CS 70	Discrete Mathematics and Probability Theory	
Spring 2015	Vazirani	Discussion 12W

1. Sanity Check!

a. Prove or give a counterexample: for any random variables X and Y, $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$.

Answer:
Let X be any random variable such that $\text{Var}[X] \neq 0$. Then, if we set $Y = X$, $\text{Var}[X+Y] = \text{Var}[2X] = \text{E}[(2X)^2] - (\text{E}[2X])^2 = 4\text{E}[X^2] - 4(\text{E}[X])^2 = 4\text{Var}[X] \neq \text{Var}[X] + \text{Var}[X]$.

b. Derive Chebyshev's inequality using Markov's inequality for some random variable X.

Answer: We're interested in the probability $\Pr(|X - \text{E}[X]| \geq k) = \Pr((X - \text{E}[X])^2 \geq k^2)$. We simply apply Markov's inequality and we find that $\Pr((X - \text{E}[X])^2 \geq k^2) \leq \text{Var}(X)/k^2$.

2. Balls and Bins Again

For this problem we toss m balls into n bins.

a. What is the expected number of collisions?

Answer: Let X be the expected number of collisions and X_{ij} the indicator for the event that balls i and j collide. Then $X = \sum_{1 \leq i < j \leq m} X_{ij}$. Since $\text{E}[X_{ij}] = 1/n$, we know that $\text{E}[X] = \binom{m}{2} \times \frac{1}{n}$.

b. Now, let's define X to be the number of collisions. At what threshold of collisions c can we ensure that the probability of having more than c collisions is less than $1/2n$?

Answer: By Markov's inequality, $\Pr[X \geq c] \leq \text{E}[X]/c$. Since we know $\text{E}[X]$ from the previous part we simply plug it in: $\Pr[X \geq c] \leq m(m-1)/2nc$, and so we set $m(m-1)/2nc = 1/2n$ and find that $c = m(m-1)$.

3. Coin flips

a. Suppose we flip a fair coin n times and we wish to understand the probability that we get at least $3n/4$ heads. Use Markov's inequality to come up with an upper bound for this probability.

Answer: Let X be a random variable for the number of heads. Let X_i be an indicator for the event that the i-th flip is heads. Since X_i is an indicator, $E[X_i] = \Pr(X_i = 1) = 1/2$. Since $X = \sum_{i=1}^{n} X_i$,

$$\text{E}[X] = \text{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \text{E}[X_i] = \frac{n}{2}$$

We want to bound the probability that $X \geq 3n/4$. Since X is nonnegative, we can use Markov's inequality to get

$$\Pr\left(X \geq \frac{3n}{4}\right) \leq \frac{\text{E}[X]}{\frac{3n}{4}} = \frac{\frac{n}{2}}{\frac{3n}{4}} = \frac{2}{3}$$

b. Use Markov's inequality to come up with a similar upper bound on the probability that the number of heads is at least n.

Answer: This time, we want to bound the probability that $X \geq n$. By Markov's inequality,

$$\Pr(X \geq n) \leq \frac{\mathrm{E}[X]}{n} = \frac{\frac{n}{2}}{n} = \frac{1}{2}$$

c. Find the true probability that there are at least n heads in a sequence of n fair coin flips. Is the bound you derived in the previous part tight?

Answer: Since X can't be greater than n,

$$\Pr(X \geq n) = \Pr(X = n) = \left(\frac{1}{2}\right)^n$$

So we can see that Markov's inequality gives a very loose bound; it bounds $\Pr(X \geq n)$ by a constant, whereas in reality this probability decreases exponentially as n increases.

4. More coin flips

a. Suppose we flip a biased coin 100 times and X is the number of heads we get. We know that $\mathrm{Var}[X] = 16$. What are the possible values for the expected vaue of X?

Answer: This is a binomial random variable, so we can use the properties we derived from last discussion. First of all, we know that the $\mathrm{Var}[X] = np(1-p)$. This means that $16 = 100p(1-p)$. Secondly, we know that $\mathrm{E}[X] = np$ which tells us that $p = \mathrm{E}[X]/100$. We plug it into our first equation and find that

$$16 = 100(\mathrm{E}[X]/100)(1 - \mathrm{E}[X]/100)$$

$$0 = E[X]^2 - 100E[X] + 1600$$

$$0 = (E[X] - 20)(E[X] - 80)$$

So we know that $E[X] = 20$ or $E[X] = 80$.

b. Now suppose $\mathrm{E}[X] = 20$. Use Chebyshev's inequality to derive an upper bound on $\Pr[X \geq 40]$.

Answer: According to Chebyshev's,

$$\Pr(|X - 20| \geq 20) \leq \mathrm{Var}(X)/20^2$$

We simply plug in the variance and we find that

$$\Pr[X \geq 40] \leq \Pr(|X - 20| \geq 20) \leq 16/400 = 1/25$$

.