Randomized Algorithms I, Spring 2018, Department of Computer Science, University of Helsinki
Homework 1: Solutions (Discussed January 25, 2018)

---

1. **Exercise 1.6:** Consider the following balls-and-bin game. We start with one black ball and one white ball in a bin. We repeatedly do the following: choose one ball from the bin uniformly at random, and then put the ball back in the bin with another ball of the same color. We repeat until there are n balls in the bin. Show that the number of white balls is equally likely to be any number between 1 and $n-1$.

Let W_n be the number of white balls when the total number of balls is n. We need to show that for all $n \geq 2$ we have

$$\Pr(W_n = k) = 1/(n-1), \qquad k = 1, \dots, n-1.$$

We prove this by induction. The base case is that $n = 2$, in which case the claim is clearly true. Now assume that the claim is true for some $n \geq 2$. For $k = 1, \dots, n$, the event $W_{n+1} = k$ can take place in two ways:

- $W_n = k$ and a black ball is chosen. (Possible only if $k < n$.)
- $W_n = k - 1$ and a white ball is chosen. (Possible only if $k > 1$.)

First note that given W_n the probability of choosing a white ball is W_n/n. Now using this and the definition of conditional probability we get the probability of the first case:

$$\Pr(\text{black} \cap W_n = k) = \Pr(\text{black} \mid W_n = k) \Pr(W_n = k) = \left(1 - \frac{k}{n}\right) \frac{1}{n-1} = \frac{n-k}{n(n-1)}$$

Note that in the special case $k = n$, the above derivation is not valid but the resulting formula still gives the right answer, namely $\Pr(\text{black} \cap W_n = n) = 0$.

Similarly we get the probability of the second case:

$$\Pr(\text{white} \cap W_n = k-1) = \Pr(\text{white} \mid W_n = k-1) \Pr(W_n = k-1) = \frac{k-1}{n} \frac{1}{n-1} = \frac{k-1}{n(n-1)}$$

Again the formula gives the right ansver for the special case $k = 1$.

Since the two cases are disjoint events, we get

$$\Pr(W_{n+1} = k) = \frac{n-k}{n(n-1)} + \frac{k-1}{n(n-1)} = \frac{1}{n},$$

Thus the claim is also true for $n+1$ and by induction principle it holds for all n.

---

2. **Exercise 1.15:** Suppose that we roll ten standard six-sided dice. What is the probability that their sum will be divisible by 6, assuming that the rolls are independent?

First note that the rolls are mutually independent. Let X_i be the result of ith roll and $S_n = \sum_{i=1}^{n} X_i$. The sum of the rolls is divisible by 6 if $S_{10} \equiv 0 \pmod 6$, where $S_{10} = X_{10} + S_9$. Assume that the results of first 9 rolls are fixed and their sum is $S_9 = s_9$. Now there is only one value of the last die that makes S_{10} divisible by 6, namely $X_{10} = 6 - (s_9 \bmod 6)$. Since each value in $\{1, 2, \dots, 6\}$ has equal probability, the probability of the event is $1/6$. More formally we have

$$\begin{aligned}
\Pr(S_{10} \equiv 0 \pmod 6) &= \Pr(X_{10} + S_9 \equiv 0 \pmod 6) \\
&= \sum_{(x_1,\dots,x_9)\in\{1,\dots,6\}^9} \Pr\left(X_{10} + \sum_{i=1}^{9} x_i \equiv 0 \pmod 6 \quad \cap \quad (X_1,\dots,X_9) = (x_1,\dots,x_9)\right) \\
&= \sum_{(x_1,\dots,x_9)\in\{1,\dots,6\}^9} \Pr\left(X_{10} + \sum_{i=1}^{9} x_i \equiv 0 \pmod 6\right) \Pr\left((X_1,\dots,X_9) = (x_1,\dots,x_9)\right) \\
&= \sum_{(x_1,\dots,x_9)\in\{1,\dots,6\}^9} \frac{1}{6} \Pr\left((X_1,\dots,X_9) = (x_1,\dots,x_9)\right) \\
&= \frac{1}{6},
\end{aligned}$$

where we first used the law of total probability and then used the fact that the rolls are mutually independent.

---

3. **Exercise 2.6:** Suppose that we independently roll two standard six-sided dice. Let X_1 be the number that shows on the first die, X_2 the number on the second die, and X the sum of the number on the two dice.

a) What is $\mathbf{E}[X \mid X_1 \text{ is even}]$?

Using the linearity of expectation:

$$\begin{aligned}
\mathbf{E}[X \mid X_1 \text{ is even}] &= \mathbf{E}[X_1 + X_2 \mid X_1 \text{ is even}] \\
&= \mathbf{E}[X_1 \mid X_1 \text{ is even}] + \mathbf{E}[X_2 \mid X_1 \text{ is even}] \\
&= \mathbf{E}[X_1 \mid X_1 \text{ is even}] + \mathbf{E}[X_2] \\
&= (2\cdot\frac{1}{3} + 4\cdot\frac{1}{3} + 6\cdot\frac{1}{3}) + (1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}) \\
&= 4 + 3\frac{1}{2} = 7\frac{1}{2}
\end{aligned}$$

b) What is $\mathbf{E}[X \mid X_1 = X_2]$?

$$\mathbf{E}[X \mid X_1 = X_2] = \sum_{x=2}^{12} x \cdot \Pr(X = x \mid X_1 = X_2) = \sum_{n=1}^{6} 2n \cdot \frac{1}{6} = 7$$

c) What is $\mathbf{E}[X_1 \mid X = 9]$?

$$\begin{aligned}
E[X_1 \mid X = 9] &= \sum_{x=1}^{6} x \cdot \Pr(X_1 = x \mid X = 9) \\
&= \sum_{x=1}^{6} x \frac{\Pr(X_1 = x \cap X = 9)}{\Pr(X = 9)} \\
&= \sum_{x=1}^{6} x \frac{1/36}{4/36} = \frac{1}{4}(3+4+5+6) = 4\frac{1}{2}
\end{aligned}$$

d) What is $\mathbf{E}[X_1 - X_2 \mid X = k]$ for k in the range $[2,12]$?

Using the linearity of expectation and the fact that X_1 and X_2 have the same distribution:

$$\mathbf{E}[X_1 - X_2 \mid X = k] = \mathbf{E}[X_1 \mid X = k] - \mathbf{E}[X_2 \mid X = k] = 0$$

4. **Exercise 2.16:** Suppose we flip a coin n times to obtain a sequence of flips $X_1, X_2, \ldots, X_n$. A *streak* of flips is a consecutive subsequence of flips that are all the same. For example, if X_3, X_4, and X_5 are all heads, there is a streak of length 3 starting at the third flip. (If X_6 is also heads, then there is also a streak of length 4 starting at the third flip.)

a) Let n be a power of 2. Show that the expected number of streaks of length $\log_2 n + 1$ is $1 - o(1)$.

Denote by $k = \log_2 n + 1$ the lengths of the streaks we are considering. Let Y_i be an indicator variable which gets value 1 if there is a streak of length k starting from the ith flip ($i = 1, \ldots, n-k+1$). Now the total number of streaks of length k is $S = \sum_{i=1}^{n-k+1} Y_i$. Using the linearity of expectation and the fact that $\mathbf{E}[Y_i] = \Pr(Y_i = 1)$, the expectation of S is

$$\mathbf{E}[S] = \sum_{i=1}^{n-k+1} \mathbf{E}[Y_i] = \sum_{i=1}^{n-k+1} \Pr(Y_i = 1).$$

For $Y_i = 1$ to be true, X_i can be either heads or tails but the next $k-1$ flips must be the same as X_i. Assuming that the flips are independent and the coin is fair, the probability that this happens is $\Pr(Y_i = 1) = (1/2)^{k-1}$. By inserting this to the above equation we get

$$\mathbf{E}[S] = \sum_{i=1}^{n-k+1} \left(\frac{1}{2}\right)^{k-1} = \sum_{i=1}^{n-k+1} \left(\frac{1}{2}\right)^{\log_2 n} = \sum_{i=1}^{n-k+1} \frac{1}{n} = 1 - \frac{k-1}{n} = 1 - \frac{\log_2 n}{n},$$

which is $1 - o(1)$ with respect to n.

b) Show that, for sufficient large n, the probability that there is no streak of length at least $\lfloor \log_2 n - 2\log_2 \log_2 n \rfloor$ is less than $1/n$. (Hint: Break the sequence of flips up into disjoint blocks of $\lfloor \log_2 n - 2\log_2 \log_2 n \rfloor$ consecutive flips, and use that the event that one block is a streak is independent of the event that any other block is a streak.)

Let's use notation $k = \lfloor \log_2 n - 2\log_2 \log_2 n \rfloor$. We break the sequence into disjoint blocks of k consecutive flips. There are $\lfloor n/k \rfloor$ such blocks (we ignore possible extra flips). For the sequence of n flips to not contain a streak of k flips (denote this event by A) it is necessary that none of the blocks contains a streak of length k (denote this event by B). Thus we have $\Pr(A) \leq \Pr(B)$.

The probability that a single block does not contain a streak is $1 - (1/2)^{k-1}$. Since the blocks are disjoint (and thus independent), the probability that none of the blocks contains a streak is

$$\Pr(B) = \left(1 - \left(\frac{1}{2}\right)^{k-1}\right)^{\lfloor n/k \rfloor}$$

Now, since $k - 1 \leq \log_2 - 2\log_2 \log_2 n$ and $\lfloor n/k \rfloor \leq n/\log_2 n - 1$, we get

$$\begin{aligned}
\Pr(B) &\leq \left(1 - \left(\frac{1}{2}\right)^{\log_2 - 2\log_2 \log_2 n}\right)^{n/\log_2 n - 1} \\
&= \left(1 - \frac{\log_2^2 n}{n}\right)^{n/\log_2 n - 1} \\
&\leq \left(\exp\left(-\frac{\log_2^2 n}{n}\right)\right)^{n/\log_2 n - 1} \\
&= \exp\left(-\frac{\log_2^2 n}{n}\left(\frac{n}{\log_2 n} - 1\right)\right) \\
&= \exp\left(-\log_2 n\left(1 - \frac{\log_2 n}{n}\right)\right),
\end{aligned}$$

where the second inequality is based on the fact that $1 - x \leq e^{-x}$. Let $g(n) = 1 - \log_2 n/n$. Since $g(4) = 3/4 > 1/\log_2 e$ and since $g(n)$ is increasing for $n >= 4$ (as the derivative $D(g(n)) = n^{-2}(\log_2 n - 1/(\ln 2)) > 0$ when $n \geq 4$), for $n \geq 4$ we have that $g(n) \geq 1/\log_2 e$ and therefore

$$\Pr(B) \le \exp\left(-\frac{\log_2 n}{\log_2 e}\right) = \exp(-\ln n) = \frac{1}{n}.$$

---

5. **Exercise 2.26:** A permutation $\pi : [1,n] \to [1,n]$ can be repsented as a set of cycles as follows. Let there be one vertex for each number i, $i = 1,\ldots,n$. If the permutation maps the number i to the number $\pi(i)$, then a directed arc is drawn from vertex i to vertex $\pi(i)$. This leads to a graph that is a set of disjoint cycles. Notice that some of the cycles could be self-loops. What is the expected number of cycles in a random permutation of n numbers?

Let X_i^k be an indicator variable which is 1 if vertex i belongs to a cycle of length k. Let $Y_k = X_1^k + X_2^k + \ldots + X_n^k$ be the total number of nodes belonging to a k-cycle and let N_k be the number of k-cycles in the graph. By definition we must have $N_k = Y_k/k$. Finally, let N be the total number of cycles in the graph, which is $N = \sum_{k=1}^{n} N_k$.

How many permutations π are there such that $X_i^k = 1$? If vertex i belongs to a k-cycle, then the next vertex $j = \pi(i)$ can not be vertex i, vertex $\pi(j)$ can be neither i nor j, and so on, until the k:th vertex is again i. Thus, we can choose the consecutive $k-1$ vertices on the same cycle in $(n-1)(n-2)\cdots(n-k+1)$ ways. The remaining $n-k$ vertices can be in any order in π, so the number of possibilities is $(n-k)(n-k-1)\cdots 2\cdot 1$. The total number of permutations such that $X_i^k = 1$ therefore $(n-1)!$. As there are $n!$ permutations in total, the probability of such event is $\Pr(X_i^k = 1) = (n-1)!/n! = 1/n$.

Now, by using the linearity of expectations twice, we can get the expecter number or cycles

$$\mathbf{E}[N] = \sum_{k=1}^{n} \mathbf{E}[N_k] = \sum_{k=1}^{n} \frac{1}{k}\mathbf{E}[Y_k] = \sum_{k=1}^{n} \frac{1}{k} \sum_{i=1}^{n} \mathbf{E}[X_i^k] = \sum_{k=1}^{n} \frac{1}{k} \sum_{i=1}^{n} \frac{1}{n} = \sum_{k=1}^{n} \frac{1}{k} = H(n).$$

---

6. **Exercise 2.24:** We roll a standard fair die over and over. What is the expected number of rolls until the first pair of consecutive sixes appears? (*Hint:* The answer is not 36.)

We partition the roll sequence into consecutive subsequences such that each subsequence begins where the previous subsequence ends, continues until the first six is encountered, then contains one extra roll more and ends. For example, a sequence $(6,4,2,4,3,6,2,1,6,6)$ consists of three such subsequences: $(6,4)$, $(2,4,3,6,2)$ and $(1,6,6)$. The first pair of consecutive sixes appears when, for the first time, the last roll of a subsequence is six. Let N be the number subsequences needed before this happens. Let E_i denote the event that last roll of ith subsequence is six. For any single subsequence the probability of this event is $P(E_i) = 1/6$. Clearly events E_i are independent, and thus N is geometrically distributed with parameter $1/6$ and its expectation is $\mathbf{E}[N] = 6$.

Now, let X_i be the length of ith subsequence. The total number of rolls until the first pair of consecutive sixes appears is then $X = \sum_{i=1}^{N} X_i$. By first using Lemma 2.5 and linearity of expectations, we get

$$\mathbf{E}[X] = \sum_{n=1}^{\infty} \Pr(N=n)\mathbf{E}[X \mid N=n] = \sum_{n=1}^{\infty} \Pr(N=n)\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{n=1}^{\infty} \Pr(N=n) \sum_{i=1}^{n} \mathbf{E}\left[X_i\right].$$

On the other hand, we have $X_i = Y_i + 1$, where Y_i is the number of rolls needed until the first six is encountered in ith subsequence. Since the rolls are independent and the probability of getting six is $1/6$ for each roll, Y_i is geometrically distributed with parameter $1/6$. Thus, we have $\mathbf{E}[Y_i] = 6$ and furthermore $\mathbf{E}[X_i] = 6 + 1 = 7$. Plugging that into the above equation, we get

$$\mathbf{E}[X] = \sum_{n=1}^{\infty} \Pr(N=n) \sum_{i=1}^{n} 7 = 7 \cdot \sum_{n=1}^{\infty} n\Pr(N=n) = 7 \cdot \mathbf{E}[N] = 7 \cdot 6 = 42.$$

— — — — — — — — — — —

Another way to approach this is to use the memorylessness property. Let X be the number of rolls until the first pair of consecutive sixes appears and let A_i denote the event that ith roll is six. Events A_i are mutually independent and we have $\Pr(A_i) = 1/6$. Now, since at least two rolls is always required, we can split up the expectation as follows:

$$\begin{aligned}
\mathbf{E}[X] &= \Pr(\bar{A}_1)\mathbf{E}[X \mid \bar{A}_1] + \Pr(A_1)\mathbf{E}[X \mid A_1] \\
&= \Pr(\bar{A}_1)\mathbf{E}[X \mid \bar{A}_1] + \Pr(A_1)\big(\Pr(\bar{A}_2 \mid A_1)\mathbf{E}[X \mid A_1, \bar{A}_2] + \Pr(A_2 \mid A_1)\mathbf{E}[X \mid A_1, A_2]\big) \\
&= \Pr(\bar{A}_1)\mathbf{E}[X \mid \bar{A}_1] + \Pr(A_1)\big(\Pr(\bar{A}_2)\mathbf{E}[X \mid A_1, \bar{A}_2] + \Pr(A_2)\mathbf{E}[X \mid A_1, A_2]\big) \\
&= \frac{5}{6}\mathbf{E}[X \mid \bar{A}_1] + \frac{1}{6}\left(\frac{5}{6}\mathbf{E}[X \mid A_1, \bar{A}_2] + \frac{1}{6}\mathbf{E}[X \mid A_1, A_2]\right).
\end{aligned}$$

Now, as the past rolls do not have any effect on the future, if our previous roll is not six, then the expected number of additional rolls we need is the same as the expected number of rolls in the beginning. That is, we have $\mathbf{E}[X \mid A_1^c] = 1 + \mathbf{E}[X]$ and $\mathbf{E}[X \mid A_1, A_2^c] = 2 + \mathbf{E}[X]$. Also, we have $\mathbf{E}[X \mid A_1, A_2] = 2$, since if A_1 and A_2 are true, then we already have two consecutive sixes. Hence, we have

$$\mathbf{E}[X] = \frac{5}{6}(1 + \mathbf{E}[X]) + \frac{1}{6}\left(\frac{5}{6}(2 + \mathbf{E}[X]) + \frac{1}{6}2\right) = \frac{35}{36}\mathbf{E}[X] + \frac{42}{36} \quad \Leftrightarrow \quad \mathbf{E}[X] = 42.$$

---

7. **Exercise 2.32:** You need a new staff assistant, and you have n people to review. You want to hire the best candidate for the position. When you interview a candidate, you can give them a score, with the highest score being the best and no ties being possible. You interview the candidates one by one. Because of your company's hiring practices, after you interview the kth candidate, you either offer the candidate the job before the next interview or you forever lose the chance to hire that candidate. We suppose the candidates are interviewed in a random order, chosen uniformly at random from all $n!$ possible orderings.
We consider the following strategy. First, interview m candidates but reject them all; these candidates give you an idea of how strong the field is. After the mth candidate, hire the first candidate you interview who is better than all of the previous candidates you have interviewed.

a) Let E be the event that we hire the best assistant, and let E_i be the event that ith candidate is the best and we hire him. Determine $\Pr(E_i)$, and show that

$$\Pr(E) = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}.$$

Since the the order of candidates is chosen uniformly at random, for the best candidate each position is equally likely. Thus, the probability that ith candidate is the best is $1/n$. Now assume, that ith candidate is the best one. If $i \le m$, then the ith canditate can not be hired and $\Pr(E_i) = 0$. Otherwise, the ith candidate is hired if and only if none of the candidates $m+1, m+2, \ldots, i-1$ is better than the best of the first m candidates, in other words, if the best of the first $i-1$ candidates is one of the first m candidates. Since each position is equally likely, the probability that this happens is $m/(i-1)$. Therefore we have:

$$\Pr(E_i) = \begin{cases} \frac{1}{n}\frac{m}{i-1} & \text{if } i > m \\ 0 & \text{if } i \le m. \end{cases}$$

Now, $E = \bigcup_{i=1}^{n} E_i$ and events E_i are disjoint, so the probability of E is

$$\Pr(E) = \sum_{i=1}^{n} \Pr(E_i) = \sum_{i=m+1}^{n} \frac{1}{n}\frac{m}{i-1} = \frac{m}{n} \sum_{i=m+1}^{n} \frac{1}{i-1}.$$

b) Bound $\sum_{j=m+1}^{n} \frac{1}{j-1}$ to obtain

$$\frac{m}{n}(\ln n - \ln m) \le \Pr(E) \le \frac{m}{n}(\ln(n-1) - \ln(m-1)).$$

We use the same method as in Lemma 2.10 to bound the probability $\Pr(E)$. In general, if $f(k)$ is monotinically decreasing, then

$$\int_a^{b+1} f(x)dx \leq \sum_{k=a}^{b} f(k) \leq \int_{a-1}^{b} f(x)dx.$$

Hence, by setting $f(i) = 1/(i-1)$, we get a lower bound

$$\frac{m}{n}\int_{m+1}^{n+1} \frac{dx}{x-1} = \frac{m}{n}(\ln n - \ln m)$$

and an upper bound

$$\frac{m}{n}\int_{m}^{n} \frac{dx}{x-1} = \frac{m}{n}(\ln(n-1) - \ln(m-1)).$$

c) Show that $m(\ln n - \ln m)/n$ is maximized when $m = n/e$, and explain why this means $\Pr(E) \geq 1/e$ for this choise of m.

We find the maximum of a function $f(m) = m(\ln n - \ln m)/n$ by finding where its derivative is zero:

$$f'(m) = \frac{\ln n - \ln m - 1}{n} = 0 \quad \Leftrightarrow \quad \ln m = \ln n - 1 \quad \Leftrightarrow \quad m = n/e$$

To check that this actually is a maximum point, we calculate the second derivative:

$$f''(m) = -\frac{1}{mn}.$$

This is negative for all $m, n > 0$, so f is concave w.r.t. m and thus $m = n/e$ is an actual maximum point. By plugging $m = n/e$ into the lower bound inequation we get

$$\Pr(E) \geq \frac{n}{en}(\ln n - \ln\frac{n}{e}) = \frac{1}{e}.$$