Mathematics for Computer Science
Spring 2019
Due: 23:59, March 18, 2019

Homework Set 3

Reading Assignments: Read Chapter 5, and Chapter 6.1 – 6.7 of LPV.

Written Assignments: Do the following exercises from LPV: 5.4.3, 5.4.4, 5.4.5, 6.3.5, 6.5.2, 6.5.4.

Special Problem 1 (counted as 2 exercises) Alice and Bob each independently tosses an unbiased coin n times. Let X and Y be the random variables corresponding to the number of HEADs in Alice' and Bob's results. Your solutions must be *closed-form* formulas in n.

(a) Determine the expected value of the random variable $X - Y$.

(b) Determine the variance of the random variable $X - Y$.

(c) Let S denote the event that $X = Y$, and let $s(n) = \Pr\{S\}$. Determine $s(n)$.

(d) Let T denote the event that $X = Y + 1$, and let $t(n) = \Pr\{T\}$. Determine $t(n)$.

Remarks A *closed-form formula* has only a bounded number of terms involving familiar functions. For example, $\log_2 x + x^5$ is a closed-form formula in x, $n^3 + \binom{n^2}{n}/(n+6) - \cos(1/n)$ is a closed-form formula in n, while $\sum_{1\le i\le n} i^2$ is *not* a closed-form formula in n. For your information, $s(1) = 1/2$, $s(2) = 3/8$, $t(1) = 1/4$, $t(2) = 1/4$.

Special Problem 2 (counted as 3 exercises) Toward the end of last class, we were analyzing the expected running time of a program that finds the minimum of an array $x[1:n]$. Assume the content of the array is a (uniformly chosen) random permutation. The program scans the array from left to right, updating the current value of minimum (stored as variable z). The analysis focuses on the random variable Y_n, defined as the number of times the updating-z line is executed. It is clear that Y_n is the number of *left-to-right minima* for a random permutation of $\{1, 2, \cdots, n\}$.

(a) Prove that $E(Y_n) = H_n$, where $H_n = \sum_{1\le j\le n} \frac{1}{j}$.

(b) Prove that $Var(Y_n) = H_n - H_n^{(2)}$, where $H_n^{(2)} = \sum_{1\le j\le n} \frac{1}{j^2}$.

(c) Prove that $\lim_{n\to\infty} \frac{Var(Y_n)}{H_n} = 1$. (Thus, as the standard deviation $\sigma(Y_n) \approx \sqrt{E(Y_n)}$ for large n, the random variable Y_n takes on value typically not too far from $E(Y_n)$.)

Special Problem 3 (counted as 4 exercises) Consider the *tree-cutting* problem for a rooted tree T of n nodes, as discussed in class. Let X_T be the random variable denoting the number of *cutting moves*. It was shown that $E(X_T) = \sum_{v\in T} \frac{1}{1+\text{depth}(v)}$.

(a) What is the tree T with the smallest value of $E(X_T)$?

(b) What is the tree T with the largest value of $E(X_T)$?

(c) An expression for $Var(X_T)$ was given in class. Give a rigorous derivation of that formula. (The argument discussed in class was not accurate, as pointed out by a student after class. Try to derive it using only the concepts we have already introduced in class; so try not to use the concept of *conditional probability* which could lead to troubles if not careful.)

(d) Determine, as closed-form formulas, for $E(X_T), Var(X_T)$ when T is a *chain* of length n. (You may consider H_n and $H_n^{(2)}$ as familiar functions.) In this case, is the value of X_T clustered around its expected value?