

What is average number of heads in 4 coin flips?

# GENERAL STRATEGY

1. See average. Figure out sample space,

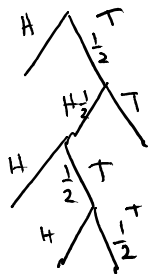
$$\{H, T\}^4$$

read off problem statement  
⇓  
random variable of interest

$$X(i) = \underline{\# \text{ heads}}$$

2a. Decide if can calculate  $E[X]$  directly:

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



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

will be same for all elements of sample space

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

⋮

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

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<sup>st</sup> flip is H  $\rightarrow$  increases  $X$  by 1

Event: 2<sup>nd</sup> flip is H  $\rightarrow$  "

" 3<sup>rd</sup> " "

" 4<sup>th</sup>

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

$$X = 1 \cdot X_1 + 1 \cdot X_2 + 1 \cdot X_3 + 1 \cdot X_4 = \sum_{k=1}^4 X_k$$

3. Use linearity of expectation:

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

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

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

---

Super powerful!

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

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

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

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

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

What is average sum of digits?

1. See average. Need sample space, random variable

$$\{1, 2, 4\}^n$$

$$X(i) = \text{sum of digits in } i$$

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

$$X = 1 \cdot X_1 + 2 \cdot X_2 + 4 \cdot X_4$$

$\uparrow$                        $\uparrow$                        $\uparrow$   
 # 1's                      # 2's                      # 4's

$$X_1 = \sum_{k=1}^n X_{1,k}$$

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

$$X_2 = \sum_{k=1}^n X_{2,k}$$

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

$$X_4 = \sum_{k=1}^n X_{4,k}$$

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

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

3. Use linearity of expectation:

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

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

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

$$\downarrow \qquad \qquad \qquad \downarrow \qquad \qquad \qquad \downarrow$$

$$\frac{1}{k} \qquad \qquad \qquad (1 - 1/k)/2 \qquad \qquad \qquad (1 - 1/k)/2$$

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

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

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