Randomized Algorithms - Homework III

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

1 Instructions

1. The homework is due on April 10, in class.
2. Each question is worth 4 points.
3. Attempt as many problems as you can. You will be given partial credit, as per the policy discussed in class.
4. The work must be entirely your own. You are expressly **prohibited** from consulting with colleagues or the internet (with the exception of the material on the course website). You may refer to books in the library, crediting the same, if used in your work.

2 Problems

1. Let X and Y be independent geometric random variables, where X has parameter p and Y has parameter q.
   (a) What is the probability that $X = Y$?
   (b) What is $\mathbf{E}[\max(X, Y)]$?
   (c) What is $P(\min(X, Y) = k)$?
   (d) What is $\mathbf{E}[X \mid X \leq Y]$?
2. Show that if $f : \Re \rightarrow \Re$ is convex, then for any $x_1, x_2, \ldots, x_n$ and $\lambda_1, \lambda_2, \ldots, \lambda_n$, with $\sum_{i=1}^{n} \lambda_i = 1$,
   $$f(\sum_{i=1}^{n} \lambda_i \cdot x_i) \leq \sum_{i=1}^{n} \lambda_i \cdot f(x_i).$$
   Let X denote a random variable that takes on finitely many values. Assuming that $f()$ is convex, show that,
   $$\mathbf{E}[f(X)] \geq f(\mathbf{E}[X]).$$
3. (a) Determine the moment generating function for the binomial random variable $B(n, p)$.
   (b) Let X be a $B(n, p)$ random variable and let Y be a $B(m, p)$ random variable, where X and Y are independent. Determine the moment generating function of $(X + Y)$.
   (c) What can we conclude from the form of the moment generating function of $(X + Y)$?
4. Suppose that we can obtain independent samples X_1, X_2, ... of a random variable X, and we want to use these samples to estimate $\mathbf{E}[X]$. Using t samples, we use $\sum_{i=1}^{t} \frac{X_i}{t}$ for our estimate of $\mathbf{E}[X]$. We want the estimate to be within $\epsilon \cdot \mathbf{E}[X]$ from the true value of $\mathbf{E}[X]$ with probability at least $(1 - \delta)$. We may not be able to use Chernoffs bound directly to bound how good our estimate is if X is not a $0-1$ random variable, and we do not know its moment generating function. Consider the following alternative approach that requires only having a bound on the variance of X.

Let $r = \frac{\sqrt{\mathbf{Var}[X]}}{\mathbf{E}[X]}$.

(a) Show using Chebyshevs inequality that $O(\frac{r^2}{\epsilon^2 \cdot \delta})$ samples are sufficient to solve the above problem.

(b) Suppose that we only need a weak estimate that is within $\epsilon \cdot \mathbf{E}[X]$ of $\mathbf{E}[X]$ with probability at least $\frac{3}{4}$. Argue that only $O(\frac{r^2}{\epsilon^2})$ samples are enough for this weak estimate.

(c) Show that by taking the median of $O(\log \frac{1}{\delta})$ weak estimates, we can obtain an estimate within $\epsilon \cdot \mathbf{E}[X]$ of $\mathbf{E}[X]$ with probability at least $(1 - \delta)$. Conclude that we only need $O(\frac{r^2 \cdot \log \frac{1}{\delta}}{\epsilon^2})$ samples.

5. Supposet that n balls are thrown independently and uniformly into n bins.

   (a) Find the conditional probability that bin 1 has one ball given that exactly one ball fell into the first three bins.

   (b) Find the conditional expectation of the number of balls in bin 1, under the condition that bin 2 received no balls.

   (c) What is the probability that every bin receives exactly one ball? (You may solve using Poisson approximation).