CS 255: Homework 3

Due Friday December 1 at 11:59 pm

Instructions:

1. You may work in groups of up to 3 students.
2. Each group makes one submission on Canvas.
3. Include the name and SJSU ids of all students in the group.
4. You may discuss high level concepts with other groups, but do not copy other's work.

Problem 1

Consider the following algorithm for shuffling a deck of n cards, initially numbered in order from 1 on the top to n on the bottom. At each step, we remove the top card from the deck and insert it randomly back into the deck, choosing one of the n possible positions uniformly at random. The algorithm ends immediately after we pick up card $n-1$ and insert it randomly into the deck.

(a) Prove that this algorithm uniformly shuffles the deck, so that each permutation of the deck has equal probability. [Hint: Prove that at all times, the cards below card $n-1$ are uniformly shuffled.]

Solution. It is convenient to label card positions starting from the bottom of the deck. That is, position 1 is the bottom card and position n is the top.

Following the hint, we show by induction that all cards below card $n-1$ are uniformly shuffled. Initially, only one card, card n, is below $n-1$. A single card is uniformly shuffled. Now suppose that for some $k \in \{2 \ldots n\}$, card $n-1$ is at position k and the $k-1$ cards below card $n-1$ (in positions $1 \ldots k-1$) are uniformly shuffled. Observe that these cards remain uniformly shuffled until the current top card, say x, is inserted into a position in $\{1 \ldots k\}$. Card x appears at any of these positions with probability $1/k$. Now consider any card y that was one of the $k-1$ cards below card $n-1$ before x is inserted. We consider the probability y is in position $j \in \{1 \ldots k\}$ after x is inserted. We will refer to the beginning position as the location before x is inserted, and the end position as the location after x is inserted.

Card y ends in position k if it begins at $k-1$ and x is inserted at any position in $\{1 \ldots k-1\}$. By the induction hypothesis, the probability y begins at position $k-1$ is $1/(k-1)$. Therefore,

$$\begin{aligned} P(y \text{ ends at position } k) &= P(y \text{ begins at } k-1) P(x \text{ inserted in } 1 \ldots k-1) \\ &= \frac{1}{k-1} \times \frac{k-1}{k} = \frac{1}{k}. \end{aligned}$$

Similarly, card y ends in position 1 if it begins at position 1 and x is inserted at any position in $\{2 \dots k\}$. By the induction hypothesis, the probability y begins at position 1 is $1/(k-1)$.

$$\begin{aligned} P(y \text{ ends at position } 1) &= P(y \text{ begins at } 1)P(x \text{ inserted in } 2 \dots k) \\ &= \frac{1}{k-1} \times \frac{k-1}{k} = \frac{1}{k}. \end{aligned}$$

Finally, y ends at position j if: (i) y begins at $j-1$ and x is inserted in $\{1 \dots j-1\}$:

$$P(y \text{ begins at } j-1)P(x \text{ inserted in } 1 \dots j-1) = \frac{1}{k-1} \times \frac{j-1}{k},$$

or (ii) y begins at j and x is inserted in $\{j+1 \dots k\}$:

$$P(y \text{ begins at } j)P(x \text{ inserted in } j+1 \dots k) = \frac{1}{k-1} \times \frac{k-(j+1)+1}{k}.$$

The sum of these probabilities is the probability y ends at position j:

$$P(y \text{ ends at } j) = \frac{1}{k-1} \times \frac{j-1}{k} + \frac{1}{k-1} \times \frac{k-(j+1)+1}{k} = \frac{k-1}{k(k-1)} = \frac{1}{k}.$$

The probability card y ends at position $j \in [1 \dots k]$ is $1/k$. Therefore, the cards below card $n-1$ are uniformly shuffled.

(b) What is the expected number of steps before this algorithm ends?

Solution. Suppose that card $n-1$ is currently in position $k \in \{2 \dots n\}$. If the top card is inserted into any position $\{1 \dots k\}$, then card $n-1$ gets closer to the top and we make progress towards shuffling the deck. Otherwise, we fail to make progress. If card $n-1$ is at position k, then we make progress with probability k/n.

Let $X_k = \#$ of times the top card is inserted while card $n-1$ is in position k. The X_k's are geometric random variables with parameter $p = k/n$ and mean $E[X_k] = 1/p = n/k$. Let $X = \sum_{k=2}^{n} X_k$, be the total number times the top card is inserted into the deck. Using linearity of expectation,

$$E[X] = \sum_{k=2}^{n} E[X_k] = \sum_{k=2}^{n} \frac{n}{k} = n(H_n - 1),$$

where H_n is the nth harmonic number.

Problem 2

Consider a random walk on a path with vertices numbered $1, 2, \dots, n$ from left to right. At each step, we flip a fair coin to decide which direction to walk, moving one step left or one step right with equal probability. The random walk ends when we fall off one end of the path, either by moving left from vertex 1 or by moving right from vertex n.

(a) Prove that if we start at vertex 1, the probability that the walk ends by falling off the *right* end of the path is exactly $1/(n+1)$.

Solution: Let p_i = probability the walk ends by falling off the *right* end of the path, given that it starts from vertex $i \in [0 \ldots n+1]$. Here, we define the base cases, $p_0 = 0$ and $p_{n+1} = 1$ where walk has ended by falling of the left and right ends of the path respectively.

By conditioning on the outcome of the first flip, we see that for all $i \in [1 \ldots n]$:

$$p_i = 0.5 p_{i-1} + 0.5 p_{i+1}.$$

Using $p_i = 0.5p_i + 0.5p_i$ in the above, we find that

$$p_{i+1} - p_i = p_i - p_{i-1}.$$

Now, since $p_0 = 0$, we obtain:

$$\begin{aligned}
p_2 - p_1 &= p_1 \\
p_3 - p_2 &= p_2 - p_1 = p_1 \\
p_4 - p_3 &= p_3 - p_2 = p_1 \\
&\vdots \\
p_{i+1} - p_i &= p_i - p_{i-1} = p_1.
\end{aligned}$$

Adding the first $i-1$ of these equations yields:

$$p_i - p_1 = \sum_{j=1}^{i-1} p_{j+1} - p_j = \sum_{j=1}^{i-1} p_1,$$

or

$$p_i = i \cdot p_1. \tag{1}$$

Now, since $p_{n+1} = 1$, we get

$$p_1 = \frac{1}{n+1}.$$

(b) Prove that if we start at vertex k, the probability that the walk ends by falling off the *right* end of the path is exactly $k/(n+1)$.

Solution: Using (1) and $p_1 = 1/(n+1)$, we find

$$p_k = \frac{k}{n+1}.$$

(c) Prove that if we start at vertex 1, the expected number of steps before the random walk ends is exactly n.

Solution: Let x_i = expected number of steps until the end of the walk starting from vertex i. The base cases are $x_0 = x_{n+1} = 0$, where the walk has already ended.

Conditioning on the first flip, for any $i \in [1 \ldots n]$:

$$x_i = 0.5(1 + x_{i-1}) + 0.5(1 + x_{i+1}).$$

Using $x_i = 0.5x_i + 0.5x_i$, and rearranging yields

$$x_{i+1} - x_i = x_i - x_{i-1} - 2.$$

Since $x_0 = 0$, it follows that

$$\begin{aligned} x_2 - x_1 &= x_1 - 2 \\ x_3 - x_2 &= x_2 - x_1 - 2 = x_1 - 4 \\ x_4 - x_3 &= x_3 - x_2 - 2 = x_1 - 6 \\ &\vdots \\ x_{i+1} - x_i &= x_i - x_{i-1} - 2 = x_1 - 2i \end{aligned}$$

Summing the first i equations gives

$$x_{i+1} - x_1 = ix_1 - \sum_{j=1}^{i} 2j,$$

or

$$x_{i+1} = (i+1)x_1 - 2\frac{i(i+1)}{2} = (i+1)(x_1 - i). \tag{2}$$

Since $x_{n+1} = 0$, then $x_1 = n$.

(d) What is the *exact* expected length of the random walk if we start at vertex k, as a function of n and k? Prove your result is correct.

Solution: From (2) and $x_1 = n$, we get

$$x_k = k(n - k + 1).$$

Problem 3

Suppose we are given a coin that may or may not be biased, and we would like to compute an accurate *estimate* of the probability of heads. Specifically, if the actual unknown probability of heads is p, we would like to compute an estimate $\hat{p}$ such that

$$P(|\hat{p} - p| > \epsilon) \leq \delta,$$

where ϵ is a given accuracy parameter, and δ is a given confidence parameter.

The following algorithm is a natural first attempt; here Flip() returns the result of an independent flip of the coin.

```
**MeanEstimate**($\epsilon$)**:**
    $count \leftarrow 0$
    **for** $i \leftarrow 0$ **to** $N$ **do**
        **if** $Flip() = Heads$ **then**
            $count \leftarrow count + 1$
    **return** $count/N$
```

(a) Let $\hat{p}$ denote the estimate returned by MeanEstimate(ϵ). Prove that $E[\hat{p}] = p$.

(b) Prove that if we set $N = \lceil \alpha/\epsilon^2 \rceil$ for some appropriate constant α, then we have $P(|\hat{p} - p| > \epsilon) < 1/4$. [Hint: use Chebyshev's Inequality.]

(c) We can increase the previous estimator's confidence by running it multiple times, independently, and returning the *median* of the estimates.

```
**MedianOfMeansEstimate**($\delta, \epsilon$)**:**
    **for** $j \leftarrow 1$ **to** $K$ **do**
        $estimate[j] \leftarrow$ MeanEstimate($\epsilon$)
    **return** Median($estimate[1 \dots K]$)
```

Let p^* denote the estimate returned by MedianOfMeansEstimate(δ, ϵ). Prove that if we set $N = \lceil \alpha/\epsilon^2 \rceil$ (inside MeanEstimate) and $K = \lceil \beta \ln(1/\delta) \rceil$, for some appropriate constants α and β, then $P(|p^* - p| > \epsilon) < \delta$. [Hint: use Chernoff bounds.]

Solution:
(a) Define the indicator random variables: $X_i = 1$ if the ith flip is heads, 0 otherwise. Then, $X = \sum_{i=1}^{N} X_i$ is the total number of heads. The value returned by MeanEstimate is: $\hat{p} = X/N$. Each X_i is a Bernoulli random variable with probability of success p and mean $E[X_i] = p$. By linearity of expectation:

$$E[\hat{p}] = E[X/N] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \frac{1}{N}\sum_{i=1}^{N} p = p.$$

(b) The variance of a Bernoulli random variable with parameter p is $Var[X_i] = p(1-p)$. Since the coin flips are independent, the X_i's are independent. It follows that:

$$Var[\hat{p}] = Var[X/N] = \frac{1}{N^2}\sum_{i=1}^{N} Var[X_i] = \frac{p(1-p)}{N}.$$

Using the result of part (a) and Chebyshev's inequality:

$$P(|\hat{p} - p| > \epsilon) \le \frac{Var[\hat{p}]}{\epsilon^2} = \frac{p(1-p)}{N\epsilon^2}.$$

The simplest bound for the above is $p(1-p) \le 1$, for any $p \in [0,1]$. In this case, choosing $N = \lceil 4/\epsilon^2 \rceil$, yields:

$$P(|\hat{p} - p| > \epsilon) \le \frac{1}{4}.$$

We can get a slightly better bound by observing that function $f(x) = x(1-x)$ is maximized at $x^* = 1/2$ where $f(x^*) = 1/4$. This shows that we only need $N = \lceil 1/\epsilon^2 \rceil$.

(c) We call each run of MeanEstimate a trial. For any $\epsilon > 0$, we say that a trial succeeds if $P(|\hat{p} - p| \leq \epsilon)$, otherwise it fails. Let $Y_j = 1$ if trial j fails, 0 otherwise. Then $Y = \sum_{j=1}^{K} Y_j$ is the number of failed trials. By part (b), if $N = \lceil 4/\epsilon^2 \rceil$ then $E[Y] \leq K/4$.

Observe that the MedianOfMeansEstimate exceeds ϵ if and only if at least $K/2$ trials fail. Recall that we can use an upper bound on the mean in a Chernoff bound. Using the bound $K/4 \geq E[Y]$ from part (b) in the simplified Chernoff bound:

$$P(Y \geq K/2) = P(Y \geq (1+1) \cdot K/4) \leq e^{-\frac{1^2 K/4}{3}} = e^{-K/12},$$

as long as $N = \lceil 4/\epsilon^2 \rceil$. Choosing $K = \lceil 12 \ln(1/\delta) \rceil$, yields:

$$P(Y \geq K/2) \leq \delta.$$

Therefore, if $N = \lceil 4/\epsilon^2 \rceil$ and $K = \lceil 12 \ln(1/\delta) \rceil$, then the MedianOfMeansEstimate satisfies $|\hat{p} - p| \leq \epsilon$, with probability at least $1 - \delta$.