

① Markov's inequality

② Chebyshev's inequality

Reminder:

- HW sessions 3 times a week (@ 39)
- If the response rate for course evals reaches 80% there will be an additional HW drop.
(Currently at 50%) Due by 8/13

Markov's Inequality Intro

Simple bound on the tail of a random variable that only uses the expected value (first moment) and the fact that the random variable is nonnegative.

X is nonnegative r.v.

$\mathbb{E}[X]$ is known

$f_x(x)$



$P(X \geq a) \leq ?$

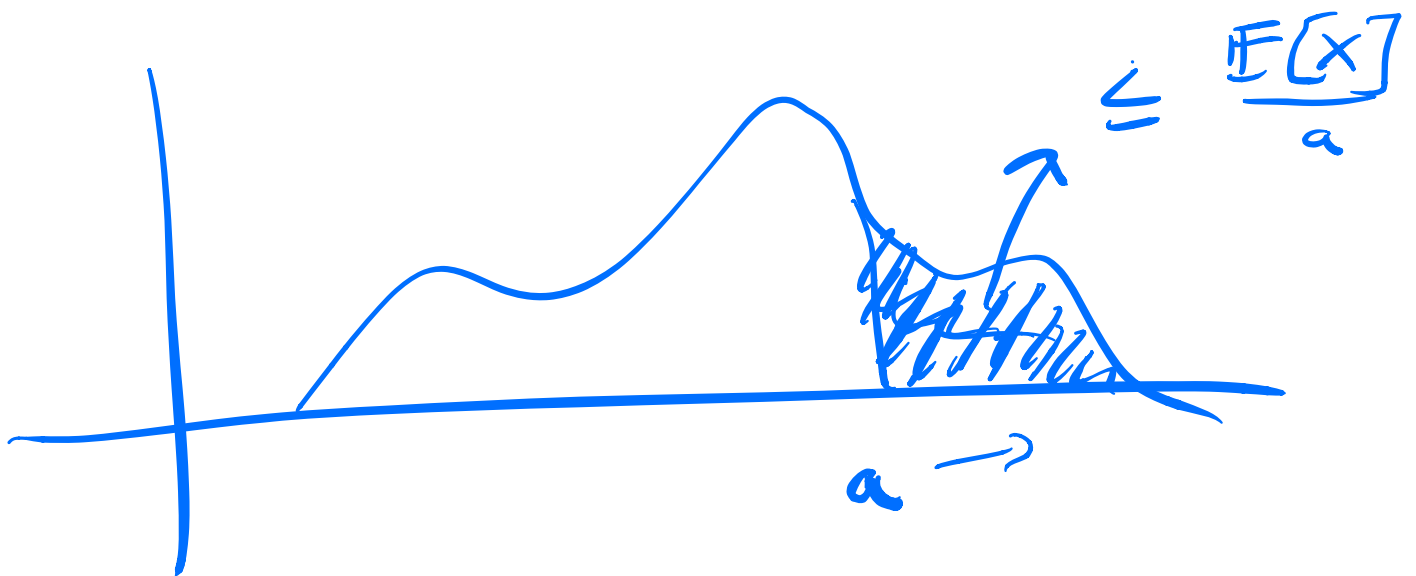
k^{th} moment of r.v. $\mathbb{E}[X^k]$



Markov's Inequality Definition

If X is a nonnegative r.v. with finite mean and $a > 0$, then the probability that X is at least a is at most the expectation of X divided by a

$$P(X \geq a) \leq \frac{E[X]}{a}$$



Markov's Inequality: Proof 1 Without loss of generality

WLOG, let X be a nonnegative continuous R.V.

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^{\infty} x \cdot f(x) dx$$

$$= \underbrace{\int_0^a x \cdot f(x) dx} + \int_a^{\infty} x \cdot f(x) dx$$

≥ 0 , since X is nonnegative.

$$\geq \int_a^{\infty} x \cdot f(x) dx$$

$$\geq \int_a^{\infty} a \cdot f(x) dx$$

$$= a \underbrace{\int_a^{\infty} f(x) dx}$$

$$= a \cdot \mathbb{P}(X \geq a)$$

$$\Rightarrow \mathbb{E}[X] \geq a \cdot \mathbb{P}(X \geq a)$$

$$\Rightarrow \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

x is at least as big as a

"If $\mathbb{E}[X]$ is small, then the probability that X is large is small"

* Assuming X is nonnegative *

Markov's Inequality: Proof II

Let I be the indicator r.v. defined as:

$$I = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{o.w.} \end{cases}$$

Only holds if X is nonnegative.

Then $X \geq a \cdot I$

$$\mathbb{E}[X] \geq \mathbb{E}[aI]$$

$$\mathbb{E}[X] \geq a \mathbb{E}[I]$$

$$\mathbb{E}[X] \geq a \cdot P(X \geq a)$$

$$\Rightarrow P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Case 1: $X \geq a$

$$X \geq a \cdot 1 \rightarrow X \geq a \checkmark$$

Case 2: $X < a$

$$X \geq a \cdot 0 \rightarrow X \geq 0$$

X is nonnegative (given) \checkmark



Markov's Inequality: Proof III

$$\mathbb{E}[X] = \underbrace{\mathbb{E}[X|X < a] \cdot P(X < a) + \mathbb{E}[X|X \geq a] \cdot P(X \geq a)}_{\text{Law of Total Expectation}}$$

nonnegative.
since X is nonnegative.

$$\geq \mathbb{E}[X|X \geq a] \cdot P(X \geq a)$$

$$\geq a \cdot P(X \geq a)$$

$$\Rightarrow P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Example: Markov & Coin Flips

Let $X \sim \text{Geom}(\frac{1}{2})$. Use Markov's inequality to bound $P(X > 10)$.

$$P(X > 10) = P(X \geq 11) \leq \frac{E[X]}{11} = \frac{2}{11}$$

Note: What is $P(X \geq 11)$ exactly?

$$P(X \geq 11) = \left(\frac{1}{2}\right)^{10} = \frac{1}{2^{10}}$$

$$\frac{1}{2^{10}}$$

↑
actual

$$< \frac{2}{11}$$

↑
bound

Markov's inequality gives a loose bound in this case.

$$P(X \geq 11) = P(X=11) + P(X=12) + P(X=13) + \dots$$



=>

Same answer.

Generalized Markov's Inequality

If X is any random variable with finite mean and $a > 0$, then for any $r > 0$:

$$\underbrace{P(|X| \geq a)} \leq \frac{E[|X|^r]}{a^r}$$

Proof left as an exercise.

(Actually the proof is in Note 18)

Chebyshev's Inequality Intro

Often times we can do better than Markov's Inequality if we use more information about the random variable.

For Chebyshev's Inequality, we use the first two moments $E[X]$ and $E[X^2]$.

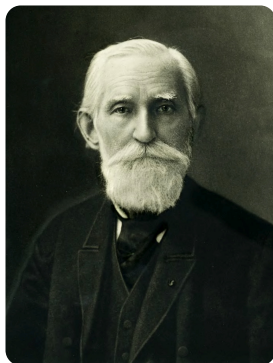
Note: The variance of a random variable

"captures" the first two moments

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

k^{th} moment $\rightarrow E[X^k]$

Markov's
Advisor.



Chebyshev's Inequality: Definition.

If X is a random variable with finite mean μ and finite variance, and $c > 0$, then the probability that X is at least c away from μ is at most $\frac{\text{Var}(X)}{c^2}$.

→ doesn't need to be nonnegative.

$$c \in \mathbb{R}$$

$$c > 0$$

$$P(|X - \mu| \geq c) \leq \frac{\text{Var}[X]}{c^2}$$

Note:

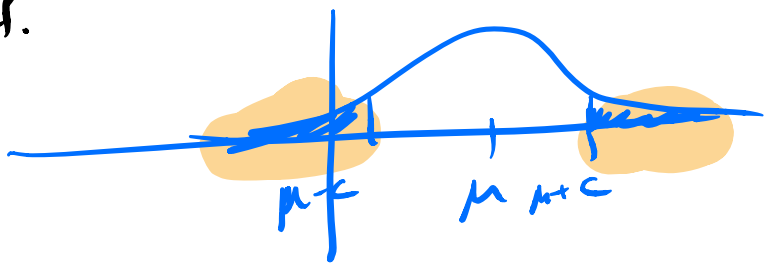
X does not need to be nonnegative.

c is positive constant.

Chebyshev's Inequality: Proof.

X is a r.v.

$$\mathbb{E}[X] = \mu$$



Let $Y = (X - \mu)^2$, then Y is a nonnegative r.v.

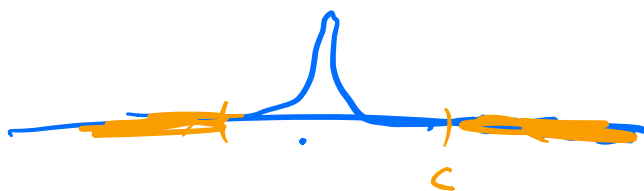
$$\mathbb{E}[Y] = \mathbb{E}[(X - \mu)^2] = \text{Var}[X]$$

Note that $\mathbb{P}(|X - \mu| \geq c) = \mathbb{P}(Y \geq c^2)$

$$\begin{aligned} \text{So, } \mathbb{P}(|X - \mu| \geq c) &= \mathbb{P}(Y \geq c^2) \\ &\leq \frac{\mathbb{E}[Y]}{c^2} \quad \left. \begin{array}{l} \text{Markov's} \\ \text{on } Y \end{array} \right\} \\ &= \frac{\text{Var}[X]}{c^2} \end{aligned}$$

$$\rightarrow \mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}[X]}{c^2}$$

"If the variance of X is small, then the probability that X is far from its mean is small"



Example: Chebyshev & Coin Flips

Let $X \sim \text{Geom}(\frac{1}{2})$. Use Chebyshev's Inequality to upper bound $\mathbb{P}(X > 10)$.

$$\mathbb{E}[X] = 2 = \mu$$

$$\text{Var}(X) = 2 = \frac{1-p}{p^2} = \frac{\frac{1}{2}}{\frac{1}{4}}$$

$$\mathbb{P}(X > 10) = \mathbb{P}(X \geq 11)$$

$$= \mathbb{P}(X \geq \mu + 9)$$

$$= \mathbb{P}(X - \mu \geq 9)$$

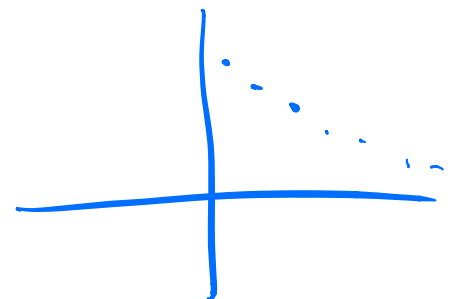
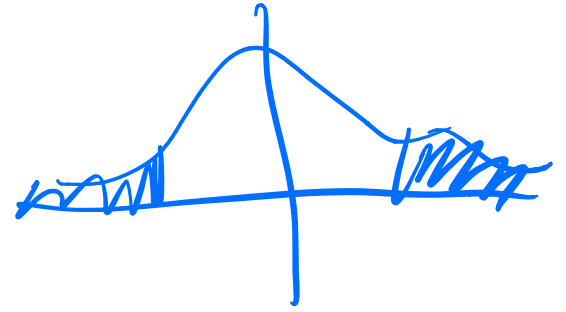
$$\leq \mathbb{P}(|X - \mu| \geq 9)$$

Chebyshev's
Inequality

$$\leq \frac{\text{Var}(X)}{9^2} = \frac{2}{81}$$

This is tighter than Markov's $(\frac{2}{11})$, but

is still far off from $\frac{1}{2^{10}}$



Chebyshev Corollary

For any random variable X with finite expectation $E[X] = \mu$ and finite standard deviation $\sigma = \sqrt{\text{Var}(X)}$:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Proof:

Plug $c = k \cdot \sigma$ into Chebyshev's

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2}$$

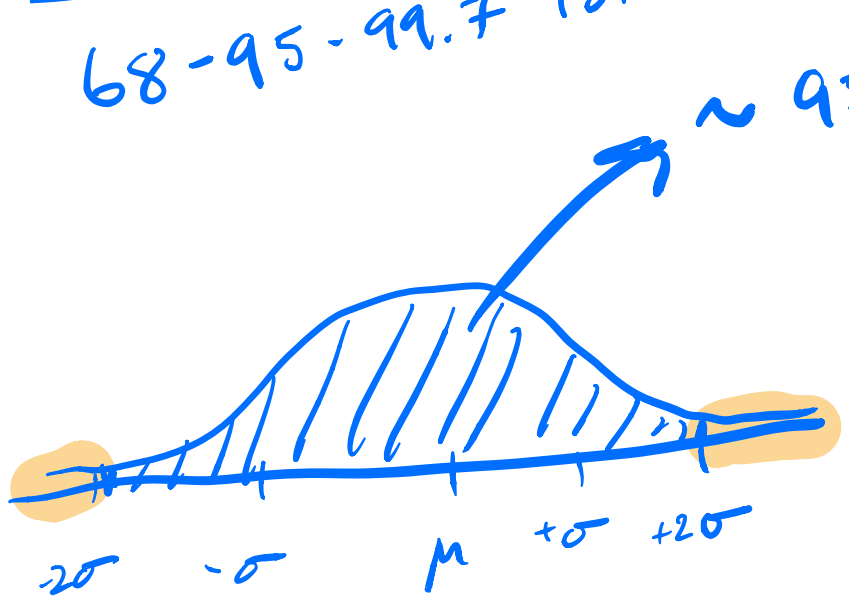
Example: Chebyshev Corollary

Let $X \sim N(\mu, \sigma^2)$. Find a bound on the probability that X is 2σ or more away from its mean μ .

$$P(|X - \mu| \geq \underset{\substack{\uparrow \\ k}}{2\sigma}) \leq \frac{1}{2^2} = \frac{1}{4}$$

Note:

68-95-99.7 "rule"



~ 95% → 5% outside.

$$\frac{1}{20} \ll \frac{1}{4}$$

↑
"rule"

Law of Large Numbers: Preview

If we observe a random variable many times and average our observations, the average will converge to the average of the random variable.

We will use Chebyshev's Inequality to prove this.

Note:

Observe random variables \rightarrow data