CSE 632 (Fall 2019): Analysis of Algorithms II : Randomized Algorithms

Lecture 10 (9/27/2019): Concentration Bounds

Lecturer: Shi Li *Scriber: Shuyan Chen*

1 Concentration bounds for bounding probability

We consider the following motivating Problem: Toss a fair coin n times, what is the probability that we get at least $0.6n$ head-up's. How fast does the probability diminish as n grows?

We shall give three concentration inequalities, each being stronger than the previous one. The upper bounds on the probability of the above event E given by the three inequalities are given below:

- Using Markov's Inequality: the $\Pr(E) \leq \frac{5}{6}$.
- Using Chebyshev's Inequality: the $\Pr(E) \leq \frac{25}{2n}$.
- Using Chernoff Bound: the $\Pr(E) \leq e^{-0.006n}$.

2 Markov's Inequality and Chebyshev's Inequality

2.1 Markov's Inequality

Lemma 1 (Markov's Inequality). *Let X be a random variable taking non-negative values and $\mu = \mathbb{E}[X]$, then for every $a \geq 1$, we have*

$$\Pr[X > a\mu] < \frac{1}{a}.$$

Proof. Assume $\Pr[X > a\mu] \leq \frac{1}{a}$ then we can get

$$\mathbb{E}[X] > \frac{1}{a} \cdot a\mu = \mu.$$

This is a contradiction with $\mathbb{E}[X] = \mu$. □

To apply Markov's Inequality on the motivating question, we define X to be the of head-up's we get, then $\mu = \mathbb{E}[X] = 0.5n$, so we have $\Pr[X \geq 0.6n] \leq \frac{\mu}{a\mu} = \frac{5}{6}$. The inequality is weak, in the sense that the upper bound $\frac{5}{6}$ does not decrease as n grows.

2.2 Chebyshev's Inequality

Lemma 2 (Chebyshev's Inequality). *Let X be a random variable, $\mu = \mathbb{E}[X]$, $\mathrm{Var}[X] = \delta^2$, where $\sigma \geq$ is the standard deviation of X. Then for $\forall a \geq 1$, we have*

$$\Pr[|x - \mu| > a \cdot \sigma] < \frac{1}{a^2}.$$

Proof. Assume $\Pr[|x - \mu| > a \cdot \sigma] \geq \frac{1}{a^2}$. Then we can get

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mu)^2] \geq \frac{1}{a^2} \cdot (a \cdot \sigma)^2 = \sigma^2.$$

This contradicts with $\mathrm{Var}[X] = \sigma^2$. □

Applying Chebyshev's Inequality to the motivating question. Because the n coin tosses are independent, we have $\mathrm{Var}[X] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \cdots + \mathrm{Var}[X_n] = \frac{n}{4}$, where X_i is the result of the i-th coin toss. Recall that $\mu = \mathbb{E}[X] = 0.5n$. So,

$$\Pr[X > .6n] = \frac{1}{2}\Pr[|X - \mu| > 0.1n] < \frac{1}{2}\left(\frac{1}{0.2\sqrt{n}}\right)^2 = \frac{25}{2n}.$$

Above, we use the Cheybshev's inequality for $a = \frac{0.1n}{\sigma} = \frac{0.1n}{\sqrt{n}/2} = 0.2\sqrt{n}$. So, unlike Markov's inequality, Chebyshev's Inequality gives a bound on $\Pr[X > 0.6n]$ that decreases as n grows.

3 Chernoff Bounds

Lemma 3. *Let X_1,X_2, ... ,X_n be independent variables taking values in [0,1]. Let $X = X_1 + X_2 + X_3 + ... + X_n$ and $\mu = \mathbb{E}[X]$, then for every $\delta > 0$ we have*

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu},$$

$$\textit{and } \Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}.$$

For a simple form with looser bounds, we have for every $\delta \in (0,1)$,

$$\Pr[X > (1+\delta)\mu] < e^{\frac{-\delta^2\mu}{3}},$$

$$\textit{and } \Pr[X < (1-\delta)\mu] < e^{\frac{-\delta^2\mu}{2}}.$$

Proof. Let t be some number whose value will be decided later. As X is the sum of n independent variables, we have

$$\mathbb{E}[e^{tX}] = \mathbb{E}[e^{t(X_1+X_2+...+X_n)}] = \mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n}\mathbb{E}[e^{tX_i}].$$

Let us define $\mu_i = \mathbb{E}[X_i]$, for every $i \in [n]$. Then as e^{tx} is a convex function of x, we have for every $i \in [n]$

$$\mathbb{E}[e^{tX_i}] \le \mathbb{E}[(1-X_i)e^{t\cdot 0} + X_i e^{t\cdot 1}] = 1 - \mu_i + e^t\mu_i = 1 + (e^t-1)\mu_i.$$

So,

$$\mathbb{E}[e^{tX}] \le \prod_{i=1}^{n}[1+(e^t-1)\mu_i] \le \prod_{i=1}^{n} e^{(e^t-1)\mu_i} = e^{\sum_{i=1}^{n}(e^t-1)\mu_i} = e^{(e^t-1)\mu}.$$

Assume $t \ge 0$, use Markov's Inequality here, we can get

$$\Pr[X > (1+\delta)\mu] = \Pr[e^{tX} > e^{t(1+\delta)\mu}] \le \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = e^{[e^t-1-t(1+\delta)]\mu}$$

Set $t = \ln(1+\delta)$, we have $e^t - 1 - t(1+\delta) = 1 + \delta - 1 - (1+\delta)\ln(1+\delta) = \delta - (1+\delta)\ln(1+\delta)$. Then $e^{e^t-1-t(1+\delta)} = \frac{e^{\delta}}{(1+\delta)^{1+\delta}}$. This gives the first inequality in the lemma.

To see the second inequality (for which we can assume $\delta \in (0,1)$), we set t be a negative number. Then

$$\Pr[X < (1-\delta)\mu] = \Pr[e^{tX} > e^{t(1-\delta)\mu}] \le \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu}} = e^{[e^t-1-t(1-\delta)]\mu}$$

Set $t = \ln(1-\delta) < 0$, we have $e^t - 1 - t(1-\delta) = 1 - \delta - 1 - (1-\delta)\ln(1-\delta) = -\delta - (1-\delta)\ln(1-\delta)$ and $e^{e^t-1-t(1-\delta)} = \frac{e^{-\delta}}{(1-\delta)^{1-\delta}}$. This gives the second bound.

For $\delta \in (0,1)$, we have

$$\delta - (1+\delta)\ln(1+\delta) \le \delta - (1+\delta)\frac{2\delta}{2+\delta} = \frac{-\delta^2}{2+\delta} \le -\frac{\delta^2}{3}.$$

$$-\delta - (1-\delta)\ln(1-\delta) \le -\delta - (1-\delta)\frac{-2\delta}{2-\delta} = \frac{-\delta^2}{2-\delta} \le -\frac{\delta^2}{2}.$$

□

For the motivating question, we have $\Pr[X > 0.6n] = \Pr[X > (1+0.2)\mu] \le \exp\left(-0.2^2\mu/3\right) \le e^{-0.006n}$.

Notice that the Chernoff bound requires the n random variables to be independent, while Cherbeshev's inequality holds without independence of the random variables. Indeed, the latter only involves one random variable X.