. This follows immediately from Inclusion-Exclusion and the fact that \( \Pr[A \cap B] \geq 0 \).

Solution.

\[
\Pr[A] = \Pr[B] - (\Pr[B] - \Pr[A])
\]

\[
= \Pr[B] - (\Pr[B] - \Pr[A \cap B]) \quad \text{(since } A = A \cap B) \\
= \Pr[B] - \Pr[B - A] \quad \text{(difference rule)} \\
\leq \Pr[B] \quad \text{(since } \Pr[B - A] \geq 0). 
\]