CSE 632: Analysis of Algorithms II: Randomized Algorithms **Spring 2024**

Lecture 6: Concentration Inequalities

Lecturer: Zongchen Chen

1 Markov's Inequality

Lemma 1 (Markov's Inequality)**.** *Let X be a* non-negative *random variable with mean $\mu = \mathbb{E}X$. Then we have*

$$\forall a \geq 0: \quad \Pr\left(X \geq a\mu\right) \leq \frac{1}{a};$$
$$\textit{Equivalently,} \quad \forall a \geq 0: \quad \Pr\left(X \geq a\right) \leq \frac{\mu}{a}.$$

Proof. We have

$$\begin{aligned}\mathbb{E}X &= \Pr(X < a)\,\mathbb{E}[X \mid X < a] + \Pr(X \geq a)\,\mathbb{E}[X \mid X \geq a] \\ &\geq \Pr(X \geq a)\,\mathbb{E}[X \mid X \geq a] \\ &\geq \Pr(X \geq a) \cdot a.\end{aligned}$$

The lemma then follows. □

Example 2. Suppose we toss a fair coin n times. What is the probability that we get at least $0.6n$ Heads? How fast does the probability diminish as n grows?

Let $X_i = 1$ if the ith toss is Head and $X_i = 0$ otherwise. So $\mathbb{E}X_i = 1/2$. Let $X = \sum_{i=1}^n X_i$ be the number of Heads in n tosses. So $\mathbb{E}X = n/2$. By Markov's Inequality, we have

$$\Pr(X \geq 0.6n) \leq \frac{0.5n}{0.6n} = \frac{5}{6}.$$

2 Chebyshev's Inequality

Lemma 3 (Chebyshev's Inequality)**.** *Let X be a random variable with mean $\mu = \mathbb{E}X$ and variance $\sigma^2 = \mathrm{Var}(X)$ where $\sigma \geq 0$ is the standard deviation. Then we have*

$$\forall a \geq 0: \quad \Pr\left(|X - \mu| \geq a\sigma\right) \leq \frac{1}{a^2};$$
$$\textit{Equivalently,} \quad \forall a \geq 0: \quad \Pr\left(|X - \mu| \geq a\right) \leq \frac{\sigma^2}{a^2}.$$

Proof. Let $Y = (X - \mu)^2$. Note that $Y \geq 0$ and $\mathbb{E}Y = \mathbb{E}[(X - \mu)^2] = \sigma^2$. We then have

$$\begin{aligned}\Pr\left(|X - \mu| \geq a\right) &= \Pr\left(Y \geq a^2\right) \\ &\leq \frac{\mathbb{E}Y}{a^2} && \text{(Markov's Inequality)} \\ &= \frac{\sigma^2}{a^2},\end{aligned}$$

as wanted. □

Example 4. As in Example 2, each X_i has variance $1/4$ and hence $\mathrm{Var}(X) = n/4$. By Chebyshev's Inequality, we have

$$\begin{aligned}\Pr(X \geq 0.6n) &= \Pr(X - \mathbb{E}X \geq 0.1n) \\ &= \frac{1}{2}\Pr(|X - \mathbb{E}X| \geq 0.1n) \\ &\leq \frac{1}{2}\frac{n/4}{(0.1n)^2} = \frac{12.5}{n} \overset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

3 Chernoff Bounds

Lemma 5 (Chernoff Bounds)**.** *Let $X_1, \dots, X_n$ be* independent *random variables taking values in* $[0,1]$. *Let $X = \sum_{i=1}^n X_i$ be their sum with mean $\mu = \mathbb{E}X = \sum_{i=1}^n \mathbb{E}X_i$. Then we have*

$$\forall\delta \geq 0: \quad \Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu \tag{1}$$

$$\text{and} \quad \Pr(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^\mu. \tag{2}$$

A looser but simpler version is

$$\forall\delta \in (0,1): \quad \Pr(X \geq (1+\delta)\mu) \leq e^{-\frac{\delta^2\mu}{3}} \tag{3}$$

$$\text{and} \quad \Pr(X \leq (1-\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2}}. \tag{4}$$

Proof. Our plan is to pick a "nice" auxiliary function φ and apply Markov's inequality to $Y = \varphi(X)$. Namely, we expect to have

$$\begin{aligned}\Pr(X \geq (1+\delta)\mu) &\overset{\text{(i)}}{\leq} \Pr(\varphi(X) \geq \varphi((1+\delta)\mu)) \\ &\overset{\text{(ii)}}{\leq} \frac{\mathbb{E}[\varphi(X)]}{\varphi((1+\delta)\mu)} && \text{(Markov's Inequality)} \\ &\overset{\text{(iii)}}{\leq} \text{nice bound},\end{aligned}$$

where (i) requires φ is monotone increasing, (ii) requires $\varphi \geq 0$ since Markov's inequality applies only to non-negative random variables, and (iii) requires we can upper bound $\mathbb{E}[\varphi(X)]$ nicely.

Our choice of φ is the exponential function $\varphi(x) = e^{tx}$ where $t \geq 0$ is a parameter to be decided. Note that

$$\mathbb{E}[\varphi(X)] = \mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t\sum_{i=1}^n X_i}\right] = \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{tX_i}\right],$$

where the last equality is because $X_1, \dots, X_n$ are independent.

Fact 6. *Suppose Z is a random variable taking values in $[0,1]$. Then we have*

$$\mathbb{E}\left[e^{tZ}\right] \leq 1 + (e^t - 1)\mathbb{E}Z.$$

Proof. For all $x \in [0,1]$, it holds $e^{tx} \leq 1 + (e^t - 1)x$. Therefore,

$$\mathbb{E}\left[e^{tZ}\right] \leq \mathbb{E}\left[1 + (e^t - 1)Z\right] = 1 + (e^t - 1)\mathbb{E}Z$$

as claimed. ☐

By Fact 6, we deduce that

$$\begin{aligned}
\mathbb{E}\left[e^{tX}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right] &\le \prod_{i=1}^{n} \left(1 + (e^t - 1)\mathbb{E}X_i\right) \\
&\le \prod_{i=1}^{n} e^{(e^t-1)\mathbb{E}X_i} \\
&= e^{\sum_{i=1}^{n}(e^t-1)\mathbb{E}X_i} \\
&= e^{(e^t-1)\mu}. \qquad (5)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Pr\left(X \ge (1+\delta)\mu\right) &\le \Pr\left(e^{tX} \ge e^{t(1+\delta)\mu}\right) && \text{(since } t \ge 0\text{)} \\
&\le \frac{\mathbb{E}\left[e^{tX}\right]}{e^{(1+\delta)t\mu}} && \text{(Markov's Inequality)} \\
&\le \frac{e^{(e^t-1)\mu}}{e^{(1+\delta)t\mu}} && \text{(by Eq. (5))} \\
&= e^{\left(e^t-1-(1+\delta)t\right)\mu}. && (6)
\end{aligned}$$

Notice that Eq. (6) holds for all $t \ge 0$. We would pick a best $t \ge 0$ that minimizes the exponent $e^t - 1 - (1+\delta)t$. By Calculus our choice is $t = \log(1+\delta) \ge 0$, and for this t Eq. (6) becomes

$$\Pr\left(X \ge (1+\delta)\mu\right) \le \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu},$$

which establishes Eq. (1).

The proof of Eq. (2) is very similar. Let $t \le 0$ to be decided, and we have

$$\begin{aligned}
\Pr\left(X \le (1-\delta)\mu\right) &\le \Pr\left(e^{tX} \ge e^{t(1-\delta)\mu}\right) && \text{(since } t \le 0\text{)} \\
&\le \frac{\mathbb{E}\left[e^{tX}\right]}{e^{(1-\delta)t\mu}} && \text{(Markov's Inequality)} \\
&\le \frac{e^{(e^t-1)\mu}}{e^{(1-\delta)t\mu}} && \text{(by Eq. (5))} \\
&= e^{\left(e^t-1-(1-\delta)t\right)\mu}. && (7)
\end{aligned}$$

We choose $t = \log(1-\delta) \le 0$ which minimizes the exponent $e^t - 1 - (1-\delta)t$, and then Eq. (7) becomes

$$\Pr\left(X \le (1-\delta)\mu\right) \le \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu},$$

establishing Eq. (2).

The bounds Eqs. (3) and (4) are obtained by upper bounding the right-hand sides of Eqs. (1) and (2) with simpler functions. □

Example 7. As in Example 2, by Chernoff Bound we have

$$\Pr\left(X \ge 0.6n\right) = \Pr\left(X \ge (1+0.2)\,\mathbb{E}X\right) \le \left(\frac{e^{0.2}}{1.2^{1.2}}\right)^{0.5n} \le 0.991^n \overset{n\to\infty}{\longrightarrow} 0.$$

In particular, the probability decays to zero exponentially fast.