Useful Inequalities for Randomized Algorithms

Konstantin Makarychev

DRAFT

1 Basic Inequalities

Theorem 1.1 (Markov's Inequality). *Consider a non-negative random variable X. For every positive t, we have*

$$\Pr\{X \geq t\} \leq \frac{\mathbf{E}[X]}{t}.$$

Theorem 1.2 (Chebyshev's Inequality). *Consider an arbitrary random variable X with a finite expectation $\mu = \mathbf{E}[X]$ and finite variance $\mathbf{Var}[X] = \mathbf{E}[(X-\mu)^2]$. For every positive t, we have*

$$\Pr\{|X-\mu| \geq t\} \leq \frac{\mathbf{Var}[X]}{t^2}.$$

Exercise: Prove Markov's and Chebyshev's inequalities.

Theorem 1.3 (Jensen's Inequality). *For every* convex *function $f : \mathbb{R} \to \mathbb{R}$ and random variable X, we have*

$$\mathbf{E}[f(X)] \geq f(\mathbf{E}[X]).$$

For every concave *function $g : \mathbb{R} \to \mathbb{R}$ and random variable X, we have*

$$\mathbf{E}[g(X)] \leq g(\mathbf{E}[X]).$$

Theorem 1.4 (The Cauchy–Schwarz Inequality). *For all random variables X and Y, the following inequality holds:*

$$\mathbf{E}[XY] \leq \sqrt{\mathbf{E}[X^2]\,\mathbf{E}[Y^2]}.$$

2 Bounds on Binomial Coefficients

Claim 2.1. *For all natural n and $k \le n$, we have*

$$\left(\frac{n}{k}\right)^k \le \binom{n}{k} < \left(\frac{en}{k}\right)^k.$$

Proof. We first show the lower bound. Write

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k \cdot (k-1) \cdots 1} = \prod_{i=0}^{k-1} \frac{n-i}{k-i}.$$

Observe that all terms $(n-i)/(k-i)$ are lower bounded by n/k. Thus,

$$\binom{n}{k} = \prod_{i=0}^{k-1} \frac{n-i}{k-i} \ge \left(\frac{n}{k}\right)^k.$$

Now we establish the upper bound. We have

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \le \frac{n^k}{k!}.$$

To finish the proof, we need to show that $k! > (k/e)^k$ (compare this inequality with Stirling's approximation for $k!$). Write the Taylor series for e^k:

$$e^k = 1 + k + \frac{k^2}{2!} + \cdots + \frac{k^k}{k!} + \cdots > \frac{k^k}{k!}.$$

We have $e^k > k^k/k!$ and, consequently, $k! > (k/e)^k$. $\square$

3 Hoeffding's Inequality

Theorem 3.1 (Hoeffding's Inequality)**.** *Let $X_1, \dots, X_n$ be i.i.d.*[1] *Rademacher random variables taking values 1 and -1 with probability $1/2$ i.e.,*

$$\Pr\{X_i = 1\} = \Pr\{X_i = -1\} = 1/2.$$

Then, for all $t \ge 0$, we have

$$\Pr\left\{\sum_i X_i \ge t\right\} \le e^{-\frac{t^2}{2n}}.$$

[1] i.i.d. stands for independent identically distributed

Proof. We first show the following lemma.

Lemma 3.2 (Bernstein)**.** *Let $X_1, \ldots, X_n$ be independent random variables. For all $\lambda > 0$ and $t \geq 0$, we have*

$$\Pr\Big\{\sum_i X_i \geq t\Big\} \leq \frac{\prod_i \mathbf{E}\big[\,e^{\lambda X_i}\,\big]}{e^{\lambda t}}.$$

Proof of Lemma 3.2. Let $f(x) = e^{\lambda x}$ and $S = \sum_i X_i$. Observe that f is a monotonically increasing non-negative function. Thus, $x \geq t$ if and only if $f(x) \geq f(t)$. In particular, $S \geq t$ if and only if $f(S) \geq f(t)$. Thus, by Markov's inequality applied to the random variable $f(S)$, we have

$$\Pr\{S \geq t\} = \Pr\{f(S) \geq f(t)\} \leq \frac{\mathbf{E}[f(S)]}{f(t)}.$$

Write,

$$\mathbf{E}[f(S)] = \mathbf{E}\Big[\exp\big(\lambda \sum_i X_i\big)\Big] = \mathbf{E}\Big[\prod_i \exp\big(\lambda X_i\big)\Big].$$

Random variables $\exp\big(\lambda X_i\big)$ $(i \in \{1, \ldots, n\})$ are independent, hence

$$\mathbf{E}\Big[\prod_i \exp\big(\lambda X_i\big)\Big] = \prod_i \mathbf{E}\Big[\exp\big(\lambda X_i\big)\Big].$$

Thus,

$$\Pr\{S \geq t\} \leq \frac{\prod_i \mathbf{E}\Big[\exp\big(\lambda X_i\big)\Big]}{f(t)}.$$

This concludes the proof. $\square$

We now use Lemma 3.2 to prove Hoeffding's inequality. To this end, we compute the expectation $\mathbf{E}[\exp(\lambda X_i)]$ for each i:

$$\mathbf{E}\Big[\exp\big(\lambda X_i\big)\Big] = \Pr\{X_i = 1\} \cdot e^{\lambda} + \Pr\{X_i = -1\} \cdot e^{-\lambda} = \frac{e^{\lambda} + e^{-\lambda}}{2}.$$

The function on the right hand side is called the hyperbolic cosine and denoted by $\cosh x$: $\cosh x = (e^{\lambda} + e^{-\lambda})/2$. We use the following simple bound on $\cosh x$.

Claim 3.3. *For all λ, we have*

$$\frac{e^{\lambda} + e^{-\lambda}}{2} \leq e^{\lambda^2/2}$$

We prove this claim below and now proceed with the proof of Hoeffding's inequality. By Claim 3.3:

$$\mathbf{E}\Big[\exp\big(\lambda X_i\big)\Big] \le e^{\lambda^2/2}.$$

Thus, by Lemma 3.2,

$$\Pr\Big\{\sum_i X_i \ge t\Big\} \le \frac{\prod_i e^{-\lambda^2/2}}{e^{\lambda t}} = e^{\lambda^2 n/2 - \lambda t}.$$

For $\lambda = t/n$, we get the desired bound. To finish the proof we need to establish Claim 3.3.

Proof of Claim 3.3. Write the Taylor series for functions $\cosh\lambda$ and $e^{\lambda^2/2}$:

$$\begin{aligned}\frac{e^{\lambda}+e^{-\lambda}}{2} &= 1 + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^{2i}}{(2i)!} + \dots\\ e^{\lambda^2/2} &= 1 + \frac{\lambda^2}{2} + \cdots + \frac{\lambda^{2i}}{2^i \cdot i!} + \dots\end{aligned}$$

Observe that $(2i)! \ge 2^i \cdot i!$. Thus, the i-th term in the first series is less than or equal to the i-th term in the second series for each i. Therefore, we have $\cosh\lambda \le e^{\lambda^2/2}$. □

□

Corollary 3.4 (Symmetric Hoeffding's Inequality)**.** *Let $X_1, \dots, X_n$ be i.i.d. symmetric Bernoulli random variables taking values 1 and -1 with probability $1/2$ i.e.,*

$$\Pr\{X_i = 1\} = \Pr\{X_i = -1\} = 1/2.$$

Then, for all $t \ge 0$, we have

$$\Pr\Big\{\Big|\sum_i X_i\Big| \ge t\Big\} \le 2e^{-\frac{t^2}{2n}}.$$

Proof. The random variable $S = \sum_i X_i$ is symmetric around 0, and consequently for every t we have $\Pr\{S \ge t\} = \Pr\{S \le -t\}$. Thus,

$$\Pr\Big\{\sum_i X_i \le -t\Big\} = \Pr\Big\{\sum_i X_i \ge t\Big\} \le e^{-\frac{t^2}{2n}}.$$

Thus,

$$\Pr\Big\{\Big|\sum_i X_i\Big| \ge t\Big\} = \Pr\Big\{\sum_i X_i \le -t\Big\} + \Pr\Big\{\sum_i X_i \ge t\Big\} \le 2e^{-\frac{t^2}{2n}}.$$

□

We now state a more general variant of Hoeffding's Inequality (without a proof).

Theorem 3.5 (Hoeffding's Inequality)**.** *Let $X_1, \dots, X_n$ be independent random variables. Suppose that each X_i takes values in the interval $[m_i, M_i]$. Let $\mu = \mathbf{E}\big[\sum_i X_i\big]$. Then, for all $t \ge 0$, we have*

$$\Pr\Big\{\Big|\sum_i X_i - \mu\Big| \ge t\Big\} \le 2e^{-\frac{2t^2}{\sum(M_i - m_i)^2}}.$$

4 Chernoff Bound

Theorem 4.1 (The Chernoff Bound)**.** *Consider independent random variables $X_1, \dots, X_n$ taking values in the interval $[0, 1]$. Let $\mu_i = \mathbf{E}[X_i]$ and $\mu = \sum_{i=1}^n \mu_i$. Then,*

$$\Pr\Big\{\sum_{i=1}^n X_i \ge t\Big\} \le e^{-\mu}\Big(\frac{e\mu}{t}\Big)^t.$$

Proof. Fix a positive λ. As in the proof of Hoeffding's inequality, we first upper bound $\mathbf{E}\big[\, e^{\lambda X_i}\,\big]$ for each i. Since $x \mapsto e^{\lambda x}$ is a convex function, the following inequality holds for all $x \in [0, 1]$:

$$e^{\lambda x} \le xe^{\lambda} + (1-x)e^0 = xe^{\lambda} + (1-x) = 1 + x(e^{\lambda} - 1).$$

Thus,

$$\mathbf{E}\big[\, e^{\lambda X_i}\,\big] \le \mathbf{E}\Big[1 + X_i(e^{\lambda} - 1)\Big] = 1 + \mu_i(e^{\lambda} - 1) \le e^{\mu_i(e^{\lambda}-1)}.$$

By Lemma 3.2,

$$\begin{aligned}\Pr\Big\{\sum_{i=1}^n X_i \ge t\Big\} &\le \frac{\prod_i \mathbf{E}\big[\, e^{\lambda X_i}\,\big]}{e^{\lambda t}} \le \frac{\prod_i \exp(\mu_i(e^{\lambda} - 1))}{e^{\lambda t}} \\ &= \frac{\exp\big(\sum_i \mu_i(e^{\lambda} - 1)\big)}{e^{\lambda t}} = \frac{e^{\mu(e^{\lambda}-1)}}{e^{\lambda t}}.\end{aligned}$$

For $\lambda = \ln(t/\mu)$, we get

$$\Pr\left\{\sum_{i=1}^{n} X_i \geq t\right\} \leq e^{t-\mu}\left(\frac{\mu}{t}\right)^t = e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

□

It is easy to use this form of the Chernoff bound in this form when $t \gg \mu$. We now derive a simpler – but less precise – upper bound for $t = (1+\delta)\mu$, $\delta > 0$. The right hand side of the inequality equals

$$e^{-\mu}\left(\frac{e\mu}{t}\right)^t = e^{-\mu}\left(\frac{e}{(1+\delta)}\right)^{(1+\delta)\mu} = \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu.$$

We estimate the term in the brackets, $e^\delta/(1+\delta)^{1+\delta}$, as follows: $\frac{e^\delta}{(1+\delta)^{1+\delta}} \leq e^{\frac{-\delta^2}{2+\delta}}$ (prove this bound!) and get the following version of the Chernoff Bound.

Corollary 4.2 (The Chernoff bound)**.** *Consider independent random variables $X_1, \dots, X_n$ taking values $\{0,1\}$. Let $\mu_i = \mathbf{E}[X_i]$ and $\mu = \sum_{i=1}^n \mu_i$. Then,*

$$\Pr\left\{\sum_{i=1}^{n} X_i \geq (1+\delta)\mu\right\} \leq e^{\frac{-\delta^2\mu}{2+\delta}}.$$

Moreover for $\delta \in [0,1]$, we have

$$\Pr\left\{\sum_{i=1}^{n} X_i \geq (1+\delta)\mu\right\} \leq e^{-\delta^2\mu/3};$$
$$\Pr\left\{\sum_{i=1}^{n} X_i \leq (1-\delta)\mu\right\} \leq e^{-\delta^2\mu/3};$$