CSCE 689: Special Topics in Modern Algorithms for Data Science

Lecture 4

Samson Zhou

Today

- Sign up for LaTeX scribe note slots

- Receive and consider list of potential projects/groups

Future

- Wednesday: Discuss potential project groups
- Friday: Email me the members/group name

- Future: Set up meetings to discuss proposed projects

Last Time: Expected Value

- The expected value of a random variable X over Ω is:

$$\mathrm{E}[X] = \sum_{x \in \Omega} \Pr[X = x] \cdot x$$

- The “average value of the random variable"

- Linearity of expectation: $\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$

Last Time: Markov's Inequality

- Let $X \geq 0$ be a non-negative random variable. Then for any $t > 0$:

$$\Pr[X \geq t \cdot \mathrm{E}[X]] \leq \frac{1}{t}$$

- Can rewrite as $\Pr[X \geq t] \leq \frac{\mathrm{E}[X]}{t}$

- "Bounding the deviation of a random variable in terms of its average"

Limitations of Markov's Inequality

- Let X be the outcome of a roll of a die. Then $\mathrm{E}[X] = 3.5 = \frac{7}{2}$

$$\Pr[X \geq 6] = \Pr\left[X \geq \frac{12}{7} \cdot \frac{7}{2}\right] \leq \frac{7}{12} \approx 0.5833$$

- We know $\Pr[X \geq 6] = \frac{1}{6} \approx 0.167$

Moments

- For $p > 0$, the p-th moment of a random variable X over Ω is:

$$\mathrm{E}[X^p] = \sum_{x \in \Omega} \Pr[X = x] \cdot x^p$$

Variance

- The variance of a random variable X over Ω is:

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- Can rewrite $\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$ since $\mathrm{E}\big[\mathrm{E}[X]\big] = \mathrm{E}[X]$

- "How far numbers are from the average"

Variance

- The variance of a random variable X over Ω is:

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- Linearity of variance for *independent* random variables: $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$

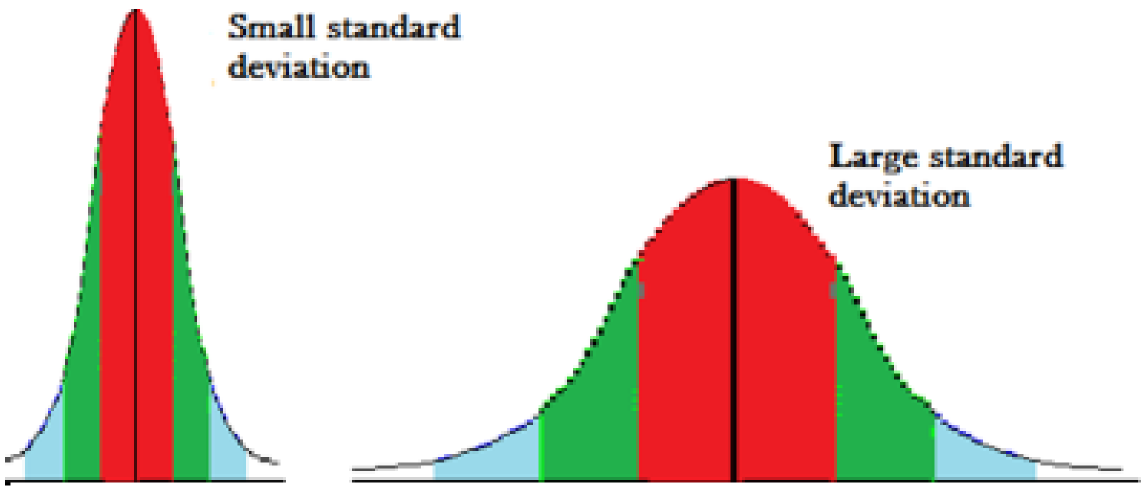
Variance and Standard Deviation

- The variance of a random variable X over Ω is:

$$\sigma^2 = \mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- The standard deviation of a random variable X is σ, and measures how far apart the outcomes are

- Standard deviation is in the same unit as the data set



Variance

- Suppose X takes the value 1 with probability $\frac{1}{2}$ and takes the value -1 with probability $\frac{1}{2}$
- What is $\mathrm{E}[X]$?
- What is $\mathrm{Var}[X]$? What is $\mathrm{std}(X)$?

Variance

- Suppose Y takes the value 100 with probability $\frac{1}{2}$ and takes the value -100 with probability $\frac{1}{2}$

- What is $\mathrm{E}[Y]$?

- What is $\mathrm{Var}[Y]$? What is $\mathrm{std}(Y)$?

Markov's Inequality

- Let $X \geq 0$ be a non-negative random variable. Then for any $t > 0$:

$$\Pr[X \geq t \cdot \mathrm{E}[X]] \leq \frac{1}{t}$$

- Can rewrite as $\Pr[X \geq t] \leq \frac{\mathrm{E}[X]}{t}$

Markov's Inequality

- Let $X \geq 0$ be a non-negative random variable. Then for any $t > 0$:

$$\Pr[X \geq t \cdot \mathrm{E}[X]] \leq \frac{1}{t}$$

- Can rewrite as $\Pr[X \geq t] \leq \frac{\mathrm{E}[X]}{t}$

- We have $\Pr[|X| \geq t] = \Pr[X^2 \geq t^2]$

Using Markov's Inequality

- We have $\Pr[|X| \geq t] = \Pr[X^2 \geq t^2]$

$$\Pr[|X| \geq t] = \Pr[X^2 \geq t^2] \leq \frac{E[X^2]}{t^2}$$

- Plug in $X - \mathrm{E}[X]$ for X

$$\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{E}[(X - \mathrm{E}[X])^2]}{t^2}$$

Toward Chebyshev's Inequality

$$\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{E}[(X - \mathrm{E}[X])^2]}{t^2}$$

Chebyshev's Inequality

$$\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{E}[(X - \mathrm{E}[X])^2]}{t^2}$$

- Recall that $\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = \mathrm{E}[(X - \mathrm{E}[X])^2]$
- $\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$

Chebyshev's Inequality

- Let X be a random variable with expected value $\mu := \mathrm{E}[X]$ and variance $\sigma^2 := \mathrm{Var}[X]$

- $\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$ becomes $\Pr[|X - \mathrm{E}[X]| \geq t] \leq \frac{\sigma^2}{t^2}$

$$\Pr[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

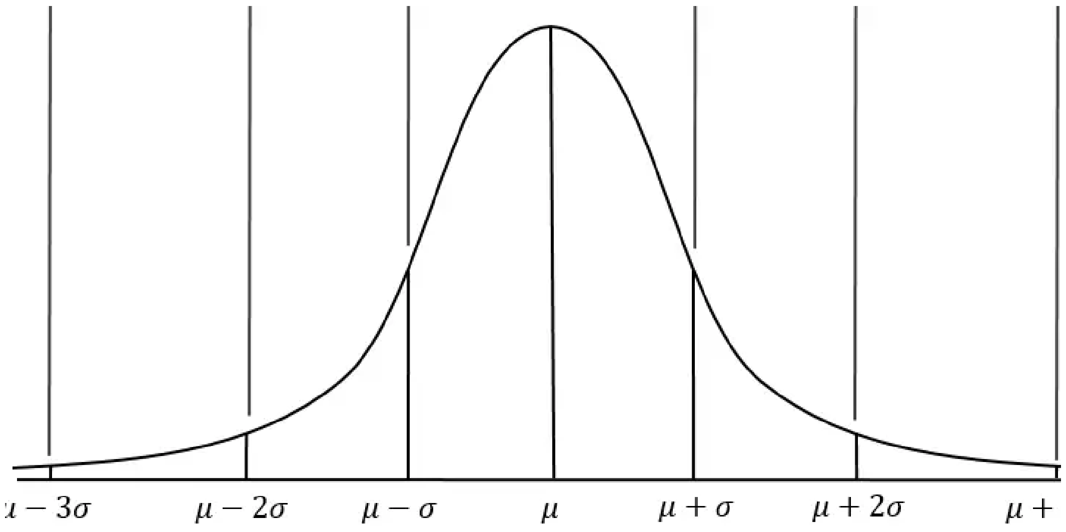
- "Bounding the deviation of a random variable in terms of its variance"

Chebyshev's Inequality

- Let X be a random variable with expected value $\mu := \mathrm{E}[X]$ and variance $\sigma^2 := \mathrm{Var}[X]$

$$\Pr[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

- Do not require assumptions about X



Chebyshev's Inequality

- Let X be the outcome of a roll of a die. Then $\mathrm{E}[X] = 3.5 = \frac{7}{2}$ and $\mathrm{Var}[X] = \frac{35}{12} \approx 2.92$ so $\mathrm{std}(X) \approx 1.71$

$$\begin{aligned} \Pr[X \geq 6] &= \Pr[X - 3.5 \geq 2.5] \\ &= \Pr[X - 3.5 \geq 1.41 \cdot 1.71] \\ &\leq \frac{1}{1.41^2} \approx 0.4667 \end{aligned}$$

- Recall that Markov's inequality bounded this 0.5833

Law of Large Numbers

- Let $X_1, \ldots, X_n$ be random variables that are independent identically distributed (i.i.d.) with mean μ and variance σ^2
- Consider the sample average $X = \frac{1}{n}\sum_i X_i$. How does it compare to μ?
- $\mathrm{Var}[X] = \frac{1}{n^2}\sum_i \mathrm{Var}[X_i] = \frac{\sigma^2}{n}$
- By Chebyshev's inequality, $\Pr[|S - \mu| \geq t] \leq \frac{\sigma^2}{nt}$

Law of Large Numbers

- By Chebyshev's inequality, $\Pr[|S - \mu| \geq t] \leq \frac{\sigma^2}{nt}$

- Law of Large Numbers: The sample average will always concentrate to the mean, given enough samples