CSCB63 TUT4 WEEK10

andrewyk.leung@mail.utoronto.ca
https://snoopysnipe.github.io/ta/b63w21/

Recall: X is a discrete random variable that takes on values $x_1, \dots, x_k$

- $E[X] = \sum_{i=k}^{k} x_i P(X = x_i)$
- If X is indicator/Bernoulli random variable
  - ↳ $X = 1$ with probability $p$
  - $X = 0$ with probability $1-p$
  - ↳ $E[X] = 1 \cdot p + 0(1-p) = p = Pr(X=1)$
- If X and Y are random variables that are not necessarily independent
  - ↳ $E[X+Y] = E[X] + E[Y]$
  - ↳ $E[aX] = aE[X]$ for $a \in \mathbb{R}$

1. What is the expected # of heads after n coin tosses?
   Probability of heads is p.

Let $X$ = # of heads after $n$ coin tosses

Want to find $E[X]$

Let $X_i$ = # of heads on i-th toss (indicator)

$\Rightarrow X = X_1 + \dots + X_n$

$$\Rightarrow E[X] = E[X_1 + \dots + X_n]$$
$$= \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

2. n customers leave their hats at coat check. Customers later receive hats back but randomly uniformly permuted. What is the expected number of customers that are lucky enough to get their own hat back?

receive hats back but randomly uniformly permuted. What is the expected number of customers that are lucky enough to get their own hat back?

Let $X_i$ be the indicator random variable for "1 if the i-th customer gets their own hat, 0 otherwise"

$E[X_i] = Pr(X_i = 1)$ [same as i-th ball problem]

$= 1 \cdot (n-1)! / n!$

$= \frac{1}{n}$

$$E[X_1 + \ldots + X_n] = \sum_{i=1}^{n} E[X_i] = n \cdot \frac{1}{n} = 1$$

3.
```
x := 0
do n times:
    if random() ≤ 3/4
        then x:=x+5
    else
        x:=x-1
```

What is the expected final value of x after the program runs?

Note 3/4 probability to run $x := x + 5$

1/4 probability to run $x := x - 1$

Let $X_i$ be the change to x on iteration i

$E[X_i] = 5 \cdot 3/4 + (-1) \cdot 1/4 = 3.5$

$$E[X_1 + \ldots + X_n] = \sum_{i=1}^{n} E[X_i] = 3.5n$$

$\therefore$ 0 + 3.5n

initial value of x

expected change to x after n iterations

4. Toss a ball into one of n bins randomly and repeat. What is the expected # of times until all bins are non-empty?

↳ Decompose the random variable "# of times until every bin is non-empty" into a sum of the easier random variables:

"# of times to go from $i-1$ non-empty bins to $i$ non-empty bins" $= X_i$

$\Rightarrow$ $X_i$ is a geometric random variable, why?

We are counting "# of trials it takes until a success"

eg: $Pr(X_3 = 5) = Pr(\text{fail})^4 \; Pr(\text{Success})^1$

Let $p_i = Pr(\text{Success})$ for $X_i$

If $X_i \sim \text{Geometric}(p_i)$, $E[X_i] = \frac{1}{p_i}$

What is $p_i$?

↳ If toss into nonempty bin, fail

↳ If toss into empty bin, success

$$p_i = \frac{\#\text{ of empty bins}}{\#\text{ of total bins}} = \frac{n - (i-1)}{n}$$

$$= \frac{n-i+1}{n}$$

$$\Rightarrow E[X_i] = \frac{n}{n-i+1}$$

$$\Rightarrow E[X] = E[X_1 + \dots + X_n]$$

$$= \sum_{i=1}^{n} E[X_i]$$

$$= \sum_{i=1} E[X_i]$$

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

$$= n \underbrace{\left( \frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{1} \right)}_{n^{th} \text{ partial sum of harmonic series}}$$

$$= n \sum_{x=1}^{n} \frac{1}{x}$$

$$= n H_n$$

$$= n \left( 1 + \sum_{x=2}^{n} \frac{1}{x} \right)$$

$$< n \left( 1 + \int_1^n \frac{1}{x} dx \right)$$

$$= n (1 + \ln n)$$

$$> n \int_1^{n+1} \frac{1}{x} dx$$

$$= n \ln(n+1)$$

$$\therefore n \ln(n+1) < n H_n < n (1 + \ln n)$$

$$\Rightarrow n H_n \in \Theta(n \ln n)$$