CS5314
Randomized Algorithms

Lecture 8: Moments and Deviations
(Common Variance, Chebyshev Inequality)

Objectives

- Variances of Bin(n,p) and Geo(p)
- Chebyshev's Inequality

Variance of Binomial RV

Lemma: Let X be a binomial random variable with parameters n and p. Then,

$$Var[X] = np(1-p)$$

How do we get that?

Recall: $Var[X] = E[(X - E[X])^2]$

$$= E[X^2] - (E[X])^2$$

First Proof (computing $E[X^2]$)

$E[X^2] = \Sigma_{0 \le j \le n} j^2 \Pr(X=j)$

$= \Sigma_{0 \le j \le n} j^2 C_j^n p^j (1-p)^{n-j}$

$= \Sigma_{0 \le j \le n} (j(j-1)+j) C_j^n p^j (1-p)^{n-j}$

$= \Sigma_{2 \le j \le n} j(j-1) C_j^n p^j (1-p)^{n-j}$

$+ \Sigma_{1 \le j \le n} j C_j^n p^j (1-p)^{n-j}$

By expanding C_j^n term, we get:

First Proof (computing $E[X^2]$)

$E[X^2]$

$= n(n-1)p^2 \sum_{2 \le j \le n} C_{j-2}^{n-2} p^{j-2} (1-p)^{n-j}$

$+ np \sum_{1 \le j \le n} C_{j-1}^{n-1} p^{j-1} (1-p)^{n-j}$

$= n(n-1)p^2 \left(p + (1-p)\right)^{n-2} + np \left(p + (1-p)\right)^{n-1}$

$= n(n-1)p^2 + np$

First Proof (computing $E[X^2]$)

Since $Var[X] = E[X^2] - (E[X])^2$, we have:

$$
\begin{aligned}
Var[X] &= n(n-1)p^2 + np - (np)^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 \\
&= np - np^2 \\
&= np(1-p)
\end{aligned}
$$

Second Proof (using indicator)

Binomial r.v. $X = \mathrm{Bin}(n,p)$ can be written as the sum of n independent indicator, X_1, X_2, ..., X_n, each succeeds with probability p

That is, $X = X_1 + X_2 + \ldots + X_n$

So, $\mathrm{Var}[X] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \ldots + \mathrm{Var}[X_n]$

$= n\, \mathrm{Var}[X_1]$

Second Proof (using indicator)

$$\begin{aligned}
Var[X_1] &= E[(X_1 - E[X_1])^2] \\
&= (1-p)^2 \Pr(X_1=1) + (0-p)^2 \Pr(X_1=0) \\
&= (1-p)^2 p + p^2(1-p) \\
&= p(1-p)(1-p + p) = p(1-p)
\end{aligned}$$

Thus,

$$Var[X] = n\, Var[X_1] = np(1-p)$$

Variance of Geometric RV

Lemma: Let X be a geometric random variable with parameter p. Then,

$$Var[X] = (1-p)/p^2$$

How do we get that?

First Proof (computing $E[X^2]$)

$$E[X^2] = \Sigma_{j \geq 0}\, j^2 \Pr(X = j)$$
$$= \Sigma_{j \geq 0}\, j^2\, p\,(1-p)^{j-1}$$
$$= [\, p/(1-p)\,] \times \Sigma_{j \geq 0}\, j^2\,(1-p)^j$$

To get $E[X^2]$, it remains to compute the value of $\Sigma_{j \geq 1}\, j^2\,(1-p)^j$

Before that, let's look at some equalities

First Proof (computing $E[X^2]$)

For $|x| < 1$,

(a) $$1/(1-x) = \sum_{j \geq 0} x^j$$

(b) By differentiating (a), we get

$$1/(1-x)^2 = \sum_{j \geq 0} j\, x^{j-1}$$

(c) By differentiating (b), we get

$$2/(1-x)^3 = \sum_{j \geq 0} j(j-1)\, x^{j-2}$$

First Proof (computing $E[X^2]$)

Using the previous equalities,

$$\sum_{j \geq 0} j^2 x^j$$

$$= \sum_{j \geq 0} j(j-1) x^j + \sum_{j \geq 0} j x^j$$

$$= 2x^2/(1-x)^3 + x/(1-x)^2$$

$$= (2x^2 + x(1-x)) / (1-x)^3$$

$$= (x^2 + x) / (1-x)^3$$

First Proof (computing $E[X^2]$)

So,

$$E[X^2] = [\, p/(1-p)\,] \times \Sigma_{j \geq 0}\, j^2 (1-p)^j$$

$$= [\, p/(1-p)\,] \times ((1-p)^2+(1-p)) \,/\, (1-(1-p))^3$$

$$= [\, p/(1-p)\,] \times ((1-p)(2-p)) \,/\, p^3$$

$$= (2-p)/p^2$$

Then,

$$Var[X] = E[X^2] - (E[X])^2$$

$$= (2-p)/p^2 - (1/p)^2 = (1-p)/p^2$$

Second Proof (by memory-less property)

Let Y be a random variable such that

$Y=1$ if the first trial succeeds, and

$Y=0$ if the first trial fails

Then,

$$E[X^2] = \Pr(Y=1)\,E[X^2|Y=1] + \Pr(Y=0)\,E[X^2|Y=0]$$
$$= p\,E[X^2|Y=1] + (1-p)\,E[X^2|Y=0]$$
$$= p + (1-p)\,E[X^2|Y=0]$$

Second Proof (by memory-less property)

We want to get $E[X^2|Y=0]$ ...

Let Z = #remaining trials until first success

In this case ($Y=0$), we have $X = Z + 1$

So, $E[X^2|Y=0] = E[(Z+1)^2]$ ... [why?]

$= E[Z^2+2Z+1] = E[Z^2] + 2E[Z] + 1$

But from the memory-less property,

$E[Z^2] = E[X^2]$ and $E[Z] = E[X]$

Second Proof (by memory-less property)

So, $E[X^2] = p + (1-p)\, E[X^2 | Y=0]$

$= p + (1-p)\,(E[X^2] + 2E[X] + 1)$

$= p + (1-p)\,(E[X^2] + 2/p + 1)$

Rearranging terms,

$p\, E[X^2] = p + 2(1-p)/p + (1-p)$

$= 1 + (2-2p)/p \;\; = (2-p)/p$

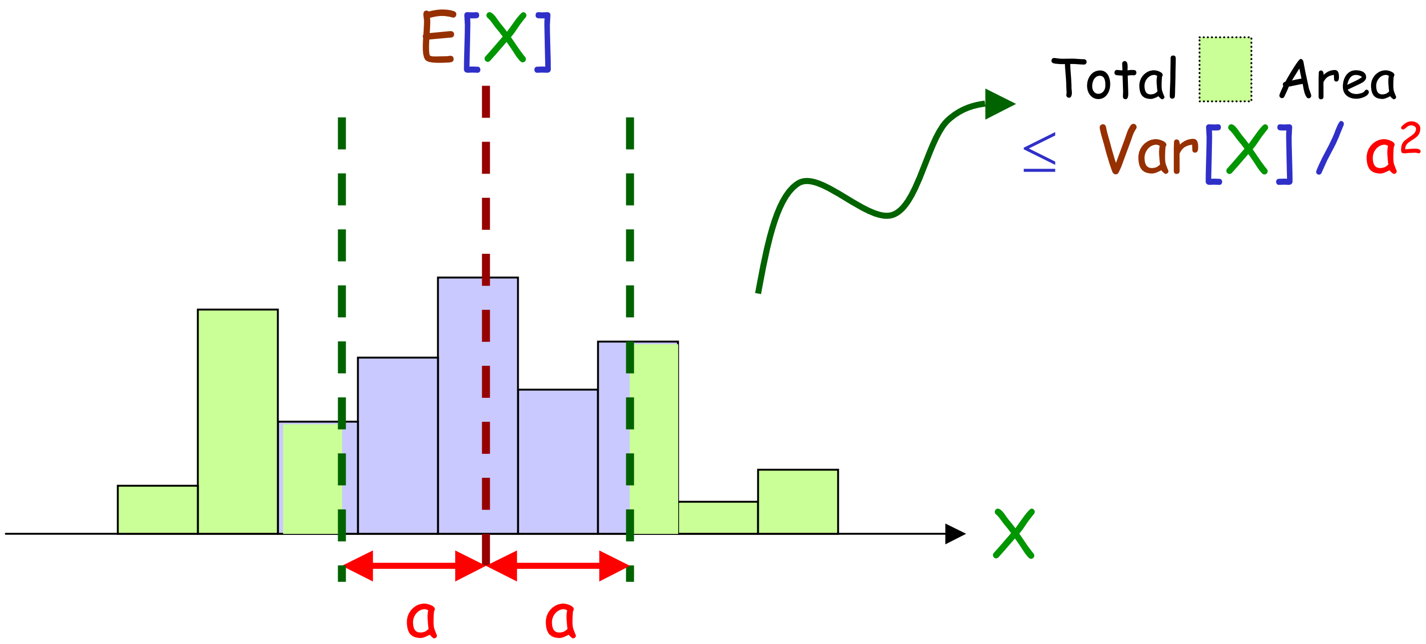
Again, $E[X^2] = (2-p)/p^2$

$\rightarrow$ $Var[X] = (1-p)/p^2$ as before

Chebyshev Inequality

Theorem: For any positive a,

$$\Pr(|X - E[X]| \geq a) \leq \mathrm{Var}[X]/a^2$$



Proof

By using Markov Inequality!!!

$$\Pr(\,|X - E[X]| \geq a\,)$$
$$= \Pr(\,(X - E[X])^2 \geq a^2\,)$$
$$\leq E[(X - E[X])^2]\,/a^2 \qquad \text{[by Markov inequality]}$$
$$= \mathrm{Var}[X]\,/a^2$$

Chebyshev Inequality
(other variations)

Corollary: For any positive r,

$$\Pr(|X - E[X]| \geq r\,\sigma[X]) \leq 1/r^2$$

Corollary: For any positive r,

$$\Pr(|X - E[X]| \geq r\,E[X]) \leq Var[X]/(r\,E[X])^2$$

Markov vs Chebyshev

- When applying Chebyshev :

1. X can take on negative values
2. Need Var[X] to get the bound
3. Often give better bounds than Markov (since it is based on more information)

Markov vs Chebyshev
(Example 1)

Suppose we flip a fair coin n times

Question: Can we bound the probability of more than $3n/4$ heads?

Let X = number of heads. So, $E[X] = n/2$

By Markov Inequality,

$$\Pr(X \geq 3n/4) \leq E[X] / (3n/4)$$
$$= (n/2) / (3n/4) = 2/3$$

Markov vs Chebyshev
(Example 1)

Let's use Chebyshev Inequality instead:

Again, X = number of heads

So, $E[X] = n/2$ and $Var[X] = n/4$ … [why?]

Then, we have

$$Pr(X \geq 3n/4)$$

$$\leq Pr(|X - E[X]| \geq n/4) \quad \text{…[why?]}$$

$$\leq Var[X] / (n/4)^2$$

$$= 4/n$$ … much better bound than 2/3!!!

Markov vs Chebyshev
(Example 2)

Let us revisit Coupon Collector's problem:

There are a total of n different cards. Each time, the card we buy is chosen independently and uniformly at random from the n cards.

Let X = number of cards we need to buy

Previously, we get $E[X] = nH(n)$

Markov vs Chebyshev
(Example 2)

Question: Can we bound the probability of buying more than $2nH(n)$ cards?

By Markov Inequality,

$$\Pr(X \geq 2nH(n))$$
$$\leq E[X] / (2nH(n))$$
$$= nH(n) / (2nH(n))$$
$$= 1/2$$

Markov vs Chebyshev
(Example 2)

Question:

How about using Chebyshev Inequality?

To apply the inequality, we need to get Var[X] … What is this value?

Markov vs Chebyshev
(Example 2)

Let X_i = #cards bought to get a new card after collecting exactly $i-1$ distinct cards

So, $X = X_1 + X_2 + \ldots + X_n$

Also, the variables X_i are all independent!

Thus,

$$Var[X] = Var[X_1] + Var[X_2] + \ldots + Var[X_n]$$

Markov vs Chebyshev
(Example 2)

What is $Var[X_k]$?

Recall: X_k is $Geo(p)$ with $p = (n-k+1)/n$

Thus,

$$\begin{aligned} Var[X_k] &= (1-p)/p^2 \\ &\leq 1/p^2 \\ &= n^2/(n-k+1)^2 \end{aligned}$$

Markov vs Chebyshev
(Example 2)

So, $Var[X] = Var[X_1] + Var[X_2] + \dots + Var[X_n]$

$$\leq n^2/(n)^2 + n^2/(n-1)^2 + \dots + n^2/(1)^2$$

$$\leq 2n^2$$

Now, by Chebyshev Inequality,

$$Pr(X \geq 2nH(n)) \leq Pr(|X - E[X]| \geq nH(n))$$

$$\leq Var[X] / (nH(n))^2$$

$$\leq 2n^2 / (nH(n))^2$$

$= O(1/\log^2 n)$ … much better than 1/2!!!