CS265, Fall 2023

Class 9: Agenda and Questions

1 Announcements

- HW4 due Friday!

2 Lecture Recap and Questions?

Questions from minilectures and pre-class quiz? (Compressed sensing; RIP; Gaussian matrices have the RIP with high probability.)

3 More practice with tail bounds; compare and contrast

For this part of the class, let's review some of the tail bounds we've seen. We'll explore when we'd like to use which.

We'll look at the following types of bounds:

1. Markov: For any non-negative random variable X, $\Pr\{X \geq t\} \leq \frac{\mathbb{E}X}{t}$.
2. Chebyshev: For any random variable X, $\Pr\{|X - \mathbb{E}X| \geq t\} \leq \frac{\mathrm{Var}(X)}{t^2}$.
3. Higher moment bounds: For any random variable X and any $c \geq 1$, $\Pr\{|X| \geq t\} \leq \frac{\mathbb{E}[|X|^c]}{t^c}$.
4. Multiplicative Chernoff bound: If $X_i \sim \mathrm{Ber}(p)$ are i.i.d. and $X = \sum_{i=1}^n X_i$, then for any $\varepsilon \in (0,1)$, $\Pr\{X \geq \mu(1+\varepsilon)\} \leq \exp(-C\varepsilon^2\mu)$ (for some constant C).
5. Additive Chernoff bound: If $X_i \sim \mathrm{Ber}(p)$ are i.i.d. and $X = \sum_{i=1}^n X_i$, then for any $\varepsilon > 0$, $\Pr\{X \geq \mu + t\} \leq \exp(-Ct^2/n)$ (for some constant C).

These are not all of the bounds we've seen (in particular, we've seen some other forms of Chernoff bounds), but hopefully this will be enough to get our intuition going.

3.1 Markov vs. Chebyshev vs. higher moments

Group Work

For this group work, define $\zeta(k) := \sum_{i=1}^{\infty} i^{-k}$. Recall that $\zeta(1) = \infty$ (that is, the sum diverges), while $\zeta(k) < \infty$ for all $k \geq 2$. (For example, $\zeta(2) = \pi^2/6$).

For this problem, it doesn't really matter what the $\zeta(k)$ are, just that they are constants (for $k \geq 2$) and that they are decreasing as $k \to \infty$.

1. Let X be a random variable so that, for each $i \geq 1$,

$$\Pr[X = i] = \frac{1}{\zeta(3)} \cdot \frac{1}{i^3}.$$

   - What is $\mathbb{E}X$?
   - Which of our tail bounds apply to an expression like $\Pr[X \geq t] \leq$ ____?
   - For all the tail bounds that apply, what do they say about how to fill in the blank? Use asymptotic notation as $t \to \infty$. (e.g., your answer should be like $O(1/t)$ or $O(1/t^2)$ or $e^{-\Omega(t)}$ or something like that).

2. Let X be a random variable so that, for each $i \geq 1$,

$$\Pr[X = i] = \frac{1}{\zeta(4)} \cdot \frac{1}{i^4}.$$

   Answer the same questions as in the previous part.

3. Let X be a random variable so that, for each $i \geq 1$,

$$\Pr[X = i] = \frac{1}{\zeta(k)} \cdot \frac{1}{i^k}.$$

   Answer the same questions as in the previous part.

4. What do you take away about Markov vs Chebyshev vs higher moment bounds? (e.g., when would you want to use one vs the others?)

3.2 Chernoff vs Chebyshev

Group Work

For this group work, say that $X_i \sim \mathrm{Ber}(p)$ are i.i.d., and let $X = \sum_{i=1}^n X_i$.

1. Suppose that $p = 1/4$, and let $\varepsilon \in (0, 1)$. What do each of the Chernoff bounds, and Chebyshev's inequality, say about $\Pr[X \geq \mu(1 + \varepsilon)]$, where $\mu = \mathbb{E}X = pn$? Give your answer in big-O notation as $n \to \infty$ and $\varepsilon \to 0$. (e.g., an expression like $O(1/(n \cdot \varepsilon))$ or $e^{-\Omega(n/\varepsilon)}$).

2. Suppose that $p = 1/4$. How big does ε have to be (asymptotically, in terms of n) to prove a statement like

$$\Pr[X \geq \mu(1 + \varepsilon)] \leq 0.001?$$

What do each {Chebyshev, multiplicative chernoff, additive chernoff} say? What if you want a statement like

$$\Pr[X \geq \mu(1+\varepsilon)] \leq \delta,$$

for some very small parameter $\delta \to 0$ that should show up in your asymptotic bound on ε?

3. Same two questions but now for $p = 1/\sqrt{n}$.
4. What take-aways do you get about when it's a good idea to use Chebyshev vs each type of Chernoff?

4 Practice with Poissonization (if time)

Group Work

Suppose you are dropping n balls into n bins.

For each of the following two bounds, explain at a very high level (e.g., a non-quantitative outline) how you would use Poissonization (and de-Poissonization) to fill in the blank. (In particular, what would be your strategy for picking the Poisson parameter? Should it be a bit bigger or a bit smaller than n?)

After you've done that for both bounds, try to fill in the quantitative details if you have time.

1. As above, you are dropping n balls into n bins. Let X_i be the indicator random variable that is 1 if there are at most two balls in bin i, and zero otherwise. Thus, $X = \sum_i X_i$ is the number of bins with at most two balls in them.

   Use Poissonization to fill in the blank as best as possible, using big-Oh notation as $n \to \infty$.

   $$\Pr[X \geq 0.9999n] \leq \_\_\_\_\_\_$$

2. Now let $Y_i = 1 - X_i$, so Y_i is 1 iff there are *more* than 2 balls in bin i. Let $Y = \sum_i Y_i$ be the number of bins with more than 2 balls in them.

   Use Poissonization to fill in the blank as best as possible, using big-Oh notation as $n \to \infty$.

   $$\Pr[Y \geq 0.3 \cdot n] \leq \_\_\_\_\_\_$$