February 17

Monday, February 17, 2020 3:47 PM

Comp Sci 212

1. Markov's Inequality

What can we say about a random variable, based on just its expected value?

$\Pr[\underset{\text{random variable}}{R} \geq \underset{\text{number}}{x}] \leq ?$

- If R represents the running time of a random algorithm, want $\Pr[R \geq x]$ to be small

Def. Expected Value Sample space S, Prob. dist $\Pr$

$E[R] = \sum_{w \in S} R(s) \cdot \Pr[w]$ $R: S \to \mathbb{R}$

Ex. In a room with 100 people, average age is 40. How many people can be at least 100 years old?

S = set of people

$\Pr[w] = \frac{1}{100}$ for all $w \in S$, uniform

$R: S \to \mathbb{R}$ $R(w)$ = age of person w

$E[R] = \sum_{w \in S} \frac{R(w)}{100} = 40$

What is the largest $\Pr[R \geq 100]$ can be?

- Not everyone can be at least 100 $\to \Pr[R \geq 100] < 1$
  - average age would be at least 100
- Not even 50 people can be over 100 $\to \Pr[R \geq 100] \leq \frac{1}{2}$
  - if 50 people are 100, 50 are 0, average is 50
- 25 people is possible 25 people 100, 75 people 20 years > avg. is 40
  - over 100 years
- largest 40 people

  40 people 100 years old, 60 people 0 years old > avg is 40

  $\Pr[R \geq 100] \leq \frac{40}{100} = \frac{E[R]}{100}$

Thm. Markov's Inequality

If R is a non-negative random variable ($R(w) \geq 0$ for all $w \in S$), $x > 0$

$\Pr[R \geq x] \leq \frac{E[R]}{x}$

Proof. Enough to prove $x \cdot \Pr[R \geq x] \leq E[R]$

$\sum_{w \in S \text{ s.t. } R(w) \geq x} \Pr[w] x$ $\sum_{w \in S} R(w) \cdot \Pr[w]$

$\sum_{w \in S} R(w)\Pr[w] = \sum_{w \in S \text{ s.t. } R(w) \geq x} R(w)\Pr[w] + \underbrace{\sum_{w \in S \text{ s.t. } R(w) < x} R(w)\Pr[w]}_{\text{at least } 0}$

$\geq \sum_{w \in S \text{ s.t. } R(w) \geq x} R(w)\Pr[w]$

$\geq \sum_{w \in S \text{ s.t. } R(w) \geq x} x \cdot \Pr[w] = x \cdot \Pr[R \geq x]$

$\square$

- Assume that everyone is at least 20 years old in our room.
  - Now 40 people over 100 is impossible
    - 40 people 100 years
    - 60 people 20 years - avg. 52
  - largest # is 25 people $\to \Pr[R \geq 40] \leq \frac{1}{4}$

    25 people 100 years old, 75 people 20 years old avg. 40

Birthday Paradox - D = set of possible days

n people in a room, $S = \underbrace{D \times D \times \cdots \times D}_{n \text{ times}}$, uniform dist.

$R: S \to \mathbb{R}$, R = # of pairs of people with the same birthday

$E_{i,j}$ = event that person i & j have the same birthday, $\Pr[E_{i,j}] = \frac{1}{|D|}$

$R = \sum_{i,j \text{ s.t. } i<j} \mathbb{1}_{E_{i,j}}$, $E[R] = \sum_{i,j \text{ s.t. } i<j} \Pr[E_{i,j}] = \sum_{i,j \text{ s.t. } i<j} \frac{1}{|D|} = \frac{n \cdot (n-1)}{2|D|} \approx 1$ if $n \geq 2\sqrt{|D|}$

If $n \geq 2\sqrt{|D|} \to \Pr[R \geq 1] \geq 1 - \frac{1}{e}$, but what about $\Pr[R \geq 10]$?

$\hookrightarrow \leq \frac{1}{10}$ by Markov's inequality

Corollary - If $R(w) \geq b$ for all w,

$\Pr[R \geq x] \leq \frac{E[R] - b}{x - b}$

Proof. Let $Y(w) = R(w) - b$

Y is always non-negative

$E[Y] = E[R - b] = E[R] - b$

$\Pr[R \geq x] = \Pr[Y \geq x - b] \leq \frac{E[Y]}{x-b} = \frac{E[R]-b}{x-b}$

$\square$

Room examples show that Markov's inequality and Corollary are tight.

$S' = \{H, T\}$, $S = \underbrace{S' \times \cdots \times S'}_{100 \text{ times}}$ uniform; sticky - all heads, all tails

$R: S \to \mathbb{R}$ $R(s)$ = # of heads. $E[R] = 50$ both cases

By Markov's inequality $\Pr[R \geq 75] \leq \frac{50}{75} \leq \frac{2}{3}$

Right answer - sticky $\Pr[R \geq 75] = \frac{1}{2}$

uniform $\Pr[R \geq 75]$ = really small