Randomized Algorithms—Problem Set #1—Fall 2021.

1. A number from 10 to 99 (both inclusive) is choose u.a.r. what is the probability this number is divisible by 5?

2. True or false: A randomized algorithm for a decision problem with one-sided-error and correctness probability 1/3 (that is, if the answer is YES, it will always output YES, while if the answer is NO, it will output NO with probability 1/3) can be amplified to a correctness probability of 99%. Justify your answer and if the amplification is possible give the numer of iterations needed to guarantee the the probability of 99% o being correct.

3. Consider the following algorithm to generate an integer $r \in \{1, \ldots, n\}$: We have n coins labelled $m_1, \ldots, m_n$, where the probability that $m_i =$ head is $1/i$. Toss the in order the coins $m_n, m_{n-1}, \ldots$ until getting the first head, if the fist head appears with coin m_i, the $r = i$. Prove that the previous algorithm yield an integer r with uniform distribution. i.e. the probability of getting any integer r is $1/n$.

4. Given a text $T = x_1 x_2, \cdots x_n$ (w.l.o.g in binary) and a pattern $S = s_1 \cdots s_m$, with $m << n$ and both chains over the same alphabet $\Sigma = \{0, 1\}$, we want to determine if S occurs as a contiguous substring of T. For ex., if $T = 10110110010101110$ and $S = 1011$ then 101 $\boxed{1011}$ 0010 $\boxed{1011}$ 10. The standard greedy takes $O(nm)$ steps. There are deterministic algorithms that work in worst-time $O(n + m)$ (Knuth-Morris-Pratt), but they are complicated, difficult to implement and with large implementation constants. The following simple probabilistic algorithm does the job using the fingerprint technique, with a small probability of error. The algorithm computes the fingerprint of S and compares with the fingerprints of successive sliding substrings of T, i.e. with $T(j) = x_j \cdots x_{j+m-1}$, for $1 \leq j \leq n - m + 1$.

   Matching (S, T)
   Express S as an integer $D(S) = \sum_{i=0}^{m-1} s_{i+1} 2^i$
   so $D(S)$ is a m-bit integer
   Choose a prime $p \in [2, \ldots, k]$,
   where $k = cmn \ln(cmn)$, for suitable $c > 1$
   Compute $\phi(S) = D(S) \mod p$
   for $j = 1$ to $n - m + 1$ **do**
   Compute $D(T(j)) = \sum_{i=0}^{m-1} x_{j+i} 2^i$
   Compute $\phi(T(j)) = D(T(j)) \mod p$
   if $\phi(T(j)) = \phi(S)$ **then**
   output match at position j
   endif
   endfor

Prove,

(a) This algorithm is one-sided-error, it may output match when there is no match. Prove the $\mathbf{Pr}\left[\textit{output match, when no match}\right] \leq 1/c$, for suitable $c > 0$.

(b) Prove that the algorithm can be implemented in $O(n + m)$ steps. Notice that computing $D(T(j))$ at every iteration costs $\Theta(m)$ if done naïvely, leading to an algorithm of cost $\Theta(n \cdot m)$.

5. (a) Suppose that we roll twice a fair k-sided die with the numbers 1 through k on the die's faces, obtaining values X_1 and X_2. What is $\mathbf{E}\left[\max(X_1, X_2)\right]$? What is $\mathbf{E}\left[\min(X_1, X_2)\right]$?

   (b) Show from your calculation in part (a) that $\mathbf{E}\left[\max(X_1, X_2)\right] + \mathbf{E}\left[\min(X_1, X_2)\right] = \mathbf{E}\left[X_1\right] + \mathbf{E}\left[X_2\right]$

   (c) Explain why the equation in part (b) must be true by using the linearity of expectations instead of a direct computation.

6. Median-of-3 quicksort is a variant of quicksort which picks uniformly at random and without replacement three elements from the array and uses the median of these three elements as the pivot for each recursive stage in which we have to sort $n \geq 3$ elements. So the probability $\pi_{n,j}$ that the pivot choosen is the j-th smallest element is not uniform, while

$$\pi_{n,j} = \frac{1}{n}, \qquad \text{for all } j,\ 1 \leq j \leq n,$$

in ordinary randomized quicksort, when the pivot of each stage is choosen u.a.r.

When analyzing the expected cost of ordinary quicksort we set up a recurrence for the expected number $q_n = \mathbf{E}\left[Q_n\right]$ of comparisons as follows:

$$\begin{aligned} q_1 &= q_0 = 0 \\ q_n &= n - 1 + \sum_{j=1}^{n} \pi_{n,j} \times \mathbf{E}\left[\#\text{ of comparisons} \mid \text{pivot is the } j\text{th element}\right] \\ &= n - 1 + \sum_{j=1}^{n} \pi_{n,j}\left(q_{j-1} + q_{n-j}\right) = n - 1 + \frac{2}{n}\sum_{j=0}^{n-1} q_j. \end{aligned}$$

Similar steps can be applied in the case of median-of-3 quicksort, but the so called *splitting probabilities* $\pi_{n,j}$ are different.

(a) Calculate the probability $\pi_{n,j}$ that the pivot is the j-th smallest element, whne it is choosen as the median of three random elements slected without replacement from the array of n elements.

(b) Set up the recurrence for the expected number $q_n^{(3)}$ of comparisons for quicksort with median-of-3.

(c) Identify a shape function for the recurrence. Apply the continuous master theorem to solve the recurrence. Show that $q_n^{(3)} = \frac{12}{7} n \ln n + o(n \log n)$ (for ordinary quicksort we have $q_n = 2n \ln n + o(n \log n)$, about a 15% more comparisons on average).