CSCE 658: Randomized Algorithms

Lecture 5

Samson Zhou

Recall: Moments

- For $p > 0$, the p-th moment of a random variable X over Ω is:

$$\mathrm{E}[X^p] = \sum_{x \in \Omega} \Pr[X = x] \cdot x^p$$

Last Time: Chebyshev’s Inequality

- Let X be a random variable with expected value $\mu \coloneqq \mathrm{E}[X]$ and variance $\sigma^2 \coloneqq \mathrm{Var}[X]$
- $\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$ becomes $\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\sigma^2}{t^2}$

$$\Pr[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

- “Bounding the deviation of a random variable in terms of its variance”

Last Time: Accuracy Boosting

- Algorithmic consequence of Law of Large Numbers

- To improve the accuracy of your algorithm, run it many times independently and take the average

Recall: Concentration Inequalities

- Concentration inequalities bound the probability that a random variable is “far away” from its expectation

- Looking at the k^{th} moment for sufficiently high k gives a number of very strong (and useful!) concentration inequalities with exponential tail bounds

- Chernoff bounds, Bernstein’s inequality, Hoeffding’s inequality, etc.

Limitations

- Suppose we flip a fair coin $n = 100$ times and let H be the total number of heads
- $\mathrm{E}[H] = 50$ and $\mathrm{Var}[H] = 25$

- Markov's inequality: $\Pr[H \geq 60] \leq 0.833$
- Chebyshev's inequality: $\Pr[H \geq 60] \leq 0.25$
- Truth: $\Pr[H \geq 60] \approx 0.0284$

Intuition for Previous Inequalities

- Recall: We proved Markov's inequality by looking at the first moment of the random variable X

$$\Pr[X \geq t \cdot \mathrm{E}[X]] \leq \frac{1}{t}$$

- Recall: We proved Chebyshev's inequality by applying Markov to the second moment of the random variable $X - \mathrm{E}[X]$

$$\Pr[|X - \mathrm{E}[X]| \geq t] = \Pr[|X - \mathrm{E}[X]|^2 \geq t^2] \leq \frac{\mathrm{Var}[X]}{t^2}$$

Generalizations

- Suppose we flip a fair coin $n = 100$ times and let H be the total number of heads

- What if we consider higher moments?
- Looking at the 4^{th} moment: $\Pr[H \geq 60] \leq 0.186$

- Markov's inequality: $\Pr[H \geq 60] \leq 0.833$
- Chebyshev's inequality: $\Pr[H \geq 60] \leq 0.25$
- Truth: $\Pr[H \geq 60] \approx 0.0284$

Concentration Inequalities

- Looking at the k^{th} moment for sufficiently high k gives a number of very strong (and useful!) concentration inequalities with exponential tail bounds

- Chernoff bounds, Bernstein's inequality, Hoeffding's inequality, etc.

Bernstein's Inequality

- Bernstein's inequality: Let $X_1, \ldots, X_n \in [-M, M]$ be independent random variables and let $X = X_1 + \cdots + X_n$ have mean μ and variance σ^2. Then for any $t \geq 0$:

$$\Pr[|X - \mu| \geq t] \leq 2e^{-\frac{t^2}{2\sigma^2 + \frac{4}{3}Mt}}$$

Bernstein's Inequality

- Bernstein's inequality: Let $X_1, \dots, X_n \in [-M, M]$ be independent random variables and let $X = X_1 + \cdots + X_n$ have mean μ and variance σ^2. Then for any $t \geq 0$:

$$\Pr[|X - \mu| \geq t] \leq 2e^{-\frac{t^2}{2\sigma^2 + \frac{4}{3}Mt}}$$

- Example: Suppose $M = 1$ and let $t = k\sigma$. Then

$$\Pr[|X - \mu| \geq k\sigma] \leq 2\exp\left(-\frac{k^2}{4}\right)$$

Bernstein's Inequality

- Suppose $M = 1$ and let $t = k\sigma$. Then

$$\Pr[|X - \mu| \geq k\sigma] \leq 2\exp\left(-\frac{k^2}{4}\right)$$

- Compare to Chebyshev's inequality:

$$\Pr[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

- Exponential improvement!

Bernstein's Inequality

- Suppose we flip a fair coin $n = 100$ times and let H be the total number of heads

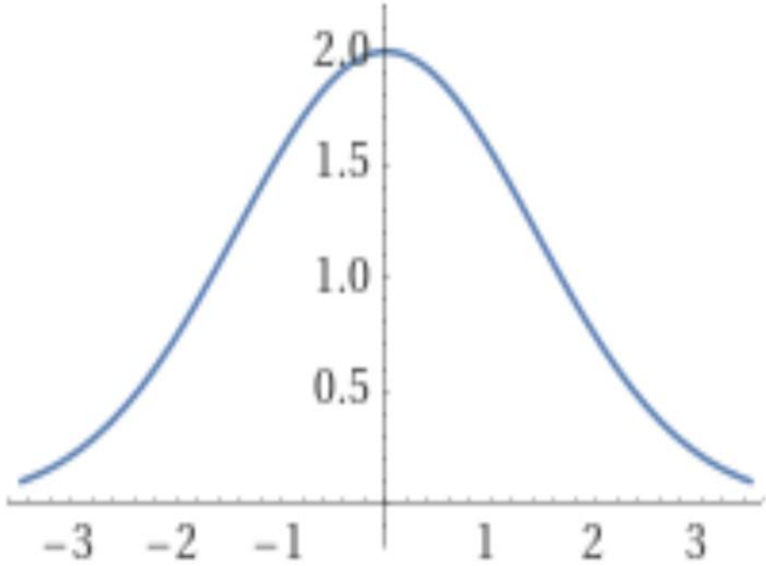
- Markov's inequality: $\Pr[H \geq 60] \leq 0.833$
- Chebyshev's inequality: $\Pr[H \geq 60] \leq 0.25$
- 4^{th} moment: $\Pr[H \geq 60] \leq 0.186$
- Bernstein's inequality: $\Pr[H \geq 60] \leq 0.15$
- Truth: $\Pr[H \geq 60] \approx 0.0284$

Bernstein's Inequality

- Suppose $M = 1$ and let $t = k\sigma$. Then

$$\Pr[|X-\mu| \geq k\sigma] \leq 2\exp\left(-\frac{k^2}{4}\right)$$

- Plot across values of k looks like normal random variable



- PDF of Gaussian $\mathrm{N}(0, \sigma^2)$ is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Central Limit Theorem

- Stronger Central Limit Theorem: The distribution of the sum of n bounded independent random variables converges to a Gaussian (normal) distribution as n goes to infinity

- Why is the Gaussian distribution is so important in statistics, data science, ML, etc.?

- Many random variables can be approximated as the sum of a large number of small and roughly independent random effects. Thus, their distribution looks Gaussian by CLT.

Trivia Question #3 (Max Load)

- Suppose we have a fair n-sided die that we roll n times. "On average", what is the largest number of times any outcome is rolled? Example: $1, 5, 2, 4, 1, 3, 1$ for $n = 7$

- $\Theta(1)$
- $\tilde{\Theta}(\log n)$
- $\tilde{\Theta}(\sqrt{n})$
- $\tilde{\Theta}(n)$

Trivia Question #4 (Coupon Collector)

- Suppose we have a fair n-sided die. “On average”, how many times should we roll the die before we see all possible outcomes among the rolls? Example: $1, 5, 2, 4, 1, 3, 1, 6$ for $n = 6$

- $\Theta(n)$
- $\Theta(n \log n)$
- $\Theta(n\sqrt{n})$
- $\Theta(n^2)$

Chernoff Bounds

- Useful variant of Bernstein's inequality when the random variables are binary
- Chernoff bounds: Let $X_1, \ldots, X_n \in \{0, 1\}$ be independent random variables and let $X = X_1 + \cdots + X_n$ have mean μ. Then for any $\delta \geq 0$:

$$\Pr[|X - \mu| \geq \delta\mu] \leq 2\exp\left(-\frac{\delta^2\mu}{2+\delta}\right)$$

Multiplicative Error Chernoff Bounds

- Chernoff bounds: Let $X_1, \ldots, X_n \in \{0, 1\}$ be independent random variables and let $X = X_1 + \cdots + X_n$ have mean μ. For $\delta \in (0,1)$:

$$\Pr[X \geq (1+\delta)\mu] \leq 2\exp\left(-\frac{\delta^2 \mu}{2+\delta}\right)$$

$$\Pr[X \leq (1-\delta)\mu] \leq \exp\left(-\frac{\delta^2 \mu}{2}\right)$$

$$\Pr[|X-\mu| \geq \delta\mu] \leq 2\exp\left(-\frac{\delta^2 \mu}{3}\right)$$

Use Case

- Suppose we design a randomized algorithm A that outputs a real number Z that is "correct" with probability $\frac{2}{3}$, e.g., $Z \in \{0,1\}$
- Suppose we want to be correct with probability 0.999 or $1 - \frac{1}{n^2}$ or $1 - \delta$
- What can we do?

Success Boosting

- Chernoff bounds: Run the algorithm A a total of $O\left(\log\frac{1}{\delta}\right)$ times and take the median. It will be correct with probability $1-\delta$

Median-of-Means Framework

- Suppose we design a randomized algorithm A to estimate a hidden statistic Θ of a dataset and we know $0 < \Theta \leq 1000$.
- Suppose each time we use the algorithm A, it outputs a number X such that $\mathrm{E}[X] = \Theta$ and $\mathrm{Var}[X] = 100\Theta^2$
- Suppose we want to estimate Θ to accuracy ε, with probability $1 - \delta$

Median-of-Means Framework

- Suppose we design a randomized algorithm A to estimate a hidden statistic Θ of a dataset and we know $0 < \Theta \leq 1000$.
- Suppose each time we use the algorithm A, it outputs a number X such that $\mathrm{E}[X] = \Theta$ and $\mathrm{Var}[X] = 100\Theta^2$
- Suppose we want to estimate Θ to accuracy ε, with probability $1 - \delta$

- Accuracy boosting: Repeat A a total of $\frac{10^{12}}{\varepsilon^2}$ times and take the mean
- Success boosting: Find the mean a total of $O\left(\log\frac{1}{\delta}\right)$ times and take the median, to be correct with probability $1 - \delta$

Max Load

- Suppose we have a fair n-sided die that we roll n times. "On average", what is the largest number of times any outcome is rolled? Example: $1, 5, 2, 4, 1, 3, 1$ for $n = 7$

- Fix a value $k \in [n]$
- Let $X_i = 1$ if the i-th roll is k and $X_i = 0$ otherwise

- $\mathrm{E}[X_i] = \frac{1}{n}$

Max Load

- The total number of rolls with value k is $X = X_1 + \cdots + X_n$
- $\mathrm{E}[X] = 1$
- Recall Chernoff bounds:

$$\Pr[X \geq (1+\delta)\mu] \leq 2\exp\left(-\frac{\delta^2\mu}{2+\delta}\right)$$

- $\Pr[X \geq 3\log n] \leq \frac{1}{n^2}$

Max Load

- Recall we fixed a value $k \in [n]$
- $\Pr[X \geq 3 \log n] \leq \frac{1}{n^2}$ means that with probability at least $1 - \frac{1}{n^2}$, we will get fewer than $3 \log n$ rolls with value k

- Union bound: With probability at least $1 - \frac{1}{n}$, no outcome will be rolled more than $3 \log n$ times

Trivia Question #3 (Max Load)

- Suppose we have a fair n-sided die that we roll n times. “On average”, what is the largest number of times any outcome is rolled? Example: $1, 5, 2, 4, 1, 3, 1$ for $n = 7$

- $\Theta(1)$
- $\widetilde{\Theta}(\log n)$
- $\widetilde{\Theta}(\sqrt{n})$
- $\widetilde{\Theta}(n)$

Coupon Collector

- Suppose we have a fair n-sided die. "On average", how many times should we roll the die before we see all possible outcomes among the rolls? Example: 1, 5, 2, 4, 1, 3, 1, 6 for $n = 6$

- Consider r rolls
- Fix a specific outcome $k \in [n]$
- Let $X_i = 1$ if the i-th roll is k and $X_i = 0$ otherwise
- $\mathrm{E}[X_i] = \frac{1}{n}$

Coupon Collector

- The total number of rolls with value k is $X = X_1 + \cdots + X_r$
- $\mathrm{E}[X] = \frac{r}{n} = 6 \log n$ for $r = 6n \log n$
- Recall Chernoff bounds:

$$\Pr[X \leq (1-\delta)\mu] \leq \exp\left(-\frac{\delta^2 \mu}{2}\right)$$

- $\Pr[X \leq \log n] \leq \frac{1}{n^2}$

Coupon Collector

- Recall we fixed a value $k \in [n]$
- $\Pr[X \leq \log n] \leq \frac{1}{n^2}$ means that with probability at least $1 - \frac{1}{n^2}$, we will at least $\log n$ rolls with value k

- Union bound: With probability at least $1 - \frac{1}{n}$, all outcomes will be rolled at least $\log n$ times

Trivia Question #4 (Coupon Collector)

- Suppose we have a fair n-sided die. "On average", how many times should we roll the die before we see all possible outcomes among the rolls? Example: $1, 5, 2, 4, 1, 3, 1, 6$ for $n = 6$

- $\Theta(n)$
- $\Theta(n \log n)$
- $\Theta(n\sqrt{n})$
- $\Theta(n^2)$