Concentration

CS 70, Summer 2019

Lecture 23, 8/1/19

Plan:

- Variance measures **deviation from mean**
- How do we **quantify** this? Can we get good **bounds** even if we don't get an exact answer?
- Tools that use **mean** and **variance**:
  - Markov's Inequality
  - Chebyshev's Inequality
  - Law of Large Numbers

Trees. (The Non-Graph Kind.)

I plant 10 trees in my yard.
After a few months, I measure their growth. I tell you that the average height of my trees is 3 ft.

Is it possible that I have 4 trees of 10 ft or taller?

Trees. (The Non-Graph Kind.)

I sample any one of my 10 trees.
Let H be the height of my tree.

What can I say about $\mathbb{P}[H \geq 10]$?

Markov Inequality

X is a RV that only takes **non-negative** values.

$$\mathbb{P}[X \geq c] \leq$$

Proof:

Example: Coin Tosses I

I flip 200 fair coins.
What is an **upper bound** on the probability that we get more than 150 heads?

Example: Generalized Markov

Let Y be a **arbitrary** RV. Let $c > 0$, $r > 0$.
Want to show:

$$\mathbb{P}[|Y| \geq c] \leq \frac{\mathbb{E}[|Y|^r]}{c^r}$$

What should we apply Markov to?

Markov: A Tight Example?

Is it possible that $\mathbb{P}[X \geq c] = \frac{\mathbb{E}[X]}{c}$?

Strategy: Go through proof of Markov.
For every **inequality** (i.e. $\geq$), determine what would give **equality.**

Markov: A Tight Example?

(Continued...)

Chebyshev Inequality

Let X be an **arbitrary** RV.
We have the following bound for two-sided tails:

$$\mathbb{P}[|X - \mu| \geq c] \leq$$

Example: Coin Tosses II

I flip 200 fair coins.
What is an **upper bound** on the probability that we get more than 150 heads?

Example: Lower Bound on Variance

Let X be a RV such that $\mathbb{E}[X] = 1$, and

$$\mathbb{P}[-2 < X < 3] = \frac{1}{2}$$

Can I get a lower bound on $\mathrm{Var}(X)$?

Example: Lower Bound on Variance

(Continued...)

Break

Whats the best Wi-Fi name youve ever seen?

Example: Bias of a Coin

I have a coin with **unknown** head probability p.
I want to **estimate** p within some **error tolerance** ϵ, and I want to be **confident** in my estimate with some probability $1 - \delta$.

To do this, I flip my coin n times, measure the number of heads, and divide by n.

Let $\hat{p}$ be my estimate.

Example: Bias of a Coin

What is $\mathbb{E}[\hat{p}]$? What is $\mathrm{Var}[\hat{p}]$?

Example: Bias of a Coin

I want $\hat{p}$ to be within ϵ (**error**) of p with probability (**confidence**) $1 - \delta$.

Chebyshev: A Tight Example?

Is it possible that $\mathbb{P}[|X - \mathbb{E}[X]| \geq c] = \frac{\mathrm{Var}(X)}{c^2}$?

Strategy: Go through proof of Chebyshev.
Where do we use inequality (i.e. $\leq$ or $\geq$) instead of equality?

Chebyshev: A Tight Example?

(Continued...)

The Law of Large Numbers

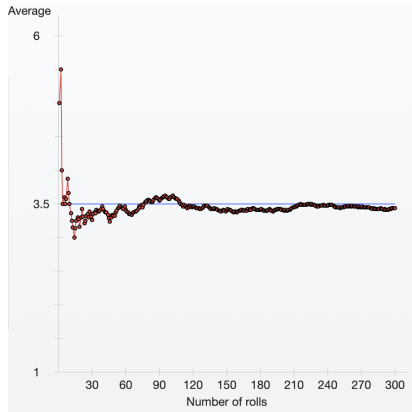
Intuition: If we observe a RV X **many times**, and **average** the observations, the average **converges** to $\mathbb{E}[X]$.

Formally: Let $X_1, X_2, \ldots$ be a sequence of i.i.d. RVs with expectation μ (where μ is finite).

Let $S_n = X_1 + X_2 + \ldots + X_n$. Then:

$$\mathbb{P}[|\frac{1}{n} \cdot S_n - \mu| < \epsilon] \to 1 \text{ as } n \to \infty$$

Visualization: Dice Rolls



Example: Coin Game

A fair coin is tossed.
1) You win if there are more than 60% heads. Which is better, 10 or 100 tosses?

2) You win if there are more than 40% heads. Which is better, 10 or 100 tosses?

Example: Coin Game

A fair coin is tossed.
1) You win if there are between than 40% and 60% heads. Which is better, 10 or 100 tosses?

2) You win if there are **exactly** 50% heads. Which is better, 10 or 100 tosses?

Exactly 50% Heads

Compare:

$$\mathbb{P}[n \text{ heads out of } 2n \text{ tosses}]$$

and

$$\mathbb{P}[(n+1) \text{ heads out of } (2n+2) \text{ tosses}]$$

Exactly 50% Heads

(Continued...)

Summary

- Markov helps find **one-sided** tail probability for **non-negative** RVs, given the **mean**.
- Chebyshev helps find **two-sided** tails for **any** RV, given the mean **and variance**.
- LLN tells us that if we observe a RV many times, the **probability that we are "close" to the mean** nears 1.