S. KIMMEL

What is average number of heads in 4 coin flips?

GENERAL STRATEGY

read off problem statement

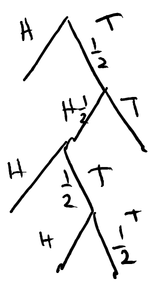
$\Downarrow$

1. See average. Figure out sample space, random variable of interest

$\{H,T\}^4$ $\qquad$ $X(i) = \#$ heads

2a. Decide if can calculate $\mathbb{E}[X]$ directly:

$$\mathbb{E}[X] = \sum_{i \in S} \Pr(i) X(i)$$



$\Pr(THTT) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16}$

will be same for all elements of sample space

$$\mathbb{E}[X] = \frac{1}{16} \sum_{i \in S} X(i)$$

$\vdots$

If getting complicated, or if I tell you to, use indicator random variables

S.KIMMEL

2. Write X as a weighted sum of indicator random variables (can do in several steps)

a) Think about what events causes X to increase

Event: 1^{st} flip is H $\rightarrow$ increases X by 1

Event: 2^{nd} flip is H $\rightarrow$ "

" 3^{rd} "

" 4^{th}

b) Create indicator random variables for each event & write X as sum, based on how much each increases

$$X = 1 \cdot \underbrace{X_1}_{\text{If flip 1 is H, add 1}} + 1 \cdot \underbrace{X_2}_{\text{If flip 2 is H, add 1}} + 1 \cdot X_3 + 1 \cdot X_4 = \sum_{k=1}^{4} X_k$$

3. Use linearity of expectation:

$$\mathbb{E}[X] = \sum_{k=1}^{4} \mathbb{E}[X_k]$$

S.KIMMEL

5. Use $\mathbb{E}[X_E] = \Pr(E)$ for indicator random variables.

$$\mathbb{E}[X] = \sum_{k=1}^{4} \Pr(k^{th} \text{ flip is heads}) = \sum_{k=1}^{4} \frac{1}{2} = 2$$

Super powerful!

ex: Time complexity of random binary search is $O(\log n)$

S. KIMMEL

Q: Consider strings of $\{1,2,4\}^n$, where

$Pr(k^{th}\text{ digit is }1) = 1/k$

$Pr(k^{th}\text{ digit is }2) = (1-1/k)/2$

$Pr(k^{th}\text{ digit is }4) = (1-1/k)/2$

What is average sum of digits?

1. See <u>average</u>. Need sample space, random variable

$\{1,2,4\}^n$ $\qquad$ $X(i)$ = sum of digits in i

2. Write X as a weighted sum of indicator random variables (can do in several steps)

$$X = 1\cdot \underbrace{X_1}_{\#\,1\text{'s}} + 2\cdot \underbrace{X_2}_{\#\,2\text{'s}} + 4\cdot \underbrace{X_4}_{\#\,4\text{'s}}$$

$$X_1 = \sum_{k=1}^{n} X_{1,k} \qquad X_{1,k} = \begin{cases} 1 & \text{if } k^{th} \text{ position is } 1 \\ 0 & \text{otherwise} \end{cases}$$

$$X_2 = \sum_{k=1}^{n} X_{2,k} \qquad X_{2,k} = \begin{cases} 1 & \text{if } k^{th} \text{ position is } 2 \\ 0 & \text{otherwise} \end{cases}$$

$$X_4 = \sum_{k=1}^{n} X_{4,k} \qquad X_{4,k} = \begin{cases} 1 & \text{if } k^{th} \text{ position is } 4 \\ 0 & \text{otherwise} \end{cases}$$

$$X = \sum_{k=1}^{n} X_{1,k} + 2\sum_{k=1}^{n} X_{2,k} + 4\sum_{k=1}^{n} X_{4,k}$$

3. Use linearity of expectation:

$$\mathbb{E}[X] = \sum_{k=1}^{n} \mathbb{E}[X_{1,k}] + 2\sum_{k=1}^{n} \mathbb{E}[X_{2,k}] + 4\sum_{k=1}^{n} \mathbb{E}[X_{4,k}]$$

S.KIMMEL

5. Use $\mathbb{E}[X_E] = \Pr(E)$ for indicator random variables.

$$\mathbb{E}[X] = \sum_{k=1}^{n} \underbrace{\mathbb{E}[X_{1,k}]}_{\frac{1}{k}} + \sum_{k=1}^{n} 2\underbrace{\mathbb{E}[X_{2,k}]}_{(1-1/k)/2} + \sum_{k=1}^{n} \underbrace{\mathbb{E}[X_{4,k}]}_{(1-1/k)/2} 4$$

$$= \sum_{k=1}^{n} \frac{1}{k} + 1 - \frac{1}{k} + 2 - 2/k$$

$$= \sum_{k=1}^{n} 3 - \frac{2}{k} = 3n - O(\log n)$$

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln(n)$$