Homework 3 Solutions

Note: These solutions are not necessarily model answers. Rather, they are designed to be tutorial in nature, and sometimes contain a little more explanation than an ideal solution. Also, bear in mind that there may be more than one correct solution. The maximum total number of points available is 35.

1. First, note that *5pts*

$$\begin{aligned} E[X \mid T] &= E[X_1 + \cdots + X_T \mid T] \\ &= E[X_1 \mid T] + \cdots + E[X_T \mid T] \quad \text{by linearity of expectation} \\ &= E[X_1] + \cdots + E[X_T] \quad \text{because the } X_i\text{'s are independent of } T \\ &= \mu T. \end{aligned}$$

Now, $E[X] = E[E[X \mid T]] = E[\mu T] = \mu E[T]$, which completes the proof. Note the use of independence in the third line above, which allows us to get rid of the conditioning.

2. (a) For each $i = 1, 2, \ldots, n$, let Y_i denote the indicator r.v. that equals 1 if i lies in a cycle of length 2, and 0 otherwise. Then, $E[Y_i] = \Pr[Y_i = 1] = \Pr[\pi(i) \neq i \wedge \pi(\pi(i)) = i] = \frac{n-1}{n} \cdot \frac{1}{n-1} = \frac{1}{n}$. Next, observe that $X = \frac{1}{2}(Y_1 + Y_2 + \cdots + Y_n)$. This is because every cycle of length 2 contributes a total of 2 to the sum $Y_1 + Y_2 + \cdots + Y_n$. It follows by linearity of expectation that $E[X] = \frac{1}{2} \cdot n \cdot \frac{1}{n} = \frac{1}{2}$. *3pts*

(b) To compute $E[X^2]$, we write X in terms of the Y_i as in part (a) and expand out the square to get $X^2 = \frac{1}{4}\left(\sum_i Y_i^2 + \sum_{i \neq j} Y_i Y_j\right)$. To take the expectation of the diagonal terms Y_i^2, note that Y_i is a 0-1 random variable, so $Y_i^2 = Y_i$ and thus $E[Y_i^2] = E[Y_i] = \frac{1}{n}$. For the cross terms $Y_i Y_j$ with $i \neq j$, there are two ways in which $Y_i = Y_j = 1$: *7pts*

- i, j are in the same 2-cycle. This probability of this is $\Pr[\pi(i) = j \wedge \pi(j) = i] = \Pr[\pi(i) = j] \cdot \Pr[\pi(j) = i \mid \pi(i) = j] = \frac{1}{n} \cdot \frac{1}{n-1} = \frac{1}{n(n-1)}$.
- i, j are in disjoint 2-cycles. There are $(n-2)(n-3)$ ways to pick i', j' such that i, j, i', j' are distinct. Then,

$$\Pr[\pi(i) = i' \wedge \pi(i') = i] = \frac{1}{n(n-1)}; \quad \text{and}$$

$$\Pr[\pi(j) = j' \wedge \pi(j') = j \mid \pi(i) = i' \wedge \pi(i') = i] = \frac{1}{(n-2)(n-3)}$$

Therefore, this case also occurs with probability $(n-2)(n-3) \cdot \frac{1}{n(n-1)} \cdot \frac{1}{(n-2)(n-3)} = \frac{1}{n(n-1)}$.

Putting the two cases together, we get $E[Y_i Y_j] = \Pr[Y_i = 1 \wedge Y_j = 1] = \frac{2}{n(n-1)}$. Combining this with the diagonal terms gives

$$E[X^2] = \frac{1}{4}\left(n \cdot \frac{1}{n} + n(n-1) \cdot \frac{2}{n(n-1)}\right) = \frac{3}{4}.$$

Finally, from part (a) we have $E[X]^2 = \frac{1}{4}$, so $\mathrm{Var}[X] = E[X^2] - E[X]^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.

Note: Many people missed Q2(b). Please read the above calculation carefully and make sure that you understand it! If not, please ask.

3. Define the r.v. $X = \frac{X_1+X_2+\cdots+X_n}{n}$. Then by linearity of expectation we have $E[X] = \mu$. Also, since the X_i's are independent, we have $\mathrm{Var}[X] = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}[X_i] = \frac{\sigma^2}{n}$. Finally, by Chebyshev's inequality, for any $\epsilon > 0$, *5pts*

$$\Pr[|X-\mu| \geq \epsilon] \leq \frac{\sigma^2}{n\epsilon^2} \to 0$$

as $n \to \infty$. This completes the proof.

4. (a) For each $e \in E$, let X_e be the indicator random variable that assumes the value 1 if e is in the cut, and 0 otherwise. Then $X = \sum_{e\in E} X_e$. In addition, $E[X_e] = \Pr[\text{endpoints of } e \text{ have different colors}] = \frac{1}{2}$. By linearity of expectation, we have $E[X] = \frac{1}{2}|E| \geq \frac{\text{OPT}}{2}$, since clearly $\text{OPT} \leq |E|$. *3pts*

Note: Many people did not understand the relationship between $|E|$ and OPT *(both in this part and in the rest of this problem). The important points to remember are: (i)* $\text{OPT} \leq |E|$ *(because clearly no cut can contain more edges than the total number of edges in the graph!); and (ii) we do not know the value of* OPT*, so we cannot use it in our algorithm (though of course we do know the value of $|E|$). Bear these points in mind when you read the solution to this problem.*

(b) Let $Y = |E| - X$, which is a non-negative random variable. Note that $E[Y] = |E| - E[X] = \frac{1}{2}|E|$. Now, $\Pr[X < 0.49|E|] = \Pr[Y > 0.51|E|]$, so we can apply Markov's inequality to Y to see that this probability is at most $E[Y]/0.51|E| = \frac{1}{2}|E|/0.51|E| = \frac{50}{51}$. Hence, $\Pr[X \geq 0.49|E|] \geq \frac{1}{51}$. Again, since $\text{OPT} \leq |E|$, we get that $p = \Pr[X \geq 0.49\text{OPT}] \geq \frac{1}{51}$. *4pts*

Some variations on the above argument are also valid. For example, we could instead use Markov's inequality to bound $\Pr[Y \geq |E| - 0.49\text{OPT}]$, which is at most $\Pr[Y \geq 0.51\text{OPT}]$. (However, note that we *cannot* use Markov to bound $\Pr[Y \geq 0.51\text{OPT}]$ because 0.51OPT may be smaller than $E[Y] = |E|/2$). Another variation is to define $Z = \text{OPT} - X$, and note that Z is a non-negative r.v. and $E[Z] \leq 0.5\text{OPT}$. We may then apply Markov's inequality to bound $\Pr[Z \geq 0.51\text{OPT}]$

A rather different argument does not use Markov's inequality directly, but instead uses the same idea as in the *proof* of Markov's inequality. It goes as follows. Note that $E[X] \leq p\cdot\text{OPT} + (1-p)\cdot 0.49\text{OPT}$. (This follows as in the proof of Markov's inequality; the first term bounds the contribution to $E[X]$ from all values of X larger than 0.49OPT, and the second term bounds the contribution from the values less than or equal to 0.49OPT.) Since $E[X] \geq \text{OPT}/2$, we can cancel OPT through the inequality to get $p \geq 1/51$.

(c) As in 2(b) above, we expand the square to compute $E[X^2]$. For the diagonal terms, we have $E[X_e^2] = E[X_e] = \frac{1}{2}$. To handle the cross terms $E[X_eX_{e'}]$ for distinct edges e, e', observe that $E[X_eX_{e'}] = \Pr[X_e = X_{e'} = 1] = \frac{1}{4}$. This is true both for the case where e, e' share an endpoint (fix the common endpoint, then the other two endpoints must both have the opposite color, which occurs with probability $\frac{1}{4}$), and for the case where e, e' do not share an endpoint (here, $X_e, X_{e'}$ are independent). It follows that $E[X^2] = \sum_e E[X_e] + \sum_{e\neq e'} E[X_eX_{e'}] = \frac{1}{2}|E| + \frac{1}{4}|E|(|E|-1) = \frac{1}{4}|E|^2 + \frac{1}{4}|E|$. Hence, using part (a), $\mathrm{Var}[X] = E[X^2] - E[X]^2 = \frac{1}{4}|E|$. *3pts*

(d) By Chebyshev's inequality, *2pts*

$$\begin{aligned}\Pr[X < 0.49|E|] &\leq \Pr\left[|X - \tfrac{1}{2}|E|| > 0.01|E|\right] \\ &\leq \tfrac{1}{4}|E|/(0.01|E|)^2 = O(1/|E|)\end{aligned}$$

Hence, $p \geq 1 - O(1/|E|)$ as required.

(e) We keep running the above algorithm until the size of the cut is at least $0.49|E|$. By the bound in part (b) and the expectation of a geometric r.v., the expected number of repetitions we need is at most 51, so the expected running time is still linear. Correctness follows from the fact that $0.49|E| \geq 0.49\text{OPT}$. We stress that the termination condition must compare the size of the cut with $0.49|E|$ and **not** with 0.49OPT because we do not know the value of OPT! *3pts*

The bound in part (e) yields a better upper bound on the expected running time; namely, it tells us that the expected number of repetitions is in fact only $1/(1 - O(|E|^{-1})) = 1 + O(|E|^{-1})$, which approaches 1 as $|E| \to \infty$ (i.e., for large graphs).