Probability in Computing

CS 237

LECTURE 22

Last time

- Markov inequality
- Chebyshev’s inequality

Today

- Applying Markov and Chebyshev’s inequalities

Reminder

- HW 10 is due Thursday

Markov and Chebyshev's Inequalities

Theorem (Markov Inequality)

Let X be a random variable taking only nonnegative values. Then, for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Andrei Markov
[1856-1922]

Theorem (Chebyshev's Inequality)

Let X be a random variable. For all $a > 0$,

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

Pafnuty Chebyshev
[1821-1894]

image source https://en.wikipedia.org/wiki/Andrey_Markov
image source https://www.britannica.com/biography/Pafnuty-Lvovich-Chebyshev

Top Hat question (Join Code: 413437)

BU Terriers win each game independently with probability 2/3.
Let X be the number of their losses in n games.
What bound on $\mathbf{Pr}\left(X \geq \frac{n}{2}\right)$ can you get using Markov inequality?

A. $\leq 1/3$

B. $\leq 1/2$

C. $\geq 1/2$

D. $\leq 2/3$

E. None of the above

Top Hat question (Join Code: 413437)

BU Terriers win each game independently with probability 2/3.
Let X be the number of their losses in n games.
What bound on $\mathbf{Pr}\left(X \geq \frac{n}{2}\right)$ can you get using Chebyshev's inequality?

A. $\leq 1/3$

B. $\leq const/n$

C. $\geq const/n^2$

D. $\leq const/n^3$

E. None of the above

For $n > 12$, Chebyshev is much better than Markov!

Estimation by Sampling: RVs

- For each $i \in \{1,2,\ldots,n\}$, define:

$$X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad\quad \text{otherwise} \end{cases}$$

- The pollster's estimate is:

$$P = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Our goal is to understand how ``close'' the estimate P is to the actual fraction of voters that support candidate A

CS 237 Sampling Estimate from Polling

- Fraction of the population supporting candidate A is $\boldsymbol{p}$
- The pollster's estimate is: $\boldsymbol{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

  where $X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$

  $X_i \sim Bernoulli(p) \qquad \mathbb{E}(X_i) = p \qquad \text{Var}(X_i) = p(1-p)$
- Expectation and variance of our estimate:

$$\mathbb{E}(P) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}\cdot np = \boldsymbol{p}$$

$$\text{Var}(P) = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}(X_i) = \frac{\boldsymbol{p(1-p)}}{\boldsymbol{n}}$$

Markov Inequality: Polling

Corollary (Markov Inequality)

Let X be a RV taking only nonnegative values. Then, for all $b > 1$,

$$\Pr(X \geq b \cdot \mathbb{E}(X)) \leq \frac{1}{b}.$$

Andrei Markov [1856-1922]

- By Markov inequality, applied to the polling estimate,

$$\Pr(P \geq 1.1p) \leq$$

$$\Pr(P \geq 2\,p) \leq$$

$$\Pr(P \geq b\,p) \leq$$

Chebyshev's Inequality: Polling

Theorem (Chebyshev's Inequality)

Let X be a random variable. For all $a > 0$,

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

Pafnuty Chebyshev

[1821-1894]

- By applying Chebyshev's inequality to the polling estimate,

$\Pr(|P - p| \geq a) \leq$

Chebyshev's Inequality: Polling

- The Chebyshev inequality bound allows us to determine how many people to poll
- Suppose we want the estimate of p to be within 0.04 with probability at least 0.95

Want:

$$\Pr(|P-p| \le 0.04) \ge 0.95$$
$$\Leftrightarrow \Pr(|P-p| > 0.04) \le 0.05$$

Chebyshev:

$$\Pr(|P-p| > 0.04) \le \frac{1}{4 \cdot (0.04)^2 \cdot n}$$

Thus it suffices to poll $n = 3125$ voters

Estimation by Sampling

- The approach we studied in the context of polling can be used in a wide range of settings
- A common scenario in CS and beyond is that we want to estimate the expectation of a distribution using sampling
- Our earlier analysis works equally well for this setting: we take several samples from the distribution and we use their average as our estimate for the expectation
- Chebyshev's inequality then tells us how close our estimate is to the actual expectation, and it also tells us how many samples we need

Random Hats: Using Markov

Number of fixed points of a permutation. Let X be the number of students that get their hats back when n students randomly switch hats, so that every permutation of hats is equally likely.

Find a bound on $\Pr(X \geq x)$ using Markov inequality.

Solution: X_i= the indicator R.V. for person i getting their hat back.

$$X = X_1 + \cdots + X_n$$

Then

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(X_1 + \cdots + X_n) \\ &= \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) \\ &= n \cdot \mathbb{E}(X_1) \\ &= n \cdot \Pr(X_1 = 1) \\ &= n \cdot \frac{1}{n} = 1 \end{aligned}$$

by linearity of expectation

by symmetry

since X_1 is an indicator

By Markov inequality, $\Pr(X \geq x) \leq$

Random Hats: Variance

Find a bound on $\Pr(X \geq x)$ using Chebyshev's inequality.

Solution: We know: $X = X_1 + \cdots + X_n$ and $\mathbb{E}(X) = 1$

To compute $\mathrm{Var}(X)$:

$$\mathbb{E}(X^2) = \mathbb{E}[(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)]$$

$$= \mathbb{E}(X_1^2) + \cdots \mathbb{E}(X_n^2) + \mathbb{E}(X_1 \cdot X_2) + \cdots + \mathbb{E}(X_{n-1} \cdot X_n)$$

by linearity of expectation

$$= n \cdot \mathbb{E}(X_1^2) + n(n-1) \cdot \mathbb{E}(X_1 \cdot X_2)$$

by symmetry

$$= n \cdot \Pr(X_1 = 1) + n(n-1) \cdot \Pr(X_1 \cdot X_2 = 1)$$

since X_1 and $X_1 \cdot X_2$ are indicators

$$= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2$$

Random Hats: Using Chebyshev

Find a bound on $\Pr(X \geq x)$ using Chebyshev's inequality.

Solution: We know: $\mathbb{E}(X) = 1$ and $\mathbb{E}\left(X^2\right) = 2$

$$\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = 2 - 1 = 1$$

$$\begin{aligned}\Pr(X \geq x) &= \Pr(X - 1 \geq x - 1) \\ &= \Pr(X - \mathbb{E}(X) \geq x - 1) \\ &\leq \Pr(|X - \mathbb{E}(X)| \geq x - 1)\end{aligned}$$