Theory of Computation Complexity

Tutorial 2

LIU yang

Outline

- Tail inequalities
  - Markov’s inequality
  - Chebyshev inequality
  - Chernoff’s bound
- Randomized communication complexity
  - About the success probability

Markov inequality

- Let X be a nonnegative random variable. Then for all a>0, $\Pr(X \geq a) \leq \frac{E[X]}{a}$.

Proof: $E[X] = \sum_{i} i \Pr(X=i)$

$$= \sum_{i<a} i \Pr(X=i) + \sum_{i \geq a} i \Pr(X=i)$$

$$\geq a \sum_{i \geq a} \Pr(X=i) = a \Pr(X \geq a) \blacksquare$$

Chebyshev inequality

$$\Pr(|X - E(X)| \geq a) \leq \frac{Var[X]}{a^2}$$

Proof:

$$\Pr(|X - E(X)| \geq a) = \Pr\left(\left(X - E(X)\right)^2 \geq a^2\right)$$

$$\leq \frac{E\left[\left(X - E(X)\right)^2\right]}{a^2} \quad \leftarrow \text{by Markov}$$

$$= \frac{Var[X]}{a^2} \quad \blacksquare$$

Corollary $\Pr(|X - E(X)| \geq t \cdot \sigma(X)) \leq \frac{1}{t^2}$

$$\Pr(|X - E(X)| \geq t \cdot E[X]) \leq \frac{Var[X]}{t^2 \left(E[X]\right)^2}$$

Chernoff bounds

- Let $X_1, \ldots, X_n$ be independent {0,1}-valued random variables, and $p_i = \Pr[X_i = 1] = E[X_i]$, $X = \sum_i^n X_i$, $u = E[X]$.
- Then for any $\delta > 0$:

$$\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu[(1+\delta)\ln(1+\delta)-\delta]};$$

$$\Pr[X < (1-\delta)\mu] < \exp(-\mu\delta^2/2).$$

$$
\begin{aligned}
\Pr[e^{tX} \geq e^{t(1+\delta)\mu}] &\leq \tfrac{1}{e^{t(1+\delta)\mu}} \cdot \mathbb{E}[e^{t\sum_i X_i}] && \text{Markov's inequality} \\
&= \tfrac{1}{e^{t(1+\delta)\mu}} \cdot \mathbb{E}[\textstyle\prod_i e^{tX_i}] \\
&= \tfrac{1}{e^{t(1+\delta)\mu}} \cdot \textstyle\prod_i \mathbb{E}[e^{tX_i}] && \text{independence of the } X_i \\
&= \tfrac{1}{e^{t(1+\delta)\mu}} \cdot \textstyle\prod_i [1 + p_i(e^t - 1)] \\
&\leq \tfrac{1}{e^{t(1+\delta)\mu}} \cdot \textstyle\prod_i e^{p_i(e^t-1)} && 1 + x \leq e^x \ \forall x \in \mathbb{R} \\
&= \tfrac{1}{e^{t(1+\delta)\mu}} \cdot e^{\mu(e^t-1)} && \textstyle\sum_i p_i = \mu \\
&= e^{-\mu(t(1+\delta)+1-e^t)}
\end{aligned}
$$

We now want to choose t to maximize $t(1+\delta)+1-e^t$. Setting the derivative equal to 0 and solving for t gives $t = \ln(1+\delta)$. Plugging this value into the last line gives the stated bound. $\square$

For $0 \leq \delta \leq 1$, $(1+\delta)\ln(1+\delta) - \delta \geq \delta^2/2$.

So, the Chernoff bound can be restated as $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/2}$ for $\delta \in [0,1]$.

$\Pr[X < (1-\delta)\mu] < \exp(-\mu\delta^2/2)$.

Randomized communication complexity

- The randomized communication complexity actually depends on the required error probability, which we will use the subscript to specify.

- For example, $R_\varepsilon(f)$ is the ε-error private-coin randomized communication complexity. When we don't specify the error probability, it's 1/3 by convention; that is, $R(f) = R_{1/3}(f)$.

- Like in randomized algorithms, the error probability can be decreased from a constant to an arbitrary ε by repeating the protocol $O(\log 1/\varepsilon)$ times.

Amplify the success probability

- Simply run the protocol (with error probability p) k times, and take a majority vote on all the k outcomes.
  - p is a constant and p<1/2.
- Then by Chernoff bound, it will give the wrong answer with probability at most

$$\Pr[X \ge k/2] = \Pr[X \ge (1+\frac{1-2p}{2p})pk]$$

$$\le e^{-pk(\frac{1-2p}{2p})^2/2} = \frac{1}{2^{\Omega(k)}}.$$

- Which is less than ε, when we repeat $k = O(\log 1/\varepsilon)$ times.

Summery

- Some tail inequalities.
- Amplify the success probability of randomized protocols.

Thanks