15-251: Great Theoretical Ideas In Computer Science

Recitation 11 : Probability and Randomized Algorithms

Lecture Review

Let $\boldsymbol{X}, \boldsymbol{Y} : \Omega \to \mathbb{R}$ be random variables.

- The expectation of $\boldsymbol{X}$ is defined to be

$$\mathbf{E}[\boldsymbol{X}] = \sum_{x \in \text{range}(\boldsymbol{X})} x \cdot \Pr[\boldsymbol{X} = x] = \sum_{\ell \in \Omega} \Pr[\ell] \boldsymbol{X}(\ell).$$

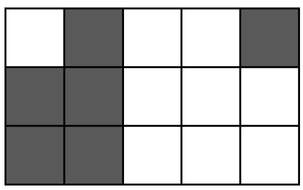
- Expectation is linear. In particular, $\mathbf{E}[\boldsymbol{X} + \boldsymbol{Y}] = \mathbf{E}[\boldsymbol{X}] + \mathbf{E}[\boldsymbol{Y}]$ even if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not independent.

Two classes of randomized algorithms:

- An algorithm A is a $T(n)$-time Las Vegas algorithm if
  - A always outputs the right answer,
  - for every input $x \in \Sigma^*$, $\mathbf{E}$[number of steps $A(x)$ takes] $\leq T(|x|)$. [1]
- An algorithm A is a $T(n)$-time Monte Carlo algorithm with error probability ε if
  - for every input $x \in \Sigma^*$, $A(x)$ gives the wrong answer with probability at most ε, [2]
  - for every input $x \in \Sigma^*$, $A(x)$ has worst-case run-time of at most $T(|x|)$.

Color Isolation

On an $r \times c$ grid of squares, color each square black or white independently with $1/2$ probability. (An example with $r = 3$ and $c = 5$ is shown below.)



Create a graph by creating a vertex for each square and putting an edge between two adjacent squares if they are of the same color (adjacent in the directions up, down, left or right). Prove that the expected number of connected components in the graph is at least $\frac{r+c}{2}$.

(Exercise: show that the expected number of connected components is also at least $\frac{rc}{16}$.)

(Hard "exercise" for the very brave: show that the expected number of connected components is also at least $\frac{rc}{8}$.)

A Hard Exam

(a) Suppose that the average score on a 150 exam was 10 points out of 100 and that 200 students took the exam. What's an upper bound on the number of students who received a perfect score? Assume that the 150 TAs are kind enough to not assign negative scores to students.

[1] In class, we said that a Las Vegas algorithm has run-time $\leq T(|x|)$ "with high probability". We meant this informally (and we aren't using any formal definition of "with high probability" that you might have seen in other classes). The definition of a Las Vegas algorithm given in this handout is the definition you should actually use.

[2] Again, the definition of a Monte Carlo algorithm from class was informal, and the definition on this handout is the one you should use.

(b) **Markov's inequality**: Let $\boldsymbol{X}$ be a non-negative random variable with non-zero expectation. For any $c > 0$,

$$\Pr[\boldsymbol{X} \geq c\,\mathbf{E}[\boldsymbol{X}]] \leq \frac{1}{c}.$$

(No need to prove this. Refer to Theorem 18.23 in the course notes to see the proof. The proof is similar to the reasoning in part (a).)

What happens in Las Vegas doesn't stay in Monte Carlo

The expected number of comparisons that the Quicksort algorithm makes is at most $2n \ln n$ (which you can cite without proof — you might see a proof of this fact if you take 15-210). Describe how to convert this Las Vegas algorithm into a Monte Carlo algorithm with the worst-case number of comparisons being $1000n \ln n$. Give an upper bound on the error probability of the Monte Carlo algorithm.

Expected Expectation

Definition: Let $\boldsymbol{X}$ be a random variable and E be an event. The conditional expectation $\boldsymbol{X}$ given the event E, denoted by $\mathbf{E}[\boldsymbol{X} \mid E]$, is defined as

$$\mathbf{E}[\boldsymbol{X} \mid E] = \sum_{x \in \text{range}(\boldsymbol{X})} x \cdot \Pr[\boldsymbol{X} = x \mid E].$$

(a) In lecture, you learned about the law of total probability. Prove the law of total expectation: given a random variable $\boldsymbol{X}$ and an event A,

$$\mathbf{E}[\boldsymbol{X}] = \mathbf{E}[\boldsymbol{X} \mid E] \Pr[E] + \mathbf{E}\left[\boldsymbol{X} \mid E^c\right] \Pr\left[E^c\right].$$

(b) Suppose $\boldsymbol{X} \sim$ Geometric(p) for some $0 < p \leq 1$. Use the Law of Total Expectation to show that $\mathbf{E}[\boldsymbol{X}] = 1/p$.

(Brain Teaser) Passive-Aggressive Passengers

Consider a plane with n seats $s_1, s_2, \ldots, s_n$. There are n passengers, $p_1, p_2, \ldots, p_n$ and they are randomly assigned unique seat numbers. The passengers enter the plane one by one in the order $p_1, p_2, \ldots p_n$. The first passenger p_1 does not look at their assigned seat and instead picks a uniformly random seat to sit in. All the other passengers, p_2, p_3, ... , p_n, use the following strategy. If the seat assigned to them is available, they sit in that seat. Otherwise they pick a seat uniformly at random among the available seats, and they sit there. What is the probability that the last passenger, p_n, will end up sitting in their assigned seat?