CS7800: Advanced Algorithms

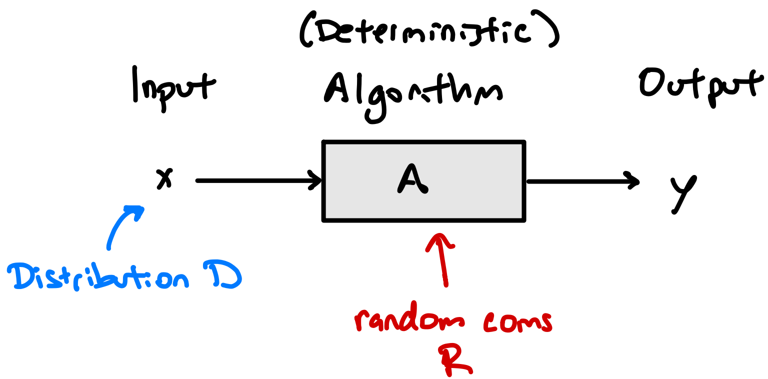
Class 20: Randomized Algorithms I

- Probability Toolkit
- Load Balancing / Balls and Bins

Jonathan Ullman

November 18, 2025

Randomness in Algorithms



Correctness: For every x,
$y = A(x, r)$ is correct 99% of the time

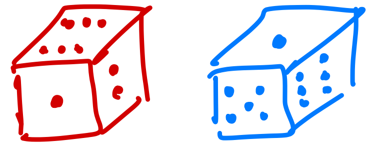
Running Time: For every x,
$A(x)$ ~~always runs in time $T(|x|)$~~
runs in time $T(|x|)$ on average

Kinds of Randomness:

- "Average-case analysis" (Deterministic algorithm, Random input) NOT OUR TOPIC
- "Randomized algorithms" (Worst case, Random algorithm) THIS CLASS
  - We don't believe randomized algorithms solve NP-hard problems
  - Randomized algs can be simpler and faster* (eg primality)
  - In many models, randomness is essential

(Discrete) Probability Toolkit

- Outcomes $\omega \in \Omega$



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

$\omega = (6, 1)$

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

① $\mathbb{P}(\omega) \geq 0$ ② $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$

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

- Events $E \subseteq \Omega$

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

- Probability of an event is $\mathbb{P}(E) = \sum_{\omega \in \Omega} \mathbb{P}(\omega)$

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

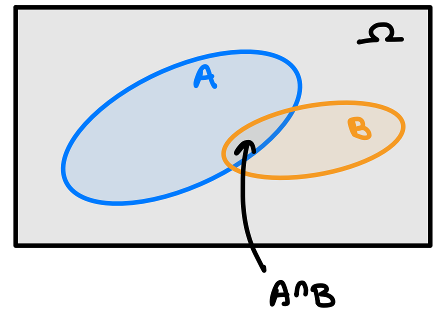
- Can take complements ("not"), unions ("or"), intersections ("and")

Conditional Probability and Independence

- Conditional probability

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

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

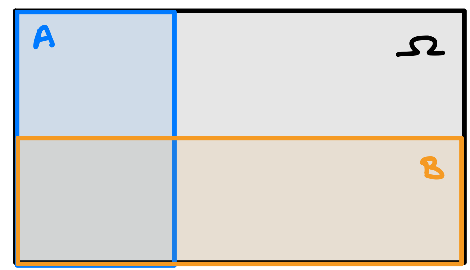


- Independence

A and B are independent

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

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



$\mathbb{P}(A) = 1/6$ $\mathbb{P}(A \cap B) = 1/3$
$\mathbb{P}(B) = 1/2$

Ex. $A = \{\omega_1 = 3\}$
$B = \{\omega_2 = 3\}$ Independent

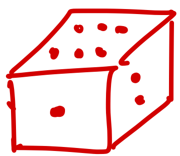
Ex. $A = \{\omega_1 = 3\}$
$B = \{\omega_1 \text{ is odd}\}$ Not independent

Random Variables (neither random nor variable)

- A random variable maps an outcome to a value

$\Omega = \{1,2,3,4,5,6\}$

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



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

- We treat integer-valued random variables as variables with unknown value

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

$$\mathbb{P}(X^2 \leq k) = \mathbb{P}(\{\omega : X(\omega)^2 \leq k\})$$

Expected Value

- The expected value of an integer random variable X is

"average" "mean"

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

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

- Does not assume independence

- X and Y are independent if $\mathbb{P}(X = k \wedge Y = \ell) = \mathbb{P}(X = k)\,\mathbb{P}(Y = \ell)$

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

Probability Case Study: Balls and Bins

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

$$\Omega = \{1, \ldots, n\}^m$$

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

$$\omega = (6, 8, 11, 2, 37, \ldots)$$

Ball 1
Bin 6

Ball 2
Bin 8

Questions:

1. How long until bin 1 gets a ball?
2. How long until no bin is empty?
3. What is the maximum number of balls in any bin?

Waiting Time

> **Fact.** If X is a non-negative integer r.v.
> $$\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k)$$

How long until bin 1 gets a ball?

> X is the random variable whose output is the first ball i that lands in bin 1.

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

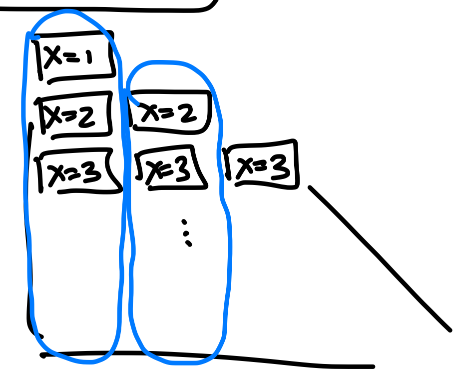
$$\sum_{k=1}^{\infty} k \cdot \mathbb{P}(X=k)$$

$$= \sum_{k=1}^{\infty} k \cdot \prod_{i=1}^{k-1} \mathbb{P}(w_i \neq 1) \cdot \mathbb{P}(w_k = 1)$$

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

$$\sum_{k=1}^{\infty} k \cdot \mathbb{P}(X=k)$$

$$= \sum_{k=1}^{\infty} \mathbb{P}(X \geq k)$$



$\mathbb{P}(X \geq 1)$ $\mathbb{P}(X \geq 2)$ ...

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k)$$

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

Waiting Time

- Suppose we repeat an experiment independently until some "desired outcome" happens
- Every time you run the experiment $\mathbb{P}(\text{"desired outcome"}) = p$
- $\mathbb{E}(\text{\# of trials until "desired outcome"}) = 1/p$

Coupon Collector

How long until no empty bins?

Directly computing $\mathbb{E}(X)$ is hard

- X = the first time t where every bin has ≥ 1 ball

$\omega = (6, 11, 17, 6, 3, 2, 3, 4, 11, 11 \ldots$

$X_3 = 1$ $\quad$ $X_5 = 1$

$X_1 = 1$ $\quad$ $X_2 = 1$ $\quad$ $X_4 = 2$ $\quad$ $X_6 = 2$

= time when 1st bin fills

= time btw 1st and 2nd bin fills

X_k is how long we waited to hit one of the $n-k+1$ empty bins

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{k=1}^{n} X_k\right)$$

$$= \sum_{k=1}^{n} \mathbb{E}(X_k) \quad \text{[linearity]}$$

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

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

$$= n \cdot \Theta(\log n) = \Theta(n \log n)$$