CS 537: *Randomness in Computing*
Boston University

Prof. Sofya Raskhodnikova
TF Gavin Brown

Discussion 4 – Wednesday, February 12, 2020

Recall Markov's and Chebyshev's inequalities:

$$\Pr[X \ge a] \le \frac{\mathbb{E}[X]}{a} \quad \text{and} \quad \Pr[|X - \mathbb{E}[X]| \ge a] \le \frac{\text{Var}[X]}{a^2}.$$

2d $\frac{1}{4}$ $a = k$

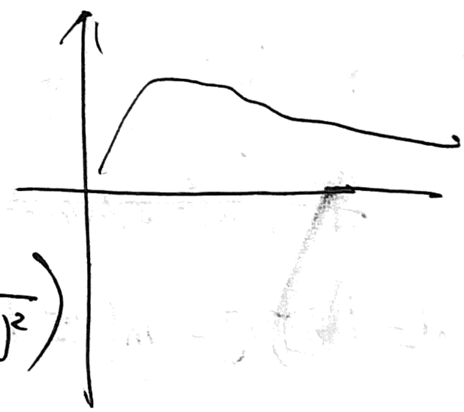
Markov's inequality only ~~holds for non~~negative random variables.

Problems

1. (**Coupon Collector**) Let random variable X be the number of draws it takes to collect all n coupons. Recall that we proved $\mathbb{E}[X] = nH_n = n \ln n + O(n)$. Upper bound $\Pr[X \ge a]$, trying different values of a to see what yields interesting and/or clean bounds. Use the following techniques:

   (a) Markov's inequality

$$Pr[X \ge a] \le \frac{n \ln n + O(n)}{a}$$



μ $\quad a = 2E[X]$

$$Pr[|X - nH_n| \ge nH_n] \le \frac{Var[X]}{n^2 (H_n)^2} \le \frac{n^2 \frac{\pi^2}{6}}{n^2 (H_n)^2} = O\left(\frac{1}{(\ln n)^2}\right)$$

   (b) Chebyshev's inequality

$$Pr[|X - E[X]| \ge \mu] \le \frac{Var[X]}{[illegible]}$$

$$Var[X] = \sum_{i=1}^{n} Var[X_i] \qquad Var[X^2] = (Var[X])^2 + \mu^2$$

$$\le \sum_{i=1}^{n} \left(\frac{n}{n-i+1}\right)^2 \qquad Var[X_i] = \frac{1-p}{p^2} \le \frac{1}{p^2}.$$

$$= n^2 \sum_{i=1}^{n} \frac{1}{(n-i+1)^2} = n^2 \sum_{j=1}^{n} \frac{1}{j^2} \le n^2 \sum_{j=1}^{\infty} \frac{1}{j^2} = n^2 \frac{\pi^2}{6}$$

   (c) The union bound

$$P[X > a] \le Pr\left[\bigcup_{i=1}^{n} (X_i > a)\right] \le \sum_{i=1}^{n} Pr[X_i > a] = \sum_{i=1}^{n} (1-p)^a$$

$Pr[X > a] = Pr[$after ath draw, still missing at least one coupon$]$

E_i ~~miss~~ missing i after a draws

$$= Pr\left[\bigcup_{i=1}^{n} E_i\right] \le \sum_{i=1}^{n} Pr[E_i] = \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^a = n\left(1 - \frac{1}{n}\right)^a \le n \cdot e^{-\frac{a}{n}}$$

try $a = 2n H_n$.

$\Rightarrow n \ln n + 2nc \quad n \exp(-2(n-2c)) \Rightarrow \frac{1}{n}$

2. (**Characterizing the Mean and Median**) Take a random variable X with finite expectation $\mathbb{E}[X]$ and finite median m. The expression $\mathbb{E}[|X-c|]$ is minimized by setting $c = m$ (proved in the book). Prove that $\mathbb{E}[(X-c)^2]$ is minimized by $c = \mathbb{E}[X]$.

$$f(c) = \mathbb{E}[(X-c)^2] = \mathbb{E}[X^2] - 2c\,\mathbb{E}[X] + c^2 .$$

$$\frac{\partial f'(c)}{\partial c} = 2c - 2\mathbb{E}[X] = 0$$

$$\Rightarrow c = \mathbb{E}[X].$$

3. (**Mean and Median are Close**) For a random variable X with finite mean μ, median m, and standard deviation σ, prove that $|\mu - m| \leq \sigma$.

$$\cancel{\mathbb{E}\left(\mu^2 - 2\mu m + m^2\right) \leq \mathbb{E}\left(\sigma^2\right)}.$$

$$|\mu - m| = |\mathbb{E}[X] - m| = |\mathbb{E}[X-m]| \leq \mathbb{E}[|X-m|] \quad \text{Jensen's inequality}$$

$$\leq \mathbb{E}[|X-\mu|].$$

$$= \mathbb{E}\left[\sqrt{(X-\mu)^2}\right]$$

$$\leq \sqrt{\mathbb{E}[(X-\mu)^2]} = \sigma \quad \boxtimes$$