lecture 19

average case

- amortization
- discrete probability
  https://class.coursera.org/algo-004/lecture/29
  - random variables
  - linearity of expectation

Common notion of "average"

$$\frac{\sum_{i=1}^{n} x_i}{n}$$

Sometimes averages do not involve random events i.e. probabilities. Sometimes they do.

Averages that don't involve probabilities

Suppose you devote 40 hours per week to your studies.
You take 5 courses: 8 work hours per week per course.

The semester is 13 weeks. Thus, you devote 104 hours per course.

For COMP 251, this 104 hours breaks down to:

39 hours of scheduled lecture time (3 hours per week)

~40 hours of review/exercises, including studying for exams
(~3 hours average per week)

~25 hours = 6 hours *4 assignments (~2 hours average per week)

When I say that you spend an average of 2 hours per week on assignments (etc),

I don't have any "randomness" in mind.

Rather, I mean that this is the amount of work you do per week, averaged or "amortized" over the semester.

Wikipedia:

Amortization is the process of accounting for an amount over a period.

Example 2

When you buy a house for 500K, you don't pay the whole amount up front and then live in it for free for n = 25 years. Rather, you pay say 100K and the bank pays 400K, and you make regular (equal) mortgage payments to the bank for n = 25 years. The amount is determined by a amortization table.

There is nothing random going on here.

Wikipedia: In computer science, **amortized analysis** is a method of analyzing the execution cost of algorithms over a sequence of operations.

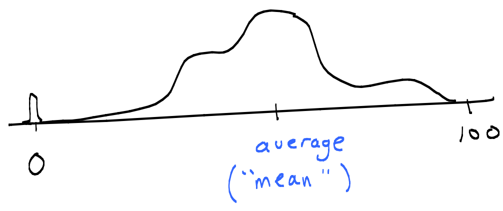
e.g.
- building a heap fast (recall lecture 4)
- building an Arraylist or Hashtable with n elements

  (See Exercises)
- performing n union-find ops.

There is nothing random going on when you do amortized analysis.

Other times when we talk about averages, e.g. average grade, we have in mind a random distribution

Here we need the language of probability theory.



When we say quicksort is $O(n \log n)$ in the "average case", we could mean two different things:

1) A quicksort algorithm that chooses pivots in some deterministic way has average performance $O(n \log n)$, namely averaging over all possible inputs which are equally like to occur. — randomized input

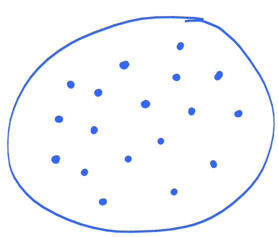
2) For any given input, a quicksort algorithm that chooses pivots randomly takes time $O(n \log n)$. — randomized algorithm (NEXT CLASS)

Probability Theory

Definition: sample space S.

a set of possible outcomes of some "experiment" (loose definition)

S

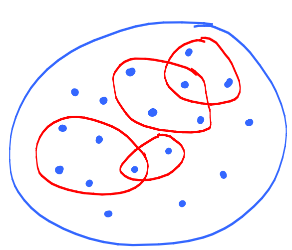


Sample space: Examples

- flip a coin $\{H, T\}$
- flip a coin 3 times $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- roll a die once $\{1, 2, 3, 4, 5, 6\}$
- roll a die twice $\{(1,1), (1,2), \ldots (5,6), (6,6)\}$
- write a midterm $\{0, \frac{1}{2}, 1, 1\frac{1}{2}, \ldots 29\frac{1}{2}, 30\}$
- take a course $\{A, A-, B+, B, \ldots C, D, F\}$
- when is your birthday? $\{1, 2, 3, 4, \ldots 365\}$
- instance of a sorting algorithm of size 3
  $\{(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)\}$

Definition:

An event is a subset of a sample space.



4 different events shown here

Q: If a sample space has n outcomes, $|S| = n$, how many events can one define?

A: 2^n.

(Usually we are interested in only some of these 2^n possible events.)

Events: Examples

- you flip a coin 3 times and you get m heads e.g. $m = 2$ $\{THH, HTH, HHT\}$
- you roll a die twice and the sum is less than m eg. $m = 5$
  $\{(1,1), (1,2), (2,1), (1,3), (2,2), (3,1)\}$
- you get some letter grade on the midterm
  e.g. A $\{25.5, 26, 26.5, \ldots\}$
- your birthday is after some date
  e.g. after February 5 $\{36, 37, \ldots 365\}$

Infinite Sample Spaces

- Discrete probability allows us to talk about finite or "countably infinite" sample spaces i.e. the integers.
  e.g. S = the number of times you flip a coin until you get heads.

- If you want to use the real numbers as your sample space, you need continuous probability (we won't go there)

Probability distribution on a sample space

definition:

mapping $p : S \to [0,1]$

$$\sum_{s \in S} p(s) = 1$$

e.g. uniform probability distribution

$|S| = n$, $p(s) = \frac{1}{n}$ for all $s \in S$.

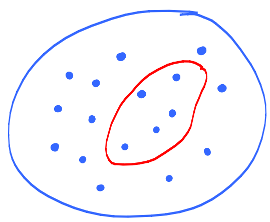
Note: if S is all integers, you cannot define a uniform probability distribution.

e.g. flip a coin

$$p(H) = p(T) = \frac{1}{2}$$

e.g. roll a die

$$p(1) = p(2) = \dots = p(6) = \frac{1}{6}$$



Let $E \subseteq S$ be some event.

$$p(E) = \sum_{s \in E} p(s)$$

Note:

- $p(\{\}) = 0$ (no outcomes), $p(S) = 1$ (all outcomes)
- If $p(\,)$ is a uniform distribution the $p(E) = \frac{|E|}{|S|}$.

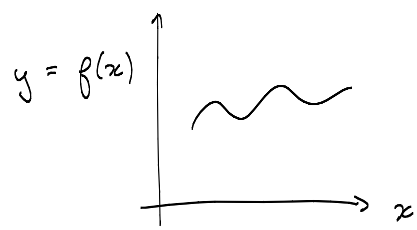
Random Variable

mapping $X : S \to \mathbb{R}$

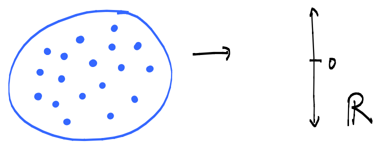
e.g. we roll a di three times.

- What is sum?
- How many times did a 6 appear?
- what is largest value?

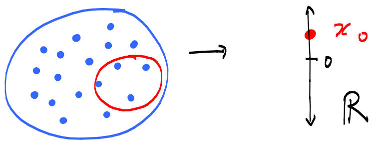
Note:
a random variable
- is not a mapping from events to numbers
- is not random
- is not a variable (well, it sort of is .... see next slide)



f is a mapping from $\mathbb{R}$ to $\mathbb{R}$.
f is not a variable.
However, when we say $y = f(x)$, then y is a "variable." In the case of a random variable, the x values (and hence the y values) occur with some probability.

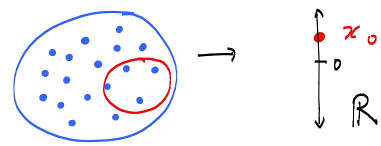
$X :$ S 

Define the event $X = x_0$ to be $\{ s \in S \text{ such that } X(s) = x_0 \}$.

$X :$ 

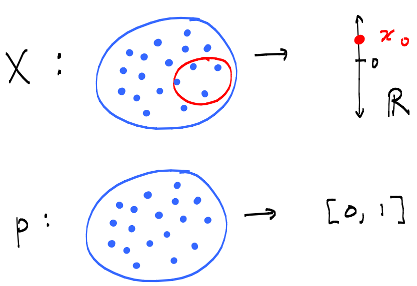
Examples $\{ X = x_0 \}$

- you flip a coin 3 times and you get $X = x_0$ heads e.g. $x_0 = 2$ { THH, HTH, HHT }
- you roll a die twice and the sum is $X = x_0$ eg. $x_0 = 5$ { (1,4), (2,3), (3,2), (4,1) }



Define $\quad Pr(X = x_0) \equiv \sum_{\{s \,:\, X(s) = x_0\}} p(s)$.

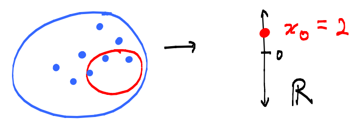
We call this a "distribution on the random variable X." Think of it as probabilities on the values of the random variable X.



Note: we need two mappings to define this distribution $Pr(X = x)$

e.g. you flip a coin 3 times. What is the distribution of the number of heads?

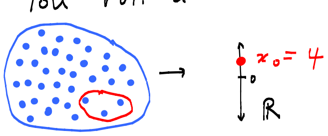
$x_0 = 2$ { THH, HTH, HHT }



$|S| = 8$

$$Pr(X = x) = \begin{cases} \frac{1}{8}, & x = 0 \\ \frac{3}{8}, & x = 1 \\ \frac{3}{8}, & x = 2 \\ \frac{1}{8}, & x = 3 \end{cases}$$

You roll a die twice and the sum is $X = x_0$. e.g. $x_0 = 4$



$|S| = 36$

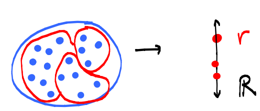
$$Pr(X = x) = \begin{cases} 1/36, & x = 2 \\ 2/36, & x = 3 \\ 3/36, & x = 4 \\ 4/36, & x = 5 \\ 5/36, & x = 6 \\ 6/36, & x = 7 \\ 5/36, & x = 8 \\ 4/36, & x = 9 \\ 3/36, & x = 10 \\ 2/36, & x = 11 \\ 1/36 & x = 12 \end{cases}$$

Expected Value of a Random Variable

Weighted average of values taken by the random variable. The weights are the probabilities.

$$\mathcal{E} X = \sum_{s \in S} p(s) \cdot X(s)$$

$$= \sum_{r \in X(S)} Pr\{ X = r \} \cdot r$$



NOTATION

$X(S) \equiv \{ X(s) : s \in S \}$

i.e. all values reached by X.

Example

What is the expected value of a roll of a di?

<u>Note:</u> for this example, $X(s) = s$.

$$\mathcal{E}(X) = \sum_{s \in S} p(s) \cdot X(s) = \sum_{r \in X(S)} \Pr\{X = r\} \cdot r$$

$$= \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6$$

$$= 3.5$$

<u>note:</u> $E(X)$ does not have to be one of the values taken by X.

Example

What is the expected value of the sum of two rolls of a di?

$$\mathcal{E}(X) = \sum_{s \in S} p(s) \cdot X(s)$$

$$= \frac{1}{36}\left(2 + 3 + 3 + \cdots + 11 + 11 + 12\right)$$

$$= \sum_{r \in X(S)} \Pr\{X = r\} \cdot r$$

$$= \frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \frac{3}{36} \cdot 4 + \cdots \frac{1}{36} \cdot 12$$

Example:

Lets play a game. I flip a coin. You pay me \$5 if its a head. I pay you \$10 if its a tail. Anyone want to play?

What is the expected value of your winnings if you play?

$$\mathcal{E}(X) = \frac{1}{2}(-5) + \frac{1}{2} \cdot 10 = 2.5$$

Lets play the game 3 times.
Let random variable Y be your total winnings.

$$\mathcal{E}(Y) = \frac{1}{8} \cdot (30 + 15 + 15 + 0 + 15 + 0 + 0 - 15)$$

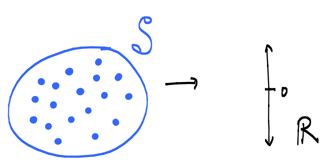
TTT, TTH, THT, THH, HTT, HTH, HHT, HHH

$$= \frac{1}{8} \cdot 30 + \frac{3}{8} \cdot 15 + \frac{3}{8} \cdot 0 + \frac{1}{8} \cdot (-15)$$

$$= 7.5$$

Let $Y = X_1 + X_2 + \cdots + X_m$ be a sum of random variables.

for $i = 1$ to m, $X_i :$



<u>Note:</u> Y also maps S to $\mathbb{R}$.

Linearity of Expectation:

$$\mathcal{E}\, Y = \sum_{i=1}^{m} \mathcal{E}\, X_i$$

"the expected value of the sum is the sum of the expected values"

Linearity of Expectation (proof):

$$\mathcal{E}Y = \sum_{s \in \mathcal{S}} p(s)\, Y(s)$$
$$= \sum_{s \in \mathcal{S}} p(s)\left(X_1(s) + X_2(s) + \cdots + X_m(s)\right)$$
$$= \sum_{s \in \mathcal{S}} \sum_{i=1}^{m} p(s)\, X_i(s)$$
$$= \sum_{i=1}^{m} \sum_{s \in \mathcal{S}} p(s)\, X_i(s)$$
$$= \sum_{i=1}^{m} \mathcal{E} X_i$$

<u>Example 2</u> (Example 1 was the coin flip game)

Suppose we roll 4 dice.
What is the expected value of the sum?
The sample space $\mathcal{S}$ has 6^4 outcomes, each with probability $1/6^4$.

Applying linearity of expectation:

$$\mathcal{E}Y = 4\,\mathcal{E}X \leftarrow \text{roll of one di}$$
$$= 4 \cdot 3.5$$
$$= 14$$

<u>Example 3</u>

Suppose we flip a coin until we get a head.

$$\mathcal{S} = \{H, TH, TTH, TTTH, \ldots\} \text{ outcomes}$$

$$\Pr(X=1) = \tfrac{1}{2}$$
$$\Pr(X=2) = \tfrac{1}{4}$$
$$\Pr(X=3) = \tfrac{1}{8}$$

$$\Pr(X=i) = \frac{1}{2^i}.$$

What is the expected value of the number X of coin flips?

$$\mathcal{E}X = \sum_{i=1}^{\infty} i \Pr(X=i)$$
$$= \sum_{i=1}^{\infty} i \left(\tfrac{1}{2}\right)^i$$

recall lecture 4 (heaps)

$$= x \sum_{i=1}^{\infty} i\,(x)^{i-1}, \quad x = \tfrac{1}{2}$$
$$= \frac{1}{2} \sum_{i=1}^{\infty} \frac{d}{dx} x^i$$
$$= \frac{1}{2} \frac{d}{dx} \sum_{i=1}^{\infty} x^i$$
$$= \frac{1}{2} \frac{d}{dx} \left(\sum_{i=0}^{\infty} x^i - 1 \right)$$
$$= \frac{1}{2} \frac{d}{dx} \left(\frac{1}{1-x} - 1 \right)$$
$$= \frac{1}{2} \frac{d}{dx} \left(\frac{x}{1-x} \right)$$
$$= \frac{1}{2} \frac{1-x-(-1)x}{(1-x)^2}$$
$$= \frac{1}{2} \frac{1}{(1-x)^2}$$

substitute $x = \frac{1}{2}$

$$= 2$$

We will use this result next class.

A few classes from now (lecture 22) I will return to this example and consider "unfair" coins, also known as Bernoulli trials.