CS1800
Discrete Structures
Fall 2019

Lecture 14
10/25/19

Last time

(Exam)

Finishing Probability

Today

- Last topic in probability
  - Birthday Paradox
  - indicator r.v.

- Algorithmic Analysis
  - sequences
  - series
  - recurrence
  - induction

Next time

- Continue Alg. Analysis

Birthday Paradox

Q: How many people do we need s.t. at least 50/50 chance that 2 (or more) people share a b-day?

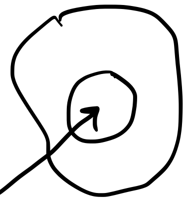
A1: PHP says that 366 people <u>guarantee</u> a collision

A2: actually closer to 23 or 24

Calculate likelihood that all 244 students in class have <u>different</u> b-days.

- # possible ways of assigning b-days to 244 students is

$$\underbrace{365 \cdot 365 \cdot \ldots 365}_{244} = 365^{244}$$



- # possible ways of assigning b-days w/ no collisions

$$365 \cdot 364 \cdot 363 \cdots = 365P_{244}$$

$$\text{Prob. of no collision} = \frac{\overbrace{365 \cdot 364 \cdot 363 \cdots 122}}{\underbrace{365 \cdot 365 \cdot 365 \cdots 365}_{244}} \approx 1.96 \times 10^{-48} \; !$$

approximation for n students...

$$= \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots$$

$$= \frac{365-0}{365} \cdot \frac{365-1}{365} \cdot \frac{365-2}{365} \cdots \frac{365-(n-1)}{365}$$

$$= \left(\frac{365}{365} - \frac{0}{365}\right) \cdot \left(\frac{365}{365} - \frac{1}{365}\right) \cdots \left(\frac{365}{365} - \frac{n-1}{365}\right)$$

$$= \left(1 - \frac{0}{365}\right) \cdot \left(1 - \frac{1}{365}\right) \left(1 - 2/365\right) \cdots \left(1 - \frac{n-1}{365}\right)$$

$$\approx e^{-0} \cdot e^{-\frac{1}{365}} \cdot e^{-2/365} \cdots e^{-\frac{n-1}{365}}$$

decent approximation when n is "small"

$$= e^{-\left(0 + 1/365 + 2/365 + 3/365 + \cdots + \frac{n-1}{365}\right)}$$

$$= e^{-\frac{1+2+\cdots(n-1)}{365}} = e^{-\frac{n(n-1)}{2 \cdot 365}} = e^{-\frac{n(n-1)}{730}}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots$$

when x small

$$e^x \approx 1 + x$$

$$1 + x \approx e^x \qquad 1 - x \approx e^{-x}$$

- prob of no collision $= e^{-\frac{n(n-1)}{730}}$

- 23 or 24

$$n=23 \qquad e^{-\frac{23(22)}{730}} = .4999998$$

- exact answer is $\frac{365P23}{365^{23}} = 0.4927$

---

another approximation + intuition

- for 2 students, chance they don't share a b-day is $\frac{364}{365} \approx 0.99726$

- but with 23 students, there are $\binom{23}{2} = \frac{23 \cdot 22}{2} = 253$ such pairs of students

- the chance that <u>all</u> such pairs simultaneously fail to share b-days is <u>approximately</u> $\left(\frac{364}{365}\right)^{\binom{23}{2}} \approx 0.4995$

the events are not perfectly independent, so this is an approximation

Q: How many collisions? (shared b-days)

- Let r.v. X = # b-days used
- want $E\{X\}$

- Let X_1 = did I use b-day 1 (Jan 1)
  X_2 = " " 2 (Jan 2)
  ⋮
  X_{365} = " " 365 (Dec 31)

  indicator r.v. 0, 1 value

- $X = X_1 + X_2 + \cdots X_{365}$

- $E\{X\} = E\{X_1 + X_2 + \cdots X_{365}\} = E\{X_1\} + E\{X_2\} + \cdots + E\{X_{365}\}$

- $E\{X_1\} = \sum_x x \cdot \Pr\{X_1 = x\} = 0 \cdot \Pr\{X_1 = 0\} + 1 \cdot \Pr\{X_1 = 1\}$
  $= \Pr(X_1 = 1) = 1 - \Pr\{X_1 = 0\}$
  $= 1 - (364/365)^{244}$

no one has Jan 1 as b-day

$$E\{x_1\} = Pr\{x_1=1\} = 1 - Pr\{x_1=0\} = 1 - \left(\frac{364}{365}\right)^{244} = .488$$

$$\forall i \quad E\{x_i\} = 0.488$$

$$E\{x\} = E\{x_1 + x_2 + \cdots + x_{365}\}$$
$$= E\{x_1\} + E\{x_2\} + \cdots E\{x_{365}\}$$
$$= 365 \cdot 0.488$$
$$= 178.12 \leftarrow \text{expected \# b-days used}$$

- but... 244 students

so, $244 - 178.12 = 65.88$ more students than b-days, in expectation, so lots of shared birthdays.

Next module: Algorithmic Analysis & Related Math

- Consider the search problem
- My dictinary has n pages
- First alg.: Linear search – go through one page at a time, in order, until find page looking for

  - best case scenario: on 1^{st} page: 1
  - worst case: on last page: n
  - average case: suppose equally lively that word is on any of $1 \to n$ pages

$$= \frac{1}{n}\cdot 1 + \frac{1}{n}\cdot 2 + \frac{1}{n}\cdot 3 + \cdots \frac{1}{n}\cdot n$$

$$= \frac{1+2+3+\cdots n}{n} = \frac{n(n+1)/2}{n} = \frac{n+1}{2}$$

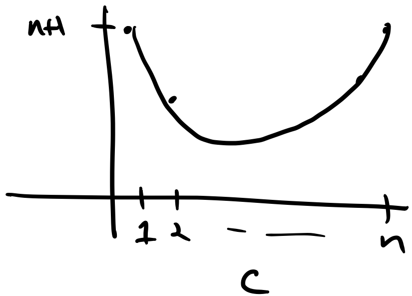
Chunk search:

- look at chunks of pages c at a time
- find correct chunk
- do linear search within chunk of size c

How fast? $\quad$ # chunks $= \frac{n}{c}$

Worst case:
- have to look at all n/c chunks
- have to look at every page in chunk

$$= \frac{n}{c} + c \qquad \frac{n}{c} = c$$



$$f(x) = \frac{n}{x} + x = n \cdot x^{-1} + x$$

$$f'(x) = -nx^{-2} + 1 = 0$$

$$nx^{-2} = 1$$

$$x^{-2} = 1/n \qquad x^2 = n \qquad x = \sqrt{n}$$

$$\boxed{\frac{n}{c} + c = \frac{n}{\sqrt{n}} + \sqrt{n} = \sqrt{n} + \sqrt{n} = 2\sqrt{n}}$$

Binary search

k cuts in half	size of book
0	n
1	$n/2$
2	$n/4 = n/2^2$
3	$n/8 = n/2^3$
$\vdots$	
k	$n/2^k$

what size k yields just one page left?

$$n/2^k = 1 \iff 2^k = n$$

$$\Rightarrow k = \log_2 n$$

T(n) = "time" (or # operations) to solve a problem of size n

		n = 1000	n = 1,000,000
l.s.	$T(n) = n$	1000	1,000,000
c.s.	$T(n) = 2\sqrt{n}$	64	2000
b.s.	$T(n) = \log_2 n$	10	20

$1000 \xrightarrow{15.75x} 64 \xrightarrow{6.4x} 10$, $1000 \xrightarrow{100x} 10$

$1{,}000{,}000 \xrightarrow{500x} 2000 \xrightarrow{100x} 20$, $1{,}000{,}000 \xrightarrow{50{,}000x} 20$

A couple of years ago--

- Fastest super computer in U.S.

  CRAY XK7 "titan" at Oak Ridge National Labs

  17.59 peta flops

  17.59×10^{15} floating point ops/sec

- Mac Pro 7 teraflops

  7×10^{12}

$$\frac{17.59 \times 10^{15}}{7 \times 10^{12}} = 2{,}513x$$

n = 1,000,000

l.s. on a supercomp.

is $\frac{50{,}000}{2{,}513} = 20x$

<u>slower</u> than

b.s. on a

Mac.