18.310 Solution for Homework 1

Due September 11th at 6PM

---

1. **Writing Assignment** Following the discussion in recitation, write up a clear and complete proof of the following theorem. Your proof should be easily understandable by another 18.310 student. Your solution *must* be word processed, using either LaTeX, Word, or some other word processor, and submitted as a pdf file (no Word files allowed). Half of the marks for this question will be allocated to the writing as opposed to simply mathematical correctness. Have a look at the resources linked from the course website for some examples and suggestions. See in particular the notes on "guiding text".

   Theorem 1. *A box contains 7 black balls and 4 white balls. Suppose that we repeatedly draw a ball at random from the box, observe its color and then discard it. We do this 4 times. For* $i \in \{1, 2, 3, 4\}$, *let* X_i *be the random variable representing the color of the* i*th ball drawn from the box. Then,*

   $$\mathbb{P}(\{X_1, X_2\} = \{\text{WHITE}, \text{BLACK}\}) \;=\; \mathbb{P}(\{X_3, X_4\} = \{\text{WHITE}, \text{BLACK}\}).$$

   Solution 1: Consider the set S_{11} of permutations of $[11] = \{1, \ldots, 11\}$. Suppose at first we color the numbers $1, \ldots, 4$ WHITE, and the rest BLACK. Choose $\sigma \in S_{11}$ uniformly at random. We claim that after permuting the elements, we have effectively chosen uniformly at random a subset of $[11]$ of size 4, corresponding to the new positions of the elements originally colored white. Indeed, take A to be an arbitrary set of size 4. Then the probability that this is the set of white balls after picking a random permutation is

   $$\mathbb{P}((\forall i \in A : X_i = \text{BLACK}) \wedge (\forall i \notin S : X_i = \text{WHITE})) = \frac{4!7!}{11!} = \frac{1}{\binom{11}{4}},$$

   independently of what set A (of size 4) is. So we can translate our problem to this new setting, taking X to be a random subset of $[11]$ of size 4.

   $$\mathbb{P}(\{X_1, X_2\} = \{\text{WHITE}, \text{BLACK}\}) = \mathbb{P}(|X \cap \{1, 2\}| = 1) = \mathbb{P}(|X \cap \{3, 4\}| = 1) = \frac{2\binom{9}{3}}{\binom{11}{4}} = \frac{28}{55}.$$

   Solution 2: Let the sample space be the set S_{11} of permutations of $[11]$. Defined a function $c : [11] \to \{\text{WHITE}, \text{BLACK}\}$, giving the color of a ball. We are interested in two subsets:

   $$A_1 = \{x \in S_{11} : \{c(x_1), c(x_2)\} = \{\text{WHITE}, \text{BLACK}\}\}$$
   $$A_2 = \{x \in S_{11} : \{c(x_3), c(x_4)\} = \{\text{WHITE}, \text{BLACK}\}\}$$

   Because all elements of S_{11} are equiprobable, it suffices to show that $|A_1| = |A_2|$ to obtain that $\mathbb{P}(A_1) = \mathbb{P}(A_2)$.

To prove that $|A_1| = |A_2|$, we construct a bijection f between these two sets. The function $f : S_{11} \to S_{11}$ takes a point (a permutation) $x \in S_{11}$ and reorders its entries by swapping x_1 with x_3 and x_2 with x_4 (doing the opposite is also fine). More formally:

$$f : x \mapsto (x_3, x_4, x_1, x_2, x_5, \ldots, x_{11}).$$

It is easy to see that f is invertible, as f is its own inverse:

$$\forall x \in S, f(f(x)) = x.$$

To complete the proof, we need to show that f maps A_1 to A_2 and vice versa that $f^{-1}(A_2) = A_1$, in order to derive that $|A_1| = |A_2|$. To do this, take any $x \in A_1$. Let $y = f(x)$. Then, by the definition of f, $\{y_3, y_4\} = \{x_1, x_2\} = \{\text{WHITE}, \text{BLACK}\}$. Hence, $y \in A_2$, as required. Vice versa, $f^{-1}(A_2) = f(A_2) = A_1$. This completes the proof.

2. **[4 + 3 + 3 = 10 points]** Let X be a uniformly random subset of $\{1, 2, \cdots, n\}$ (there are 2^n possible subsets, and each is chosen with probability $1/2^n$). Let Y be another independently chosen random subset.

   (a) For $i \in \{1, 2, \cdots, n\}$, let A_i be the event that $i \in X$, and similarly let B_i be the event that $i \in Y$. Argue that the $2n$ events $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_n$ are all independent.

Solution: Choose two arbitrary sets of indices $I, J \subseteq [n]$. Then, we can calculate

$$\mathbb{P}((\bigwedge_{i\in I} A_i)\wedge(\bigwedge_{j\in J} \text{BLACK}_j)) = \mathbb{P}((I \subseteq X)\wedge(J \subseteq Y)) = \mathbb{P}(I \subseteq X)\mathbb{P}(J \subseteq Y) = \frac{2^{n-|I|}2^{n-|J|}}{2^n 2^n} = \frac{1}{2^{|I|+|J|}},$$

because of independence, and on the other hand

$$\prod_{i\in I} \mathbb{P}(A_i) \prod_{j\in J} \mathbb{P}(B_j) = \prod_{i\in I} \frac{2^{n-1}}{2^n} \prod_{j\in J} \frac{2^{n-1}}{2^n} = \frac{1}{2^{|I|+|J|}},$$

proving independence.

Also determine:

   (b) the probability that $X \cap Y = \emptyset$,
   Solution: Note that by independence

$$\mathbb{P}(X \cap Y = \emptyset) = \mathbb{P}(\bigwedge_{i\in[n]} (\neg A_i \vee \neg B_i)) = \prod_{i\in[n]} \mathbb{P}(\neg A_i \vee \neg B_i) = \prod_{i\in[n]} (1 - \mathbb{P}(A_i \wedge B_i)),$$

   and again by independence this is equal to

$$\prod_{i\in[n]} (1 - \mathbb{P}(A_i)\mathbb{P}(B_i)) = (3/4)^n.$$

(c) $\mathbb{P}(X \cap Y = \emptyset$ and X has k elements), where k is an integer between 1 and n.
Solution: In this case we can just count all possible cases. After choosing a set of size k, we to get a set Y that doesn't intersect we need to choose some set among the remaining $n-k$ elements:

$$\mathbb{P}(X \cap Y = \emptyset \wedge |X| = k) = \frac{\binom{n}{k} 2^{n-k}}{2^{2n}} = \frac{\binom{n}{k}}{2^{n+k}}.$$

Remark: finding (b) using (c) might not be the simplest solution.

3. **[4 + 6 = 10 points]** A coin is tossed 10 times. This will give a sequence of H's and T's, such as HTTHHTHTTT.

(a) What is the expected number of times TTH appears consecutively in this sequence.
Solution: If interpreted as just the number of times the sequence TTH appears, this is just an application of linearity of expectation. Let X_i, with $i = 1, \ldots, 8$ be the indicator variable for the event that there is a TTH substring starting at i. Then the number of times TTH appears is just the sum of the X_i and by linearity of expectation we get

$$\mathbb{E}(\sum_{i=1}^{8} X_i) = \sum_{i=1}^{8} \mathbb{E}(X_i) = \sum_{i=1}^{8} \mathbb{P}(X_i = 1) = 8\frac{1}{8} = 1.$$

There is also the possibility of interpreting the question as asking for the number of consecutive times TTH appears in the string. e.g. for $TTHHTTH$ would be 2 by the first interpretation and 1 by the second. If they did this, the calculation is trickier, but shouldn't be counted as wrong.

(b) Use the inclusion-exclusion formula to compute the probability that TTH appears (consecutively) at least once in the sequence.
Solution: This is just a calculation:

$$\mathbb{P}(\text{TTH appears at least once}) = \mathbb{P}(\exists i : X_i)$$

We note that there are at most 3 appearances of TTH, so

$$\mathbb{P}(\exists i : X_i = 1) = \sum_i \mathbb{P}(X_i = 1) - \sum_{i<j} \mathbb{P}(X_i = 1 \wedge X_j = 1) + \sum_{i<j<k} \mathbb{P}(X_i = 1 \wedge X_j = 1 \wedge X_k = 1).$$

The first one we already calculated to be 1. For the second term observe that $\mathbb{P}(X_i \wedge X_j) = \frac{1}{2^6}$ as long as $i \leq j - 3$. Since there are 15 such pairs the second term is $\frac{15}{2^6}$. For the third term we do something analogous, and since there are 4 choices of (i, j, k) such that $i \leq j - 3$ and $j \leq k - 3$, we get that the third term is $\frac{4}{2^9}$ so that the required probability is

$$\mathbb{P}(\exists i : X_i = 1) = 1 - \frac{15}{32} + \frac{4}{512}.$$

4. **10 points** Suppose N people enter an elevator in the basement of a building with K floors above it. Assume each person gets off at a floor with probability $1/K$ and independent of any other person. What is the expected number of stops the elevator will make in unloading everyone?

Solution: We can use linearity of expectation again here. Let X_i be the indicator random variable for the event in which someone gets off in floor i. Then the required expectation is that of the sum of the X_i. Since

$$\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(\text{no one gets off in floor i}),$$

and the actions are independent we get

$$\mathbb{P}(X_i = 1) = 1 - \prod_{i=1}^{n} \frac{K-1}{K}$$

so that the required expectation is

$$K\left(1 - \left(\frac{K-1}{K}\right)^N\right).$$

The following exercises should **not** be handed in, but we nevertheless encourage you to do them.

1. (a) We have a box with two black balls and one white ball. We play a game in which we win if we draw the white ball. We draw a ball at random from the box. Before we can see which color it is, the ball is taken away from us and we draw another ball at random from the box. Suppose that we see this second ball and it happens to be black. At this point, we may choose either the first ball we drew or the ball remaining in the box. Which should we choose? **Hint:** Prove that both strategies have the same success probability.
   (b) Consider the same setup, but now the second ball is not drawn at random, but selected to be black (i.e. someone opens the box, finds a black ball and removes it). Do both strategies still have the same success probability? Prove or disprove. **Hint:** What is the relevant subset of the sample space?

2. You roll two dice, each having numbers 1 to 6 on its faces. Consider the following events:
   (a) A: The number on the first die is even.
   (b) B: The number on the second die is 1.

   Find:

   (a) an event C so that A, B and C are 3 independent events.
   (b) an event D so that A, B and D are 3 events that are pairwise independent but not independent.