CS 388R: Randomized Algorithms, Fall 2019 September 1, 2019

Lecture 2: Concentration inequalities

Prof. Eric Price *Scribe: Will Sherwood, Feifei Zhang*

NOTE: THESE NOTES HAVE NOT BEEN EDITED OR CHECKED FOR CORRECTNESS

1 Concentration inequalities

1.1 Markov's Inequality

Let X be a nonnegative random variable. Then, for any $t > 0$,

$$\Pr[X \geq t] \leq \frac{1}{t}\mathbb{E}[x].$$

Proof.

$$\mathbb{E}[X] = \int_0^\infty f_X(s) \cdot s \text{ ds} \geq \int_t^\infty f_X(s) \cdot s \text{ ds} \geq \int_t^\infty f_X(s) \cdot t \text{ ds} = \Pr[X \geq t].$$

□

1.2 Chebyshev's Inequality

Let X be a random variable with variance σ^2 and finite expectation. Then, for any $t > 0$,

$$\Pr[|X - \mathbb{E}[X]| > t\sigma] \leq \frac{1}{t^2}.$$

Proof. Let Y denote the non-negative random variable $|X - \mathbb{E}[X]|$. By Markov's Inequality,

$$\Pr[|X - \mathbb{E}[X]| \geq t\sigma] = \Pr[Y \geq t\sigma] = \Pr[Y^2 \geq t^2\sigma^2] \leq \frac{1}{t^2\sigma^2}\mathbb{E}[Y^2] = \frac{1}{t^2}.$$

□

1.3 Chernoff Bounds

Let X be a random variable with expectation μ. Then,

$$\Pr[X - \mu \geq t] \leq \min_{\lambda > 0} \frac{\mathbb{E}[e^{\lambda(X-\mu)}]}{e^{\lambda t}}.$$

Proof.

$$\Pr[X - \mu \geq t] = \Pr[e^{X-\mu} \geq e^t] \leq \frac{\mathbb{E}[e^{\lambda(X-\mu)}]}{e^{\lambda t}} \text{ for } \lambda > 0 \text{ by Markov's inequality.}$$

□

Example: Let $x_i \sim [0, 1]$ be independently and identically distributed random variables and $x = \sum_{i \in [n]} x_i$. Then, for any $t \geq 0$,

$$\Pr[x \geq E[x] + t] \leq e^{-2t^2/n}$$
$$\Pr[x \leq E[x] - t] \leq e^{-2t^2/n}$$

These are the *additive* Chernoff bounds.

$$\Pr[x \geq (1 + t)E[x]] \leq e^{-\mathbb{E}[x]t^2/(2+t)}$$
$$\Pr[x \leq (1 - t)E[x]] \leq e^{-\mathbb{E}[x]t^2/2}$$

These are the *multiplicative* Chernoff bounds, which often achieve a better lower bound on probability when $E[x]$ is small.

2 Examples

2.1 Unbiased coin

Flip a fair coin 1000 times. Let X be the number of heads.

- $\mathbb{E}[X] = 1000 \cdot 0.5 = 500$.
- $\Pr[X = 500] = \binom{1000}{500}/2^{1000} \approx 0.025$.
- Question: Pr[X ≥ 600] = ?

  We apply additive Chernoff bounds ($n = 1000, \mu = 500$)

  $$\begin{aligned} \Pr[X \geq 600] &\leq e^{\frac{-2 \cdot 100^2}{1000}} \\ &= e^{-20} \end{aligned}$$

  Note that we can get the same bound with the central limit theorem.

  $$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 250)$$

$$\Pr[X \geq 600] \leq e^{\frac{\left(\frac{100}{\sqrt{250}}\right)^2}{2}}$$
$$= e^{-20}$$

Applying the multiplicative Chernoff bound gives a less precise upper bound.

$$\begin{aligned}\Pr[X \geq 600] &= Pr[X \geq (1+0.2)\cdot 500] \\ &\leq e^{-\frac{0.2^2}{2+0.2}\cdot 500} \\ &= e^{-9.091}\end{aligned}$$

2.2 Biased coin

Suppose n p-biased coins are flipped. Let X denote the number of heads shown, and $\mu := \mathbb{E}[X]$. In this example, we directly compare the bounds received from both additive and multiplicative Chernoff: For any $k \geq 0$,

(Multiplicative Chernoff). $\Pr[X \geq \mu + k] \leq e^{-k^2/(2\mu+k)}$.

(Additive Chernoff). $\Pr[X \geq \mu + k] \leq e^{-2k^2/n}$.

In particular, using the multiplicative bound we have

$$\Pr[X = n] = p^n \leq e^{-\frac{(p-1)^2}{p+1}n},$$

and using the additive bound gives

$$\Pr[X = n] = p^n \leq e^{-2(p-1)^2 n}$$

Proof.

$$\begin{aligned}\Pr[X \geq \mu + k] &= \Pr[X \geq (k/\mu + 1)\mu] \\ &\leq e^{\frac{-\mu(k/\mu)^2}{2+k/\mu}} && \text{(Multiplicative Chernoff bound)} \\ &= e^{\frac{-k^2}{2\mu+k}}.\end{aligned}$$

The tail bounds above comes from the application $k = n - \mu$ and $\mu = np$. □

2.3 Amplification with confidence $1 - \delta$

Suppose we have an algorithm A which solves some optimization problem correctly with probability $\frac{2}{3}$ with no guarantees otherwise. For any $\delta > 0$, the following algorithm solves the optimization problem correctly with probability $1 - \delta$:

Algorithm: Given an input to the optimization problem x, execute $A(x)$ at least $N = 18 \cdot \log(1/\delta)$ times and return the majority output. If there is no majority, return nothing.

Proof. Suppose we run A n times and take the majority and let X be the number of successful times. The probability that the algorithm fails would be

$$\begin{aligned}
&\Pr\left[\text{Amplification algorithm fails}\right] \\
\leq &\Pr\left[X \leq \frac{n}{2}\right] \\
= &\Pr\left[X \leq E\left[X\right] - \frac{n}{6}\right] && (E\left[X\right] = 2n/3) \\
\leq &e^{-\frac{-2\cdot(n/6)^2}{n}} = e^{-\frac{n}{18}} && \text{(Chernoff Bound)}
\end{aligned}$$

Hence, setting

$$n \geq 18 \cdot \log(1/\delta)$$

lets the majority algorithm succeed with probability $1 - \delta$. □

2.4 Biased coin: estimate of p

Suppose biased coin with unknown $p \in [0, 1]$.
How many trials ($n = ???$) are needed to estimate p to some error $\pm\epsilon$, with probability $1 - \delta$?
Given n i.i.d. samples, with $\Pr[x_i] = p$

$$x_1, ...x_n \in \{0, 1\}$$

For some set error ϵ and probability $1 - \delta$, for what n it is thre that

$$\Pr\left[\left|\frac{1}{n}\sum x_i - p\right| < \epsilon\right] \geq 1 - \delta?$$

We can provide an upper bound with additive Chernoff bounds.

$$\begin{aligned}
\Pr\left[\left|\frac{1}{n}\sum x_i - p\right| \geq \epsilon\right] &= \Pr\left[\left|\sum x_i - p\right| \geq \epsilon \cdot n\right] \\
&\leq 2 \cdot e^{-2\cdot\frac{(\epsilon\cdot n)^2}{n}} \\
&= 2 \cdot e^{-2\epsilon^2\cdot n}
\end{aligned}$$

Note that the constant factor of 2 comes from adding the two tails of this distribution

$$\Pr\left[\frac{1}{n}\sum X_i - p > \epsilon\right] + \Pr\left[\frac{1}{n}\sum X_i - p < -\epsilon\right]$$

Now we solve for n. To make

$$\Pr\left[\left|\frac{1}{n}\sum x_i - p\right| \geq \epsilon\right] \leq 2 \cdot e^{-2\epsilon^2\cdot n} \leq \delta,$$

it is enough to set

$$n \geq \frac{1}{2 \cdot \epsilon^2} \log \frac{2}{\delta}$$

We can also provide an estimate using multiplicative Chernoff bounds.

$$\Pr\left[\left|\frac{1}{n}\sum x_i - p\right| \geq \epsilon \cdot p\right] \leq 2 \cdot e^{-\frac{\epsilon^2}{3} \cdot p \cdot n}, \text{ for } \epsilon < 1$$

Substituting ϵp by p, we have

$$\Pr\left[\left|\frac{1}{n}\sum x_i - p\right| \geq \epsilon\right] \leq 2 \cdot e^{-\frac{\epsilon^2}{3} \cdot \frac{n}{p}}, \text{ for } \epsilon < p$$

To let the probability bounded by ϵ, set

$$n \geq \frac{p}{\epsilon^2} \log(\frac{2}{\delta})$$

Note that this is similar to the additive Chernoff bound. However, the multiplicative bound may provide a better upper bound in the case of small enough p.