Design of Algorithms - Homework II (Solutions)

K. Subramani
LCSEE,
West Virginia University,
Morgantown, WV
$\{$ksmani@csee.wvu.edu$\}$

1 Problems

1. Each of n customers gives a hat to a hat-check person at a restaurant. The hat-check person gives the hats back to the customers in a random order. What is the expected number of customers that get their own hat back?

   Solution: We solve this problem using indicator random variables. For $1 \leq i \leq n$, let

$$X_i = \mathbf{I}\{\text{Customer } i \text{ gets his own hat back}\}.$$

   Let the random variable X be the number of customers which get their own hats back. We wish to compute $E[X]$.

   Clearly,

$$X = \sum_{i=1}^{n} X_i.$$

   It is easy to see that the probability that customer i receives his own hat is $\frac{1}{n}$. This implies that $E[X_i] = \frac{1}{n}$, by basic properties of indicator random variables (Lemma 5.1 in the text). Thus,

$$\begin{aligned} E[X] &= E\Big[\sum_{i=1}^{n} X_i\Big] && (1)\\ &= \sum_{i=1}^{n} E[X_i] && (2)\\ &= \sum_{i=1}^{n} \frac{1}{n} && (3)\\ &= 1. && (4) \end{aligned}$$

   Thus the expected number of customers which receive their own hats back is 1. Note that Line (2) is linearity of expectation. Also note that linearity of expectation holds *even when the corresponding random variables are not independent*. The X_i's in this problem are not independent, which can easily be seen by considering the case when $n = 2$. In this case, if customer 1 gets his own hat back, then customer 2 must get his own hat back as well — there is no other customer to receive it.

   $\square$

2. Let $\mathbf{A}[1 \cdot \cdot n]$ represent an array of n distinct numbers that have been randomly permuted. If $i < j$ and $A[i] > A[j]$, then the pair (i, j) is called an *inversion* of $\mathbf{A}$. Use indicator random variables to compute the expected number of inversions of $\mathbf{A}$.

Solution: For $1 \le i < j \le n$,

$$X_{ij} = \mathbf{I}\{A[i] > A[j]\}.$$

That is, X_{ij} is the indicator random variable for the event that the pair (i, j), with $i < j$, is inverted. Now, Pr$\{X_{ij} = 1\}$ is equal to $\frac{1}{2}$, because in any random permutation of distinct numbers, there are precisely two possibilities, viz., $A[i] > A[j]$ and $A[j] > A[i]$, with each of them having probability $\frac{1}{2}$. Arguing as we did in problem 1, we can conclude that $E[X_{ij}] = \frac{1}{2}$.

Let X be the random variable denoting the total number of inverted pairs in the array, so the solution to the problem is $E[X]$. Note that X is equal to the sum of all the X_{ij}'s.

The total number of inversions is clearly,

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij} \text{ (Why?)}$$

It follows that the expected number of inversions is:

$$\begin{aligned}
E[X] &= E[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}] \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}] \text{ (Why?)} \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{2} \\
&= \frac{1}{2} \sum_{i=1}^{n-1} (n - (i+1) + 1) \\
&= \frac{1}{2} \sum_{i=1}^{n-1} (n - i) \\
&= \frac{1}{2} [\sum_{i=1}^{n-1} n - \sum_{i=1}^{n-1} i] \\
&= \frac{1}{2} [n \cdot (n-1) - \frac{1}{2} \cdot n \cdot (n-1) \\
&= \frac{1}{2} [\frac{1}{2} \cdot n \cdot (n-1)] \\
&= \frac{n \cdot (n-1)}{4}
\end{aligned}$$

□

3. A miner is trapped in a mine containing three doors. The first door leads to a tunnel that will allow the miner to reach safety in two hours. The second door leads to a tunnel that will bring him back to the mine after three hours. The third door leads to a tunnel that will bring him back to the mine after five hours. Assume that the miner chooses one of the three doors uniformly and at random, whenever he is confronted with the three choices. In how much time can he expect to get to safety?

   Solution: Let X denote the random variable that represents the time taken by the miner to reach safety. Observe that $X \in \{2, \cdots \infty\}$. We are interested in $E[X]$.

   Let Y be the random variable that assumes the values 1, 2 or 3, depending upon whether the miner chooses the first door, the second door, or the third door respectively.

From the discussion in class, we know that

$$E[X] = E[E[X \mid Y]]$$

As per the given information, $E[X \mid Y = 1] = 2$, $E[X \mid Y = 2] = 3 + E[X]$ and $E[X \mid Y = 3] = 5 + E[X]$.

It follows that,

$$\begin{aligned} E[X] &= E[E[X \mid Y]] \\ &= E[X \mid Y = 1] \cdot Pr[Y = 1] + E[X \mid Y = 2] \cdot Pr[Y = 2] + E[X \mid Y = 3] \cdot Pr[Y = 3] \\ &= 2 \cdot \frac{1}{3} + (3 + E[X]) \cdot \frac{1}{3} + (5 + E[X]) \cdot \frac{1}{3} \\ &= \frac{10}{3} + \frac{2}{3} \cdot E[X] \end{aligned}$$

Hence, $\frac{1}{3}E[X] = \frac{10}{3}$ and therefore, $E[X] = 10$. In other words, the miner can expect to spend ten hours in the mine, before he gets to safety.

□

4. Show that

$$\sum_{k=1}^{n-1} k \cdot \log k \leq \frac{1}{2}n^2 \log n - \frac{1}{8}n^2.$$

Solution: First, note that it must be the case that $n \geq 2$. Since $f(k) = k \cdot \log k$ is a monotonically increasing function of k, we know that

$$\sum_{k=1}^{n-1} k \cdot \log k \leq \int_1^n x \cdot \log x \, dx$$

Integrating by parts with $u = \log x$, $du = \frac{1}{x}\frac{1}{\ln 2}dx$, $dv = x\ dx$, and $v = \frac{x^2}{2}$, we have

$$\int_1^n x \cdot \log x\ dx = \frac{x^2}{2} \cdot \log x\Big|_1^n - \int_1^n \frac{x^2}{2} \cdot \frac{1}{x}\frac{1}{\ln 2}\ dx = \frac{n^2}{2} \cdot \log n - \Big(\frac{1}{\ln 2}(\frac{n^2}{4} - \frac{1}{4})\Big).$$

Therefore it suffices to show

$$\frac{1}{8}n^2 \leq \frac{1}{\ln 2}(\frac{n^2}{4} - \frac{1}{4})$$

for $n \geq 2$. It is easy to see that at $n = 2$, the LHS is less than the RHS, and that on $[2, \infty)$, the derivative of the LHS is less than the derivative of the RHS. Hence the inequality holds.

□

5. Show that the second smallest of n elements in an array can be found using at most $n + \lceil \log n \rceil - 2$ element to element comparisons.

Solution: Consider the following algorithm (Algorithm 1.1).

This algorithm returns the smallest element in the whole array min_w and a set S_{2min} of candidate elements for the second minimum element.

The number of comparisons is characterized by the following recurrence relation:

$$\begin{aligned} T(1) &= 0 \\ T(2) &= 1 \\ T(n) &= 2 \cdot T(\frac{n}{2}) + 1 \end{aligned}$$

```
Function FIND-2MIN(A, low, high)
n = high − low + 1
S_2min = φ
if (n = 1) then
    min_w = A[1]
    S_2min = φ.
    return (min_w, S_2min)
end if
if (n = 2) then
    if (A[1] ≤ A[2]) then
        min_w = A[1]
        Add A[2] to S_2min
        return (min_w, S_2min)
    else
        min_w = A[2].
        Add A[1] to S_2min
        return (min_w, S_2min)
    end if
end if
{We know that n ≥ 3}
mid = ⌊(high+low)/2⌋
(lmin_w, lS_2min) = FIND-2MIN(A, low, mid)
(rmin_w, rS_2min) = FIND-2MIN(A, mid + 1, high)
if (lmin_w ≤ rmin_w) then
    min_w = lmin_w
    S_2min = lS_2min ∪ rmin_2
else
    min_w = rmin_w
    S_2min = rS_2min ∪ lmin_2
end if
return (min_w, S_2min)
```

Algorithm 1.1: Finding the two smallest elements in an array

This recurrence is easily solved to get $T(n) = n - 1$. The size of the candidate set S_{2min} can be characterized by the following recurrence:

$$\begin{aligned} G(1) &= 0 \\ G(2) &= 1 \\ G(n) &= 1 + G(\frac{n}{2}) \end{aligned}$$

$G(n)$ is easily seen to be at most $\lceil \log_2 n \rceil$. We can find the smallest element in S_{2min} using at most $\lceil \log_2 n \rceil - 1$ comparisons. It thus follows that the second smallest element can be found in $n + \lceil \log_2 n \rceil - 2$ comparisons.

□