# CS7800: Advanced Algorithms

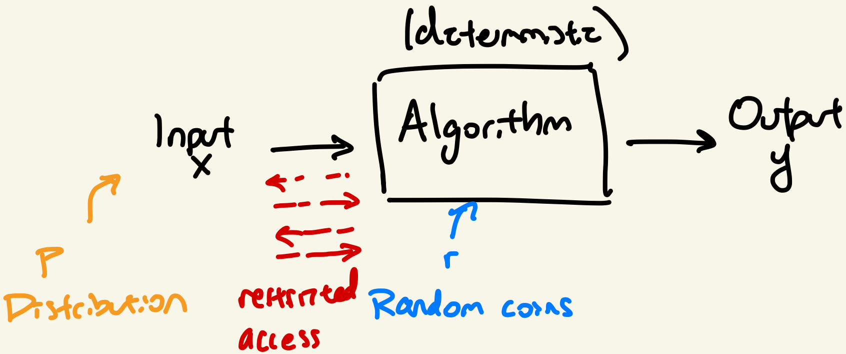
## Lecture 16: Randomized Algorithms I

- Overview
- Probability Workout

Jonathan Ullman

11-04-2022

# Randomness in Algorithms



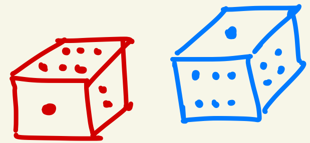

**Correctness**: for **every** input $x$, $y = A(x)$ is correct

**Running time**: for **every** $x$, $A(x)$ runs in time $T(|x|)$

How does randomness change this picture

1. "Average-case running time"
2. "Randomized algorithms"
   for every $x$, $\Pr(A(x,r) \text{ is correct}) \approx 1$
   for every $x$, $\mathbb{E}(\text{time on input } x) \leq T(|x|)$
3. Restricted access to input
4. Secrecy and security

# (Discrete) Probability Toolkit

- Outcomes $\omega \in \Omega$ 

  e.g. $\Omega = \{1,2,3,4,5,6\}^2$

  $\omega = (6, 1)$

- Probability $\mathbb{P} : \Omega \to \mathbb{R}_{\geq 0}$

  $\left( \sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1 \right)$

  e.g. $\mathbb{P}(\omega) = \frac{1}{36}$ for $\omega \in \Omega$

- Events $E \subseteq \Omega$

  e.g. $E = \{\omega : \omega_1 + \omega_2 = 7\}$

- Probability of an event is

  $$\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega)$$

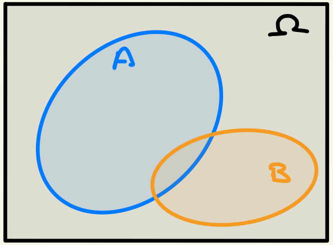
  e.g. $\mathbb{P}(E) = 6 \times \frac{1}{36} = \frac{1}{6}$

- Can take negations ("not"), unions ("or"), and intersections ("and")

# Combining Events

- Conditional Probability

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \wedge B)}{\mathbb{P}(B)}$$




- Independence

A and B are <u>independent</u>

$$\Leftrightarrow \quad \mathbb{P}(A \wedge B) = \mathbb{P}(A)\mathbb{P}(B)$$

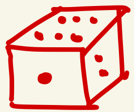
$$\Leftrightarrow \quad \mathbb{P}(A|B) = \mathbb{P}(A) \quad \text{(and vice versa)}$$


- Independence is an assumption (e.g. "Roll two independent dice")

# Random Variables (neither random nor variable)

- A random variable maps an <u>outcome</u> to a <u>value</u>

outcome $\omega \in \Omega = \{1, 2, \dots, 6\}$



$\omega = 6$

$$\xrightarrow{X(\omega) = \omega^3 + 1} 217$$

$$\xrightarrow{X(\omega) = \text{King Henry the } \omega}$$

- We treat integer-valued r.v.s as variables taking unknown values

$$\mathbb{P}(X = x) = \mathbb{P}(\{\omega : X(\omega) = x\})$$

$$\mathbb{P}(X \leq x) = \mathbb{P}(\{\omega : X(\omega) \leq x\})$$

$$\mathbb{P}(X = Y) = \mathbb{P}(\{\omega : X(\omega) = Y(\omega)\})$$

# Expected Value

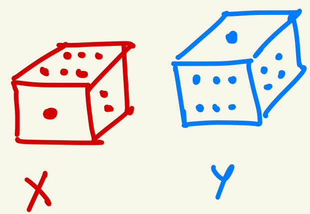
- The expected value of an integer-valued r.v. $X$ is

$$\mathbb{E}(X) = \sum_{x=-\infty}^{\infty} x \cdot \mathbb{P}(X=x)$$

- Expectation is linear $\mathbb{E}(aX+bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

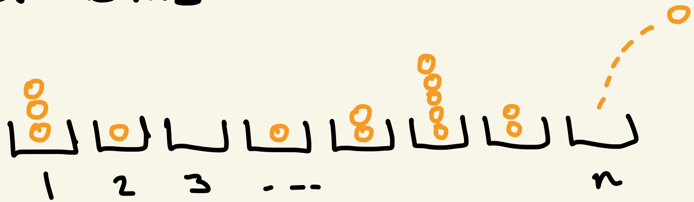
- R.v.s $X$ and $Y$ are independent if

$$\mathbb{P}(X=x \wedge Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y)$$




- If $X$ and $Y$ are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$

# Balls and Bins




Throw $m$ balls into $n$ bins <u>independently</u>

$$\omega = (1, 8, 19, 23, 6, 2)$$

ball 1 in bin 1

ball 2 in bin 8

ball 3 in bin 19

$$\mathbb{P}(\omega_1, \omega_2, \ldots, \omega_m) = \frac{1}{n^m}$$

<u>Questions:</u>

1. How long until bin 1 gets a ball (in expectation)?
2. How long until no bins are empty (in expectation)?
3. What is the most number of balls in any bin (in expectation)?

# Waiting Time

How long until bin 1 gets a ball?

Given $\omega = (\omega_1, \omega_2, \ldots)$ $\quad X(\omega) =$ minimum $i$ s.t. $\omega_i = 1$

$$\mathbb{E}(X) = \sum_{x=1}^{\infty} x \cdot \underbrace{\mathbb{P}(X=x)}$$

What is the probability I wait exactly $x$

$$= \sum_{x=1}^{\infty} x \cdot \left(1-\frac{1}{n}\right)^{x-1} \frac{1}{n}$$

$$= n$$

$$\mathbb{P}(X=x) = \mathbb{P}\Big(\overbrace{(\omega_1 \neq 1)}^{E_1 = \{\omega : \omega_1 \neq 1\}} \wedge (\omega_2 \neq 1) \wedge \ldots \wedge (\omega_{x-1} \neq 1) \wedge (\omega_x = 1)\Big)$$

$$= \left(\prod_{i=1}^{x-1} \mathbb{P}(\omega_i \neq 1)\right) \cdot \mathbb{P}(\omega_x = 1)$$

$$= \left(1-\frac{1}{n}\right)^{x-1} \frac{1}{n}$$

# Waiting Time

Suppose you have independent events $E_1, E_2, \ldots$

$$\mathbb{P}(E_1) = \mathbb{P}(E_2) = \ldots = p$$

$$\mathbb{E}(\text{first } i \text{ s.t. } E_i \text{ occurs}) = \frac{1}{p}$$

# Coupon Collector

How long until no more empty bins?

$\omega = (\omega_1, \omega_2, \ldots)$

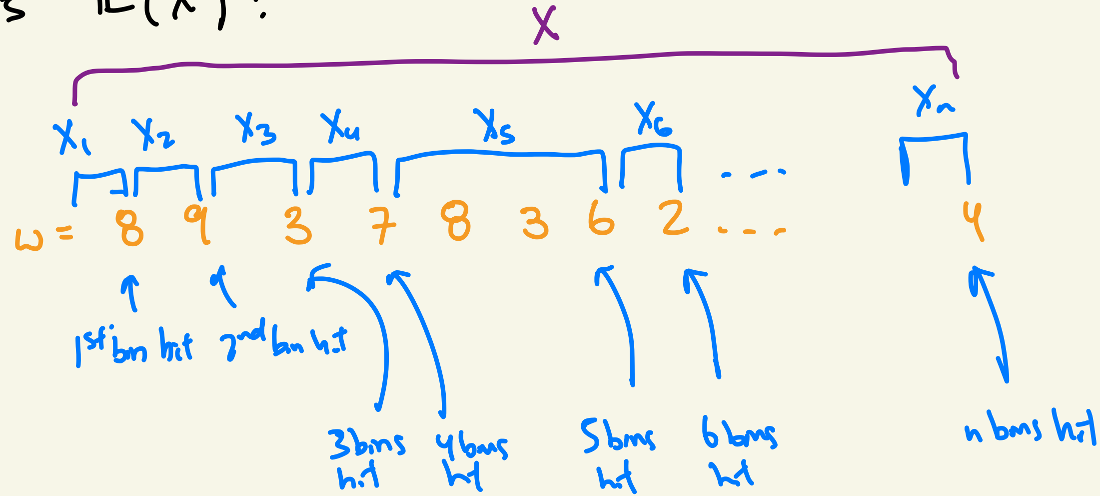
$X(\omega) =$ first ball $i$ such that every bin got at least one ball $\leq i$

What is $\mathbb{E}(X)$?

$$\omega = \overbrace{\overbrace{8}^{X_1}\ \overbrace{9}^{X_2}\ \overbrace{\quad 3}^{X_3}\ \overbrace{7}^{X_4}\ \overbrace{8\ \ 3\ \ 6}^{X_5}\ \overbrace{2}^{X_6}\ \cdots \qquad \overbrace{4}^{X_n}}^{X}$$

1st bin hit, 2nd bin hit, 3 bins hit, 4 bins hit, 5 bins hit, 6 bins hit, n bins hit

What is $\mathbb{E}(X)$?

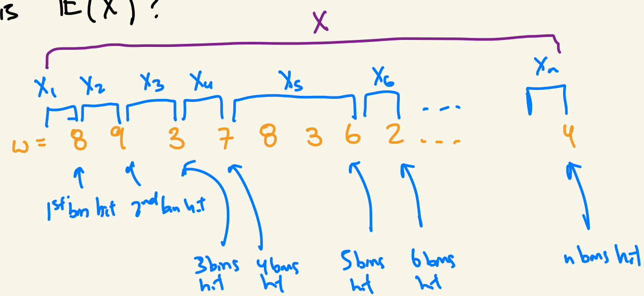



What is $\mathbb{E}(X_i)$?

$$\mathbb{E}(X_i) = 1/\mathbb{P}(\text{hit one of the } n-i+1 \text{ empty bins})$$

$$= 1\Big/\left(\frac{n-i+1}{n}\right) = \frac{n}{n-i+1}$$

What is $\mathbb{E}(X)$?





What is $\mathbb{E}(X_i)$?

$$\mathbb{E}(X_i) = 1/\mathbb{P}(\text{1 hit one of the } n-i+1 \text{ empty bins})$$

$$= 1/\left(\frac{n-i+1}{n}\right) = \frac{n}{n-i+1}$$

Magic of linearity

$$\mathbb{E}(X) = \mathbb{E}(X_1 + X_2 + \ldots + X_n) = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \ldots \frac{n}{3} + \frac{n}{2} + \frac{n}{1}$$

$$= n \cdot \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n}\right)$$

$$\approx n \cdot \ln(n)$$

Maximum Load

# Markov's Inequality

# Chebyshev's Inequality

# Applying Chebyshev to Balls and Bins

# Chernoff Bounds

Let $Z_1, \ldots, Z_n$ be independent r.v.'s such that

$$Z_i = \begin{cases} 1 & \text{w.p. } p_i \\ 0 & \text{w.p. } 1-p_i \end{cases}$$

and $Z = Z_1 + \ldots + Z_n$. Let $\mu = \mathbb{E}(Z) = p_1 + \ldots + p_n$

<u>Thm:</u> $\mathbb{P}(Z > (1+\delta)\mu) \leq \left(\dfrac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$

# Chernoff Bound Proof

Let $Z_1, \dots, Z_n$ be independent r.v.'s such that

$$Z_i = \begin{cases} 1 & \text{w.p. } p_i \\ 0 & \text{w.p. } 1-p_i \end{cases}$$

and $Z = Z_1 + \dots + Z_n$. Let $\mu = \mathbb{E}(Z) = p_1 + \dots + p_n$

<u>Thm:</u> $\mathbb{P}(Z > (1+\delta)\mu) \leq \left( \frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \right)^{\mu}$

$$\mathbb{P}(Z > a\mu) = \mathbb{P}(e^{tZ} > e^{ta\mu})$$

$$\leq e^{-ta\mu} \cdot \mathbb{E}(e^{tZ})$$

$$= e^{-ta\mu} \cdot \prod_{i=1}^{n} \mathbb{E}(e^{tZ_i}) = e^{-ta\mu} \cdot \prod_{i=1}^{n} (p_i e^t + 1 - p_i) = e^{-ta\mu} \cdot \prod_{i=1}^{n} (1 + p_i(e^t - 1))$$

$$\leq e^{-ta\mu} \cdot \prod_{i=1}^{n} e^{p_i(e^t-1)} = e^{-ta\mu} \cdot e^{(e^t-1) \cdot \sum_{i=1}^{n} p_i} = e^{-ta\mu} \cdot e^{(e^t-1)\mu}$$

$$= e^{(e^t-1)\mu - ta\mu}$$

$(a = (1+\delta))$

$$= \left( e^{e^t - 1 - ta} \right)^{\mu} \longleftarrow$$

Set $a = 1+\delta$, plug in the right value of $t$

# Applying Chernoff to Balls and Bins