CMPSCI 611: Advanced Algorithms

Lecture 15: Tail Inequalities

Andrew McGregor

Last Compiled: April 11, 2024

Variance Refresher

Variance Refresher

- Expectation: $\mathbb{E}[X] = \sum_r r\mathbb{P}[X = r]$

Variance Refresher

- Expectation: $\mathbb{E}[X] = \sum_r r\mathbb{P}[X = r]$
- Linearity of expectation: $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Variance Refresher

- Expectation: $\mathbb{E}[X] = \sum_r r\mathbb{P}[X = r]$
- Linearity of expectation: $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- Variance random variable: $\mathbb{V}[X] = \sigma_X^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$

Variance Refresher

- Expectation: $\mathbb{E}[X] = \sum_r r\mathbb{P}[X = r]$
- Linearity of expectation: $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- Variance random variable: $\mathbb{V}[X] = \sigma_X^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$
- Linearity of variance if X and Y are independent:

$$\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y]$$

Examples of Random Variables

Example

Let X have the binomial distribution $Bin(n, p)$:

$$\mathbb{P}[X = i] = \binom{n}{i} p^i (1-p)^{n-i}$$

"How many heads do we see when we toss a coin with probability p of heads n times?" $\mathbb{E}[X] = np$ and $\mathbb{V}[X] = np(1-p)$.

Examples of Random Variables

Example

Let X have the binomial distribution $Bin(n, p)$:

$$\mathbb{P}[X = i] = \binom{n}{i} p^i (1-p)^{n-i}$$

"How many heads do we see when we toss a coin with probability p of heads n times?" $\mathbb{E}[X] = np$ and $\mathbb{V}[X] = np(1-p)$.

Example

Let X have the geometric distribution $Geom(p)$:

$$\mathbb{P}[X = i] = (1-p)^{i-1} p$$

"How many times do we toss a coin with probability p of heads until we see a heads." $\mathbb{E}[X] = 1/p$, $\mathbb{V}[X] = (1-p)/p^2$.

Outline

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \ .$$

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \ .$$

Proof.

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \ .$$

Proof.

- $\mathbb{E}[Y] = \sum_r r \cdot \mathbb{P}[Y = r]$

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \; .$$

Proof.

- $\mathbb{E}[Y] = \sum_r r \cdot \mathbb{P}[Y = r] \geq \sum_{r \geq t\mu} r \cdot \mathbb{P}[Y = r]$

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \ .$$

Proof.

- $\mathbb{E}[Y] = \sum_r r \cdot \mathbb{P}[Y = r] \geq \sum_{r \geq t\mu} r \cdot \mathbb{P}[Y = r] \geq \mathbb{P}[Y \geq t\mu] \cdot t \cdot \mu$

Markov Inequality

Theorem (Markov)

Let Y be a positive random variable and let $\mu = \mathbb{E}[Y]$ be strictly positive. Then, for $t > 0$,

$$\mathbb{P}[Y \geq t\mu] \leq 1/t \ .$$

Proof.

- $\mathbb{E}[Y] = \sum_r r \cdot \mathbb{P}[Y = r] \geq \sum_{r \geq t\mu} r \cdot \mathbb{P}[Y = r] \geq \mathbb{P}[Y \geq t\mu] \cdot t \cdot \mu$
- Therefore, $\mathbb{P}[Y \geq t\mu] \leq 1/t$.

□

Chebyshev Inequality

Theorem (Chebyshev)

Let X be a random variable with expectation μ and variance σ^2 that is strictly positive. Then for $t > 0$,

$$\mathbb{P}\left[|X - \mu| \geq t\sigma\right] \leq 1/t^2 \ .$$

Chebyshev Inequality

Theorem (Chebyshev)

Let X be a random variable with expectation μ and variance σ^2 that is strictly positive. Then for $t > 0$,

$$\mathbb{P}\left[|X - \mu| \geq t\sigma\right] \leq 1/t^2 \ .$$

Proof.

- Note that $\mathbb{P}\left[|X - \mu| \geq t\sigma\right] = \mathbb{P}\left[(X - \mu)^2 \geq t^2\sigma^2\right]$

□

Chebyshev Inequality

Theorem (Chebyshev)

Let X be a random variable with expectation μ and variance σ^2 that is strictly positive. Then for $t > 0$,

$$\mathbb{P}\left[|X - \mu| \geq t\sigma\right] \leq 1/t^2 \ .$$

Proof.

- Note that $\mathbb{P}\left[|X - \mu| \geq t\sigma\right] = \mathbb{P}\left[(X - \mu)^2 \geq t^2\sigma^2\right]$
- Let $Y = (X - \mu)^2$ and note $\mathbb{E}\left[Y\right] = \sigma^2$

□

Chebyshev Inequality

Theorem (Chebyshev)

Let X be a random variable with expectation μ and variance σ^2 that is strictly positive. Then for $t > 0$,

$$\mathbb{P}\left[|X - \mu| \geq t\sigma\right] \leq 1/t^2 \ .$$

Proof.

- Note that $\mathbb{P}\left[|X - \mu| \geq t\sigma\right] = \mathbb{P}\left[(X - \mu)^2 \geq t^2\sigma^2\right]$
- Let $Y = (X - \mu)^2$ and note $\mathbb{E}\left[Y\right] = \sigma^2$
- Use Markov's inequality to show $\mathbb{P}\left[Y \geq t^2\mathbb{E}\left[Y\right]\right] \leq 1/t^2$

□

Outline

Chernoff Bound

Theorem

Let $X_1, \ldots, X_n$ be independent boolean random variables such that $\mathbb{P}[X_i = 1] = p_i$. Then, for $X = \sum_i X_i$, $\mu = \mathbb{E}[X]$, and $\delta > 0$,

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

Chernoff Bound

Theorem

Let $X_1, \ldots, X_n$ be independent boolean random variables such that $\mathbb{P}[X_i = 1] = p_i$. Then, for $X = \sum_i X_i$, $\mu = \mathbb{E}[X]$, and $\delta > 0$,

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

Other versions: For $0 < \delta \leq 1$

$$\mathbb{P}[X \geq (1+\delta)\mu] \leq e^{-\delta^2\mu/3}$$

$$\mathbb{P}[X \leq (1-\delta)\mu] \leq e^{-\delta^2\mu/2}$$

and so by the union bound, $\mathbb{P}[|X-\mu| \geq \delta\mu] \leq 2e^{-\delta^2\mu/3}$.

Chernoff Bound: Proof of Upper Tail (1/2)

Proof.

- For any $t > 0$: $\mathbb{P}[X > (1+\delta)\mu] = \mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right]$

Chernoff Bound: Proof of Upper Tail (1/2)

Proof.

- For any $t > 0$: $\mathbb{P}[X > (1+\delta)\mu] = \mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right]$
- Apply Markov inequality:

$$\mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right] < \mathbb{E}\left[e^{tX}\right] / e^{t(1+\delta)\mu}$$

Chernoff Bound: Proof of Upper Tail (1/2)

Proof.

- For any $t > 0$: $\mathbb{P}[X > (1+\delta)\mu] = \mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right]$
- Apply Markov inequality:

$$\mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right] < \mathbb{E}\left[e^{tX}\right]/e^{t(1+\delta)\mu}$$

- By independence:

$$\mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t\sum_i X_i}\right] = \mathbb{E}\left[\prod_i e^{tX_i}\right] = \prod_i \mathbb{E}\left[e^{tX_i}\right]$$

Chernoff Bound: Proof of Upper Tail (1/2)

Proof.

- For any $t > 0$: $\mathbb{P}[X > (1+\delta)\mu] = \mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right]$
- Apply Markov inequality:

$$\mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right] < \mathbb{E}\left[e^{tX}\right]/e^{t(1+\delta)\mu}$$

- By independence:

$$\mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t\sum_i X_i}\right] = \mathbb{E}\left[\prod_i e^{tX_i}\right] = \prod_i \mathbb{E}\left[e^{tX_i}\right]$$

- We will prove $\prod_i \mathbb{E}\left[e^{tX_i}\right] \leq e^{(e^t-1)\mu}$ in the next slide.

□

Chernoff Bound: Proof of Upper Tail (1/2)

Proof.

- For any $t > 0$: $\mathbb{P}[X > (1+\delta)\mu] = \mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right]$
- Apply Markov inequality:

$$\mathbb{P}\left[e^{tX} > e^{t(1+\delta)\mu}\right] < \mathbb{E}\left[e^{tX}\right]/e^{t(1+\delta)\mu}$$

- By independence:

$$\mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t\sum_i X_i}\right] = \mathbb{E}\left[\prod_i e^{tX_i}\right] = \prod_i \mathbb{E}\left[e^{tX_i}\right]$$

- We will prove $\prod_i \mathbb{E}\left[e^{tX_i}\right] \le e^{(e^t-1)\mu}$ in the next slide.
- For $t = \ln(1+\delta)$:

$$\mathbb{E}\left[e^{tX}\right]/e^{t(1+\delta)\mu} \le e^{(e^t-1)\mu}/e^{t(1+\delta)\mu} = \left[\frac{e^\delta}{(1+\delta)^{1+\delta}}\right]^\mu$$

□

Chernoff Bound: Proof of Upper Tail (2/2)

Lemma

$\prod_i \mathbb{E}\left[e^{tX_i}\right] \leq e^{(e^t-1)\mu}$

Proof.

- Using $1 + x \leq e^x$:

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

□

Chernoff Bound: Proof of Upper Tail (2/2)

Lemma

$\prod_i \mathbb{E}\left[e^{tX_i}\right] \le e^{(e^t-1)\mu}$

Proof.

- Using $1 + x \le e^x$:

$$\mathbb{E}\left[e^{tX_i}\right] = p_i e^t + (1 - p_i) = 1 + p_i(e^t - 1) \le \exp(p_i(e^t - 1))$$

- Using $\mu = \mathbb{E}\left[\sum_i X_i\right] = \sum_i p_i$:

$$\prod_i \exp(p_i(e^t - 1)) = \exp(\sum_i p_i(e^t - 1)) = \exp((e^t - 1)\mu)$$

□

Outline

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

1. *Add each element in S to a set R with probability $p = 1/n^{1/4}$.*

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

1. *Add each element in S to a set R with probability $p = 1/n^{1/4}$.*
2. *Call this set R, Sort R and let*

$$a = \ (n^{3/4}/2 - 5\sqrt{n}) \text{ smallest element in } R.$$

$$b = \ (n^{3/4}/2 + 5\sqrt{n}) \text{ smallest element in } R.$$

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

1. *Add each element in S to a set R with probability $p = 1/n^{1/4}$.*
2. *Call this set R, Sort R and let*

$$a = (n^{3/4}/2 - 5\sqrt{n}) \text{ smallest element in } R.$$

$$b = (n^{3/4}/2 + 5\sqrt{n}) \text{ smallest element in } R.$$

3. *Construct $S' = \{i \in S : a < y < b\}$ and let t be the number of values less or equal to a amongst S.*

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

1. *Add each element in S to a set R with probability $p = 1/n^{1/4}$.*
2. *Call this set R, Sort R and let*

$$a = \ (n^{3/4}/2 - 5\sqrt{n}) \textit{ smallest element in } R.$$

$$b = \ (n^{3/4}/2 + 5\sqrt{n}) \textit{ smallest element in } R.$$

3. *Construct $S' = \{i \in S : a < y < b\}$ and let t be the number of values less or equal to a amongst S.*
4. *Sort S' and return $(k - t)$th smallest value in S'.*

Lazy Select

Let S be set of $n = 2k$ distinct values. Want to find k-th smallest value.

Algorithm

1. *Add each element in* S *to a set* R *with probability* $p = 1/n^{1/4}$.
2. *Call this set* R*, Sort* R *and let*

$$a = \ (n^{3/4}/2 - 5\sqrt{n}) \textit{ smallest element in } R.$$

$$b = \ (n^{3/4}/2 + 5\sqrt{n}) \textit{ smallest element in } R.$$

3. *Construct* $S' = \{i \in S : a < y < b\}$ *and let* t *be the number of values less or equal to* a *amongst* S.
4. *Sort* S' *and return* $(k - t)$*th smallest value in* S'.

We'll analyze it next time...