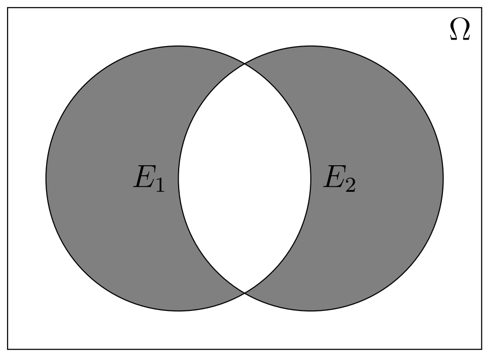
CSE 2331 Homework 4

Please do not write on the question sheet. You must show all work to receive any credit.

1. The symmetric difference is a binary operation on sets. That is, it that takes as input two sets and returns a set. For any pair of sets A and B their symmetric difference, denoted $A \triangle B$, is given by $(A - B) \cup (B - A)$. Let $(\Omega, \Pr)$ be a discrete probability space and let E_1, $E_2 \subset \Omega$ be events.

   (a) Draw a Venn diagram involving Ω, E_1, and E_2 and indicate regions which contribute to the symmetric difference.

   (b) Use the event algebra to express $\Pr[E_1 \triangle E_2]$ in terms of $\Pr[E_1]$, $\Pr[E_2]$, and $\Pr[E_1 \cap E_2]$.
   Note $E_1 - E_2$ and $E_2 - E_1$ are disjoint.

$$\begin{aligned}
\Pr[E_1 \triangle E_2] &= \Pr[(E_1 - E_2) \cup (E_2 - E_1)] \\
&= \Pr[E_1 - E_2] + \Pr[E_2 - E_1] \\
&= \Pr[E_1] - \Pr[E_1 \cap E_2] + \Pr[E_2] - \Pr[E_2 \cap E_1] \\
&= \Pr[E_1] + \Pr[E_2] - 2\Pr[E_1 \cap E_2].
\end{aligned}$$

2. Consider a single roll of a fair 5-sided dice:

   (a) Let $(\Omega, \Pr)$ denote the corresponding (discrete) probability space. Assuming $\Omega = [5]$, write down the function $\Pr : \Omega \to [0, 1]$.

$$\Pr(i) = \frac{1}{5}.$$

   (b) Let E be the event that an odd face is rolled. Write down E in set form.

$$E = \{1, 3, 5\}.$$

   (c) Write down a predicate to encode rolling an even.

$$p(i) = \begin{cases} \mathrm{T} & \text{if } i \in \{2, 4\} \\ \mathrm{F} & \text{otherwise.} \end{cases}$$

(d) Write down the indicator variable corresponding to the predicate for rolling an even.

$$\mathbb{1}_p(i) = \begin{cases} 1 & \text{if } p(i) \text{ (alternatively, } i \in \{2,4\}) \\ 0 & \text{otherwise.} \end{cases}$$

3. Consider a 5-sided biased-dice such the probability of rolling each odd face is p, and the odds of rolling each even face is $(1-3p)/2$.

(a) Write down $(\Omega, \Pr)$ and verify that is a discrete probability space. In other words, check that $\Pr : \Omega \to [0,1]$ sums to what it should.

The discrete probability space is $(\Omega, \Pr)$ where $\Omega = [5]$, (that is, $\Omega = \{1,2,3,4,5\}$) and

$$\Pr(i) = \begin{cases} p & \text{if } i \in \{1,3,5\} \\ (1-3p)/2 & \text{otherwise.} \end{cases}$$

Note that

$$\sum_{i \in \Omega} \Pr(i) = \sum_{i=1}^{5} \Pr(i) = 3p + 2(1-3p)/2 = 3p + 1 - 3p = 1,$$

as required.

(b) Write down an event, O, corresponding to rolling an odd and calculate $\Pr[O]$. Note that O is just the collection of outcomes in Ω consistent with rolling an odd. In this case, $O = \{1,3,5\}$ and

$$\Pr[O] = \sum_{i \in O} \Pr(i) = \Pr(1) + \Pr(3) + \Pr(5) = 3p.$$

(c) Write down an indicator variable $\mathbb{1}_o$ for rolling an odd and compute $\mathbb{E}[\mathbb{1}_o]$.

$$\mathbb{1}_o(i) = \begin{cases} 1 & \text{if } i \in O \\ 0 & \text{otherwise.} \end{cases}$$

Observe,

$$\mathbb{E}[\mathbb{1}_o] = \sum_{i=1}^{5} \mathbb{1}_o(i) \Pr(i) = \sum_{i \in \{1,3,5\}} \Pr(i) = 3p.$$

4. Consider n rolls of a 5-sided biased-dice such the probability of rolling a 1 or a 3 is p (each), and the odds of rolling any other face is $(1-2p)/3$.

(a) Write down $(\Omega, \Pr)$. The elements of Ω should be n-tuples over $[5]$. You do not need to verify that Pr sums to 1.

The discrete probability space is $(\Omega, \Pr)$ where $\Omega = [5]^n$, (i.e., $\Omega = \{1,2,3,4,5\}^n$) and

$$\Pr(e) = p^k((1-2p)/3)^{n-k},$$

where k denotes the combined number of 1s and 3s rolled in event e.

If this is not clear, an equivalent way to express Pr is to let n_1, n_2, n_3, n_4, and n_5 denote the (respective) numbers of 1s through 5s rolled in event e. Then

$$\Pr(e) = ((1-2p)/3)^{n_1} p^{n_2} ((1-2p)/3)^{n_3} p^{n_4} p^{n_5} = p^{n_2+n_4+n_5} ((1-2p)/3)^{n_1+n_3}.$$

(b) Given an event $e \in \Omega$ and $i \in [n]$, let $r_i(e)$ give the outcome of i-th roll.[1] Let X_i be the indicator variable that the i-th roll is an odd. Write a formula for X_i.

$$X_i(e) = \begin{cases} 1 & \text{if } r_i(e) \in \{1,3,5\} \\ 0 & \text{otherwise.} \end{cases}$$

(c) Let X denote the number of odd rolls. What is $\mathbb{E}[X]$? (Hint: How can the indicator variables $\{X_i\}_{i=1}^n$ these be used to express X? Use linearity of expectation.)

Observe that the total number of odds, X, is given by $X = \sum_{i=1}^n X_i$. By linearity of expectation, it follows that:

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i].$$

Now $\mathbb{E}[X_i]$ is the probability that the i-th roll is odd. Since the rolls are independent, this is just the probability of rolling an odd. That is,

$$\mathbb{E}[X_i] = \Pr(1) + \Pr(3) + \Pr(5) = 2p + (1-2p)/3 = \frac{1}{3}(4p+1).$$

Returning to the calculation above,

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \frac{1}{3}(4p+1) = \frac{n}{3}(4p+1).$$

That ends the solution.

When you do something like this you should always try a few values to see if it makes sense. Note that the biggest p can be is $1/2$ where this represents the situation that 1 is rolled with probability $1/2$ and 3 is rolled with probability $1/2$. All other values have probability 0. In this case the expected number of odd rolls should be n and indeed it is since $n/3(4/2+1) = n$. What about when $p = 0$, then we can only roll 2s, 4s, and 5s, which are all equally likely. Therefore we expect $n/3$ of them would be odd, and indeed that's what we get.

[1] For example, if $n = 5$ and $e = (5,2,4,1,1)$ then $r_1(e) = 5$, $r_2(e) = 2$, $r_3(e) = 4$, $r_4(e) = 1$, and $r_5(e) = 1$. In other words, in general, $e = (r_1(e), r_2(e), \ldots, r_n(e))$.

5. As a group we all order meals for our fancy catered lunch. We're a particular bunch so we all have something different. Unfortunately, due to the complexity of our order, our meals arrive late and nothing is labeled. Instead, there are a lot of us, n identical-looking boxes, and little time to sort things out. We decide to stop being picky and just eat whatever box we grab, which we do uniformly at random. In expectation how many people get to eat the meal they ordered?

   Since there are n individuals. For each individual, let X_i be an indicator variable that indicates whether or not they got their own meal. It follows that $X = \sum_{i=1}^{n} X_i$ is a random variable representing the total number of these individuals. By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i].$$

   Thus we need only to determine the odds with which an individual gets their own meal. Since there are n items assigned uniformly at random to n people, the odds that any individual gets any particular item is $1/n$. Therefore, the odds that they get their own is also $1/n$. Thus,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \frac{1}{n} = 1.$$

6. You have m bins, and n balls. Each time you throw a ball it lands in one of the m bins with probability $1/m$. What is the expected number of empty bins? (Hint: Let X_i be the indicator variable that the i-th bin is empty. Let X be the sum of these indicator variables. What is $\mathbb{E}[X]$? Use linearity of expectation.)

   Let X_i be the indicator variable that indicates when the i-th bin is empty. It follows that $X = \sum_{i=1}^{m} X_i$ is a random variable representing the total number of empty bins. By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} \mathbb{E}[X_i].$$

   We know that $\mathbb{E}[X_i]$ is the probability that a given bin is empty. The probability that a single throw ends up in a bin is $1/m$. Therefore, the probability that a single throw is missed is $1 - 1/m$. A bin is empty if all n throws are missed. Therefore, it follows that the bin is empty with probability $(1 - 1/m)^n$.

   Returning to the above calculation,

$$\mathbb{E}[X] = \sum_{i=1}^{m} (1 - 1/m)^n = m\left(1 - \frac{1}{m}\right)^n.$$