2-8 Probabilistic Analysis

"No Expectation, No Disappointment."

Hengfeng Wei

hfwei@nju.edu.cn

April 21, 2020

```
RANDOMIZE-IN-PLACE(A)
n = A.length
for i = 1 to n
    swap A[i] with A[RANDOM(i, n)]
```

Fisher–Yates shuffle & Knuth shuffle (TAOCP Vol. 2)

Sampling without Replacement

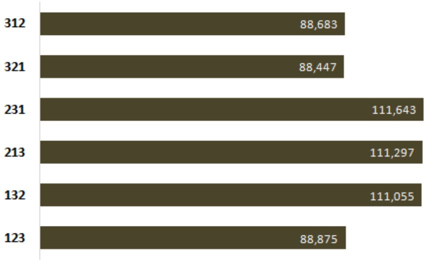
Permute-With-All (TC 5.3-3)

Permute-With-All(A)

$n = A.length$
for $i = 1$ **to** n
    swap $A[i]$ with $A[\text{Random}(1, n)]$

$\boxed{N^N \text{ vs. } N!}$

$3^3 = 27 \text{ vs. } 3! = 6$



$\# = 600,000$

The Danger of Naïveté @ Coding Horror

Definition (Expectation)

$$\mathbb{E}[X] = \sum_{x} x \Pr(X = x)$$

Theorem (A Second Way of Computing Expectation)

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

Theorem (Linearity of Expectation)

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathbb{E}[cX] = c\,\mathbb{E}[X]$$

$$\mathbb{E}[X+Y] = \sum_{s \in S} \Big(X(s) + Y(s) \Big) \Pr(s) = \mathbb{E}[X] + \mathbb{E}[Y]$$

Definition (Indicator Random Variable)

$$I_E = \begin{cases} 1, & \text{if } E \text{ occurs} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbb{E}[I_E] = \Pr(E)$$

Hat-check Problem (TC 5.2-4)

X : # of customers who get back their own hat $\qquad \mathbb{E}[X]$

$$I_i = \begin{cases} 1 & \text{customer } c_i \text{ gets back his/her hat} \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{n} I_i \qquad \mathbb{E}[I_i] = \Pr\Big(c_i \text{ gets back his/her hat}\Big) = \frac{1}{n}$$

$$\mathbb{E}[X] = 1$$

Inversions (TC 5.2-5)

$$A[1 \cdots n] \text{ of } n \text{ distinct numbers}$$

$$(i, j) \text{ is an inversion of } A : i < j \land A[i] > A[j]$$

$$X : \# \text{ of inversions in } A$$

$$\mathbb{E}[X] \quad (A \text{ is randomly ordered})$$

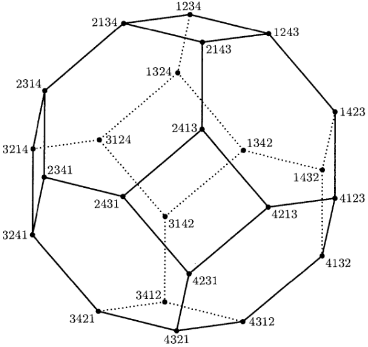
$$I_{ij} = \begin{cases} 1 & (i, j) \text{ is an inversion} \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{n-1} \sum_{j>i}^{n} I_{ij} \qquad \mathbb{E}[I_{ij}] = \Pr\Big((i, j) \text{ is an inversion}\Big) = \frac{1}{2}$$

$$\mathbb{E}[X] = \frac{n(n-1)}{4}$$

$$\mathbb{E}[X] = \sum_{s \in S} X(s) \Pr(s) = \frac{1}{n!} \sum_{s \in S} X(s)$$

$$= \frac{1}{n!} \left(\frac{n!}{2} \cdot \frac{n(n-1)}{2} \right) = \frac{n(n-1)}{4}$$



\# of inversions in $\langle 3214 \rangle$ + # of inversions in $\langle 4123 \rangle$

Searching an Unsorted Array (TC Problem 5-2 (f))

```
procedure DETERMINISTIC-SEARCH(A[1 ··· n], x)
    i ← 1

    while i ≤ n do
        if [A[i] = x] then
            return true
        i ← i + 1

    return false
```

$$\exists!_k \; i : A[i] = x$$

Y : # of comparisons

$$\exists!_k \; i : A[i] = x$$

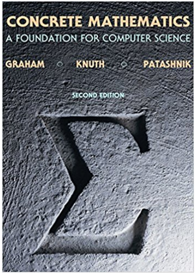
Y : # of comparisons

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{i=1}^{n-k+1} i \Pr\{Y = i\} \\
&= \sum_{i=1}^{n-k+1} i \Pr\{i \text{ is the first index among } k \text{ indices } s.t.\ A[i] = x\} \\
&= \sum_{i=1}^{n-k+1} i \frac{\binom{n-i}{k-1}}{\binom{n}{k}} = \frac{1}{\binom{n}{k}} \sum_{i=1}^{n-k+1} i \binom{n-i}{k-1} = \cdots \\
&= \frac{1}{\binom{n}{k}} \binom{n+1}{k+1} = \frac{n+1}{k+1}
\end{aligned}$$

$$k = 1 \implies \mathbb{E}[Y] = \frac{n+1}{2}, \qquad k = n \implies \mathbb{E}[Y] = 1$$

How Did I Evaluate this Summation:

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



$$r\binom{r-1}{k-1} = k\binom{r}{k}$$

$$\sum_{0 \le k \le n} \binom{k}{m} = \binom{n+1}{m+1}$$

```
Sum[i Binomial[n - i, k - 1], i, 1, n - k + 1]
```

$$
\begin{aligned}
\sum_{i=1}^{n-k+1} i\binom{n-i}{k-1} &= \sum_{i=0}^{n-k}(i+1)\binom{n-i-1}{k-1} \\
&= \sum_{i=0}^{n-k}\left((n+1)-(n-i)\right)\binom{n-i-1}{k-1} \\
&= \sum_{i=0}^{n-k}(n+1)\binom{n-i-1}{k-1} - \sum_{i=0}^{n-k}(n-i)\binom{n-i-1}{k-1} \\
&= (n+1)\sum_{i=0}^{n-k}\binom{n-i-1}{k-1} - k\sum_{i=0}^{n-k}\binom{n-i}{k} \\
&= (n+1)\sum_{m=k-1}^{n-1}\binom{m}{k-1} - k\sum_{m=k}^{n}\binom{m}{k} \\
&= (n+1)\binom{n}{k} - k\binom{n+1}{k+1} = \binom{n+1}{k+1}
\end{aligned}
$$

Theorem (A Third Way of Computing Expectation)

Let X be a discrete random variable that takes on only nonnegative integer values $\mathbb{N}$.

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} i \Pr(X = i) = \sum_{i=1}^{\infty} \Pr(X \geq i)$$

$$\begin{array}{lll} \Pr(X = 1) & & \\ \Pr(X = 2) & \Pr(X = 2) & \\ \Pr(X = 3) & \Pr(X = 3) & \Pr(X = 3) \\ \dots & & \end{array}$$

$$\sum_{j=1}^{\infty} \sum_{i=1}^{j} \Pr(X = j) = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \Pr(X = j)$$

$$
\begin{aligned}
\mathbb{E}[Y] &= \sum_{i=1}^{n-k+1} \Pr\{Y \geq i\} \\
&= \sum_{i=1}^{n-k+1} \frac{\binom{n-i+1}{k}}{\binom{n}{k}} \\
&= \frac{1}{\binom{n}{k}} \sum_{i=1}^{n-k+1} \binom{n-i+1}{k} \\
&= \frac{1}{\binom{n}{k}} \sum_{r=k}^{n} \binom{r}{k} \\
&= \frac{1}{\binom{n}{k}} \binom{n+1}{k+1} = \frac{n+1}{k+1}
\end{aligned}
$$

Definition (Conditional Expectation on an Event)

$$\mathbb{E}[X \mid E] = \sum_x x \Pr(X = x \mid E)$$

$$\Big(\mathbb{E}[X] = \sum_x x \Pr(X = x) \Big)$$

Theorem (A Fourth Way of Computing Expectation (CS 5.6-8))

Let X be a random variable defined on a sample space Ω.

Let $E_1, E_2, \cdots, E_n$ be a partition of Ω.

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid E_i] \Pr(E_i)$$

"The Law of Total Expectation"

$$\begin{aligned}
\mathbb{E}[X] &= \sum_x x \Pr(X = x) = \sum_x x \sum_{i=1}^{n} \Pr(X = x \mid E_i) \Pr(E_i) \\
&= \sum_x \sum_{i=1}^{n} x \Pr(X = x \mid E_i) \Pr(E_i) = \sum_{i=1}^{n} \sum_x x \Pr(X = x \mid E_i) \Pr(E_i) \\
&= \sum_{i=1}^{n} \mathbb{E}[X \mid E_i] \Pr(E_i)
\end{aligned}$$

(∄) Rational Person Playing a Card Game (CS 5.6 – 4)

A : \$1.00; Repeat

J : \$2.00; End

K : \$3.00; End

Q : \$4.00; End

Conditioning on the first draw c

$$\mathbb{E}[X] = \frac{1}{4}\Big(\mathbb{E}[X \mid c = A] + \mathbb{E}[X \mid c = J] + \mathbb{E}[X \mid c = K] + \mathbb{E}[X \mid c = Q]\Big)$$

$$\mathbb{E}[X \mid c = A] = \mathbb{E}[X] + 1$$

$$\mathbb{E}[X] = \frac{1}{4}\Big(\mathbb{E}[X] + 1 + 2 + 3 + 4\Big) = \frac{10}{3}$$

In-class Exercise: Coin Pattern (Provided by Yifan Pei)

X : # of tosses to get 3 consecutive heads (HHH)

$$\mathbb{E}[X]$$

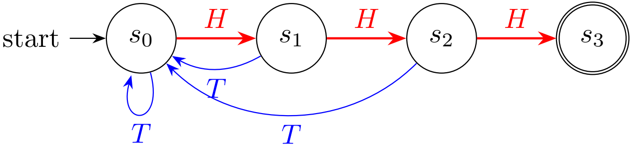
Conditioning on the first 3 tosses

$$T, \quad HT, \quad HHT, \quad HHH$$

$$\mathbb{E}[X] = \frac{1}{2}(\mathbb{E}[X] + 1) + \frac{1}{4}(\mathbb{E}[X] + 2) + \frac{1}{8}(\mathbb{E}[X] + 3) + \frac{1}{8} \times 3 = 14$$

$\boxed{X : \#\text{ of tosses to get } HHH}$

$$T, \quad HT, \quad HHT, \quad HHH$$

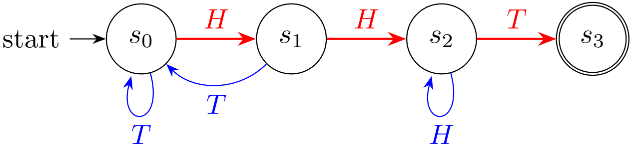


$$\mathbb{E}[X] = \frac{1}{2}(\mathbb{E}[X]+1) + \frac{1}{4}(\mathbb{E}[X]+2) + \frac{1}{8}(\mathbb{E}[X]+3) + \frac{1}{8} \times 3$$

$$\mathbb{E}[X_{H^n}] = \cdots = 2(2^n - 1)$$

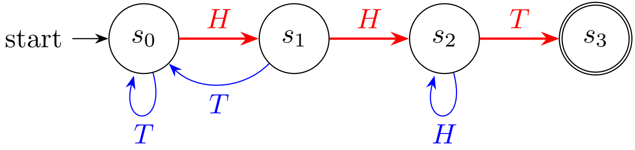
$$\boxed{X : \# \text{ of tosses to get } HHT}$$

$$\mathbb{E}[X_{HHH}] \ vs. \ \mathbb{E}[X_{HHT}]$$



$$T, \quad HT, \quad HHH, \quad HHT$$

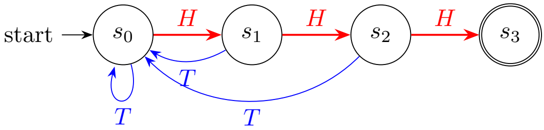
$$\mathbb{E}[X] = \frac{1}{2}(\mathbb{E}[X] + 1) + \frac{1}{4}(\mathbb{E}[X] + 2) + \frac{1}{8}(\mathbb{E}[X] + 3) + \frac{1}{8} \times 3$$

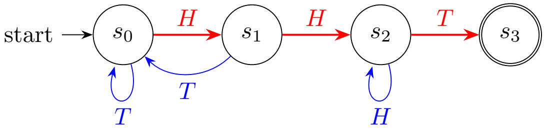


$\boxed{S_i: \text{ Expected number of tosses from state } s_i \text{ to reach state } s_n}$

$$
\begin{aligned}
S_0 &= \frac{1}{2}(1 + S_0 + 1 + S_1) \\
S_1 &= \frac{1}{2}(1 + S_0 + 1 + S_2) \\
S_2 &= \frac{1}{2}(1 + S_2 + 1 + S_3) \\
S_3 &= 0
\end{aligned}
$$

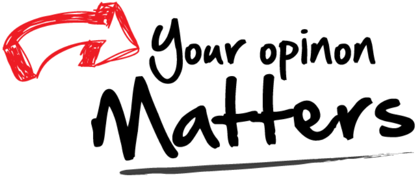
$$S_0 = 8$$





$$\mathbb{E}[X_{HHH}] = 14 > \mathbb{E}[X_{HHT}] = 8$$


Thank
You!



Office 302

Mailbox: H016

hfwei@nju.edu.cn