600.464 Randomized and Big Data Algorithms
Homework #2 Answers

Ravindra Gaddipati

October 18, 2016

Problem 1 (3 points)

Alice and Bob play checkers often. Alice is a better player. so the probability that she wins any given game is 0.6, independent of all other games. They decide to play a tournament of n games. Bound the probability that Alice loses the tournament using a Chernoff bound.

Answer:

Let $X_i, ..., X_n$ be the outcome of each game, where $X_i = 1$ if Alice wins, and $X_i = 0$ if Alice loses or it is a tie. Further, Let $X = \sum_{i=0}^{n} X_i$. Alice wins 0.6 of the games, so we can compute $E[X] = \mu = n(0.6 * 1 + 0.4 * 0) = (3/5)n$. If Alices loses the tournament, then we want to bound $P(X < n/2)$, where a tied tournament is not considered a loss for Alice. We use the multiplicative form of the Chernoff bound,

$$P(X < (1-\delta)\mu) < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$$

For the desired bound, we compute the upper bound of δ.

$$\begin{aligned} n/2 &= (1-\delta)\mu \\ &= (1-\delta)\frac{3}{5}n \\ \delta &= 1/6 \end{aligned}$$

The Chernoff bound then gives us

$$\begin{aligned} P(lose) = P(X < n/2) &< \left(\frac{e^{-1/6}}{(1-1/6)^{1-1/6}}\right)^{0.6n} \\ &< (6/5)^{0.5n} e^{-0.1n} \end{aligned}$$

Problem 2 (3 points)

a. In an election with two candidates using paper ballots, each vote is independently misrecorded with probability p = 0.02. Use a Chernoff bound to bound the probability that more than 4% of the votes are misrecorded in an election of 1,000,000 ballots.

b. Assume that a misrecorded ballot always counts as a vote for the other candidate. Suppose that candidate A received 510,000 votes and that candidate B received 490,000 votes. Use Chernoff bounds to bound the probability that candidate B wins the election owing to misrecorded ballots. Specifically, let X be the number of votes for candidate A that are misrecorded and let Y be the number of votes for candidate B that are misrecorded. Bound $Pr((X > k) \cap (Y < l))$ for suitable choices of k and l.

Answer:

a.

Let X be the number of incorrect votes. For an election of 1,000,000 votes, then

$$\mu = E[X] = 1000000 \times p = 1000000 \times 0.02 = 20000$$

We want to bound the probability more than 4% of the votes are misrecorded, $P(X > 0.04 * 1000000) = P(X > 40000$. We use the multiplicative form the Chernoff bound:

$$P(X > (1+\delta)\mu) < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

For the desired bound, we compute the upper value of δ

$$\begin{aligned}(1-\delta)\mu &= 40000\\ \delta &= 1\end{aligned}$$

Applying the Chernoff bound,

$$\begin{aligned}P(X > 40000) &< \left(\frac{e^{1}}{(1+1)^{1+1}}\right)^{20000}\\ &< \left(\frac{e}{4}\right)^{20000}\end{aligned}$$

b.

Let X be the number of votes for candidate A that are misrecorded, and let Y be the number of votes for candidate B that are misrecorded. We want to bound the probability that candidate B wins, $P(X + \bar{Y} > 500000)$, where $\bar{Y}$ is the number of correctly recorded votes for candidate B. The expected value is,

$$\begin{aligned}\mu = E[X + \bar{Y}] &= 510000p + 490000(1 - p)\\ &= 510000(0.02) + 490000(1 - 0.02)\\ &= 490400\end{aligned}$$

We find δ for the desired bound,

$$\begin{aligned}500000 &= (1 + \delta)\mu\\ 500000 &= (1 + \delta)490400\\ \delta &= \frac{12}{613}\end{aligned}$$

Using the same Chernoff bound form in **(a)**,

$$\begin{aligned}P(Bwins) = P(X + \bar{Y} > 500000) &< \left(\frac{e^{12/613}}{(1 + 12/613)^{1+12/613}}\right)^{490400}\\ &< 2.855 \times 10^{-41}\end{aligned}$$

Since each vote is an independent event,

$$P(X + \bar{Y} > 500000) = P((X > k) \cap (Y < l)) = P(X > k) + P(Y < l)$$

where k and l satisfy the constraint $k - l > 10000$.

Problem 3 (3 points)

Recall that a function f is said to be *convex* if for any x_1, x_2 and for $0 \leq \lambda \leq 1$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

a. Let Z be a random variable that takes on a (finite) set of values in the interval [0,1], and let $p = E[Z]$. Define the Bernoulli random variable X by $Pr(X = 1) = p$ and $Pr(X = 0) = 1 - p$, show that $E[f(Z)] \leq E[f(X)]$ for any convex function f.

b. Use the fact that $f(x) = e^{tx}$ is convex for any $t \geq 0$ to obtain a Chernoff-like bound for Z based on a Chernoff bound for X.

Answer:
a.
We represent Z as a convex combination of 0 and 1, for $\lambda \in [0,1]$

$$\begin{aligned} Z &= x_1\lambda + (1-\lambda)x_2 \\ &= \lambda(x_1 - x_2) + x_2 \\ \lambda &= \frac{Z - x_2}{x_1 - x_2} \end{aligned}$$

By the definition of a convex function, this holds for any x_1, x_2. We pick $x_1 = 1$ and $x_2 = 0$ and take the expectation. We use $E[Z] = p$ and the linearity of the function to obtain,

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda f(x_1) + (1-\lambda)f(x_2) \\ f(Z) &\leq \frac{Z - x_2}{x_1 - x_2} f(x_1) + \left(1 - \frac{Z - x_2}{x_1 - x_2}\right) f(x_2) \\ &\leq \frac{Z - 0}{1} f(1) + \left(1 - \frac{Z - 0}{1}\right) f(0) \\ &\leq Zf(1) + (1 - Z)f(0) \\ E[f(Z)] &\leq E[Zf(1) + (1-Z)f(0)] \\ &\leq E[Z]f(1) + E[1-Z]f(0) \\ &\leq f(1)p + f(0)(1-p) \end{aligned}$$

Similarly, we take the expectation of $f(X)$.

$$\begin{aligned} E[f(X)] &= \sum_{i=0}^{1} p_i X_i \\ &= f(1)p + f(0)(1-p) \end{aligned}$$

We see that these forms match the definition of a convex function, thus it follows that $E[f(Z)] \leq E[f(X)]$

$$(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$
$$E[f(Z)] \leq E[f(X)]$$

b.
We represent Z as a sum $\sum_{i=1}^{n} Z_i$ where $\mu = E\left[\sum_{i=1}^{n} Z_i\right]$. Using the Chernoff bound, we see:

$$P\left[\sum_{i=1}^{n} Z_i > (1-\delta)\mu\right] = P\left[e^{\sum_{i=1}^{n} Z_i} > e^{(1-\delta)\mu}\right] \leq \frac{E\left[\sum_{i=1}^{n} Z_i\right]}{e^{t(1-\delta)\mu}}$$

From **(a)** we know $E[f(Z)] \leq E[f(X)]$. Using this and the fact that $E[\sum_{i=1}^{n} Z_i] = E[Z]$, it follows that

$$\frac{E\left[\sum_{i=1}^{n} Z_i\right]}{e^{t(1-\delta)\mu}} \leq \frac{E[e^{tX}]}{e^{(1+\delta)\mu}}$$
$$P\left[e^{Z} > e^{(1-\delta)\mu}\right] \leq \frac{E[e^{tX}]}{e^{(1+\delta)\mu}}$$

Problem 4 (3 points)

Let $X_0 = 0$ and for $j \geq 0$ let X_{j+1} be chosen uniformly over the real interval $[X_j, 1]$. Show that, for $k \geq 0$, the sequence

$$Y_k = 2^k(1 - X_k)$$

is a martingale.

Answer:

To show that the sequence Y_n is a martingale, we demonstrate that it satisfies the 3 properties of a martingale:

- $E[Y_{n+1}|X_1, ..., X_n] = Y_n$
- $E[|Y_n|] < \infty$
- Y_n is a function of $X_0, ..., X_n$

To show that $E[Y_{n+1}|X_1, ..., X_n] = Y_n$, we take the expectation, where X is a uniform continuous distribution with $\mu = \frac{X_n+1}{2}$.

$$\begin{aligned} E[Y_{n+1}|X_1, ..., X_n] &= E[2^{n+1}(1 - X_{n+1})] \\ &= 2^{n+1}(1 - E[X_{n+1}]) \\ &= 2^{n+1}(1 - \frac{X_n + 1}{2}) \\ &= 2^n(1 - X_n) = Y_n \end{aligned}$$

To show that $E[|Y_n|] < \infty$, we take the expectation again where $E[X_n] = 1 - 1/2^n$ since it is a uniform continuous distribution over X_{n-1} and 1. This holds as $X_0 = 0$.

$$\begin{aligned} E[Y_n] &= E[2^n(1 - X_n)] \\ &= 2^n E[1 - X_n] \\ &= 2^n - 2^n E[X_n] \\ &= 2^n - 2^n(1 - 1/2^n) \\ &= 1 \end{aligned}$$

We know Y_k is a function of X_k, where X_{j+1} is chosen in the interval $[X_j, 1]$. As such, we know $X_{j+1} \geq X_j$ for all X_j with $X_0 = 0$. It follows that Y_k is a function of $X_0, ..., X_n$, thus we have shown Y_k is a martingale.

Problem 5 (3 points)

Consider a random graph from $G_{n,N}$, where $N = cn$ for some constant $c > 0$. Let X be the expected number of isolated verticies (i.e., verticies of degree 0).

a. Determine $E[X]$

b. Show that:

$$P(|X - E[X]| \geq 2\lambda\sqrt{cn}) \leq 2e^{-\lambda^2/2}$$

Answer:

a.

Consider a set of $N = cn$ vertices. We randomly select two vertices to construct an edge n independent times. The probability of selecting a particular vertex is then $1/n$ as self-loops are allowed. Let k_i be the event that a particular vertex does not have an edge, with a degree 0. With $p = 1/n$, the probability that a vertex is not selected on a specific trial is $1 - 1/n$.

$$P(k_i) = \left(1 - \frac{1}{n}\right)^{2cn}$$

For a bin to have degree 0, it must have not been selected for cn trials, were each trial involves 2 random selections.

$$\begin{aligned}
E[X] &= \sum_{i=1}^{n} E[P[k_i]] \\
&= \sum_{i=1}^{n} (1 - 1/n)^{2cn} \\
&= n(1 - 1/n)^{2cn}
\end{aligned}$$

Thus the expected number of vertices with degree 0 is

$$E[X] = n(1 - 1/n)^{2cn}$$

b.

We use the edge exposure martingale $Z_i = E[F(G)|Y_0, ..., Y_{i-1}]$ where $F(G)$ is the number of verticies with degree 0, and $Y_i = 1$ if an edge is revealed, otherwise 0. When there are zero edges revealed, then $Z_0 = E[X]$. Similarly after all the edges are revealed, $Z_{cn} = X$. Because an edge can either connect 1 vertex (self loop) or 2 verticies, we can bound k from above as $k \leq 2$.

We know for a martingale $X_0, X_1, ...$ and for all $k \geq 1$ such that (Mitzenmacher, Upfal, Corollary 12.5):

$$|X_k - X_{k-1}| \leq k$$

Then, for all $t \geq 1$ and $\lambda > 0$,

$$P(|X_t - X_0| \geq \lambda k \sqrt{t}) \leq 2e^{-\lambda^2/2}$$

$$P(|X - E[X]| \geq 2\lambda \sqrt{cn}) \leq 2e^{-\lambda^2/2}$$