Randomness + probability uses in CS:

- randomized algorithms
- data structures using randomness
- modeling real-world phenomena

But first, we need to learn to count!

Sum rule: If $A \cap B = \emptyset$, then $|A \cup B| = |A| + |B|$.

Product rule: The number of pairs (x, y) with $x \in A$, $y \in B$ is $|A| \cdot |B|$.

$$|A \times B| = |A| \cdot |B|$$

ex A restaurant has 2 lunch specials.

1. soup or salad
2. soup and salad

If A = set of soups = {chicken noodle, ...}
B = set of salads = {caesar, ...}

How many possibilities are there for

1: $|A \cup B| = |A| + |B|$

2: $|A \times B| = |A| \cdot |B|$

More general product rule:

$$|A_1 \times A_2 \times A_3 \times \cdots \times A_k| = |A_1| \cdot |A_2| \cdot |A_3| \cdots |A_k|$$

ex How many 32-bit string are there?

$$\underbrace{010 \cdots 001}_{\text{32-bit string}}$$

2^{32} by the generalized product rule.

$$|\underbrace{\{0,1\} \times \{0,1\} \times \{0,1\} \times \cdots \times \{0,1\}}_{\text{32 times}}|$$

$$= \underbrace{|\{0,1\}| \cdot |\{0,1\}| \cdot |\{0,1\}| \cdots |\{0,1\}|}_{\text{32 times}} = 2^{32}$$

ex How many MAC addresses are there?

(you don't need to evaluate the value)

16^{12}

$$\underbrace{12:AC:D9:03:F9:7B}_{\text{12 digits}}$$

$$|\underline{\{0,1,2,\cdots,9,A,B,\cdots,F\}} \times \{0,1,2,\cdots,9,A,B,\cdots,F\} \times \cdots|$$

per pair: 16^2 possibilities.

$$(16^2)^6 = 16^{12}$$

Inclusion-Exclusion rule:

$|A \cup B| = |A| + |B| - |A \cap B|$ → 0 if empty

ex let $O = \{1,3,5,7,9\}$ and $P = \{2,3,5,7\}$
what is $|O \cup P|$?

$|O \cap P| = |\{3,5,7\}| = 3$

$|O \cup P| = |O| + |P| - |O \cap P| = 5 + 4 - 3 = 6$

(double check: $O \cup P = \{1,2,3,5,7,9\}$)

0123, 7980, 0111

ex How many invalid PINs are there?

hint: it's btwn 150 and 250

Let S denote the set of PINs starting w/ 3 repeated digits.

eg 1110 $\quad |S| = 10^2 = 100$
2223
3332

Let E denote the set of PINs ending w/ 3 repeated digits.

eg 0111 $\quad |E| = 100$
2333

$S \cap E$: all digits same

$|S \cap E| = 10$

$|S \cup E| = |S| + |E| - |S \cap E| = 100 + 100 - 10$
$= 190$

<u>Def</u> Given some random process, the <u>sample space</u> S is the set of all possible <u>outcomes</u>.

A <u>probability function</u> $Pr : S \to \mathbb{R}$ describes the fraction of the time that $s \in S$ occurs.

$f : A \to B$
$f(a) = b$

$Pr[s]$

$$\sum_{s \in S} Pr[s] = 1$$

$$Pr[s] \geq 0 \quad \forall s \in S$$

<u>ex</u>

flipping a fair coin

$S = \{\text{heads}, \text{tails}\}$

$Pr[\text{heads}] = 0.5$
$Pr[\text{tails}] = 0.5$

$$\sum_{s \in S} Pr[s] = Pr[\text{heads}] + Pr[\text{tails}] = 0.5 + 0.5 = 1$$

drawing a card

$S = \{2 \text{ clubs}, 3 \text{ clubs}, \ldots \}$

$Pr[s] = \frac{1}{52} \quad \forall s \in S$

flipping 2 fair coins

$S = \{(H,H), (H,T), (T,H), (T,T)\}$

each has $Pr[s] = 0.25$

<u>ex</u> let $S = \{0, 1, 2, \ldots, 7\}$. Choose from S by flipping 7 coins and counting # of heads.

HHHHHHH $\rightarrow$ 7

$Pr[7] = 0.0078$
$Pr[4] = 0.2734$

<u>Def</u> A set of outcomes is called an <u>event</u>.

$E \subseteq S$, $Pr[E] = \sum_{s \in E} Pr[s]$

<u>ex</u>

when flipping 2 coins, the probability that at least one is heads is

$0.25 + 0.25 + 0.25 = 0.75$

$S = \{(H,H), (H,T), (T,H), (T,T)\}$

$E \subseteq S$

$E = \{(H,H), (H,T), (T,H)\}$

when drawing 1 card from a 52-card deck, the prob. that it is an ace is $4/52$.

Theorem 10.4. (Properties of event probabilities)

Let S be a sample space and $A \subseteq S$, $B \subseteq S$ be events. Let $\bar{A} = S - A$.

$Pr[S] = 1$ $\qquad$ $Pr[\bar{A}] = 1 - Pr[A]$

$Pr[\emptyset] = 0$

ex When drawing 1 card, what is the probability that it's not an ace?

$S = \{\text{all cards}\}$

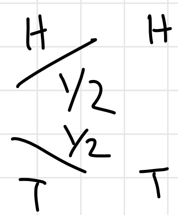
$A = \{$A clubs, A spades, A hearts, A diamonds$\}$

$$Pr[\bar{A}] = 1 - Pr[A] = 1 - 4/52 = 48/52 = 24/26 = 12/13$$

Tree Diagrams in Probability

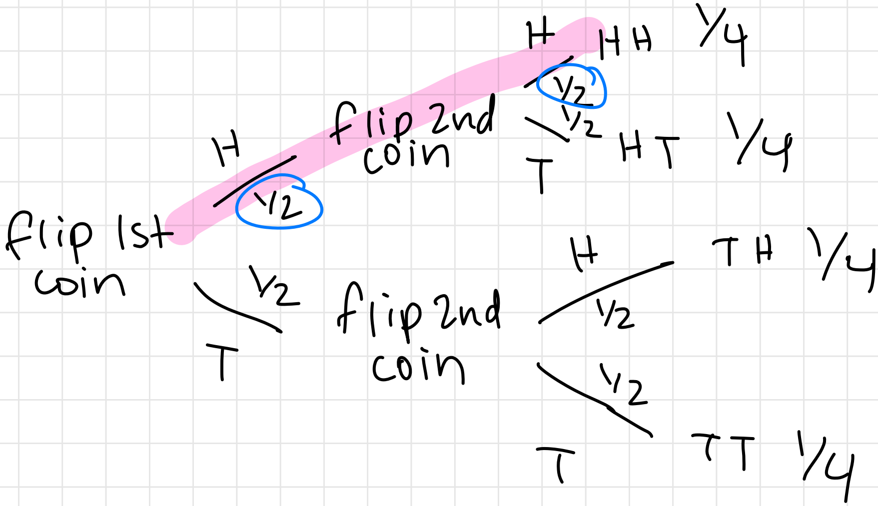
- internal nodes represent random choices, labeled w/ probability of each outcome

flip coin



Leaves are outcomes

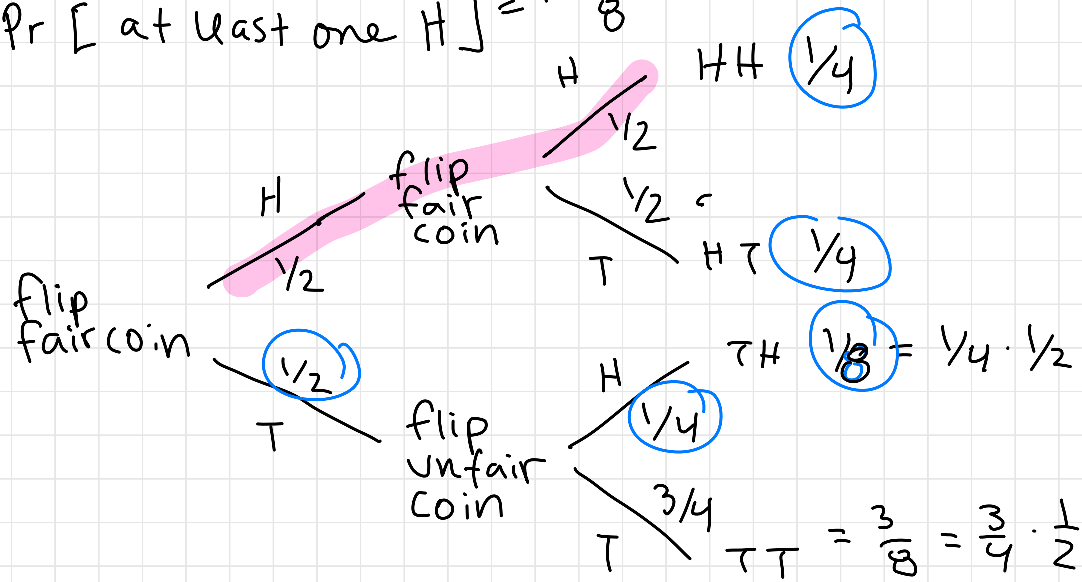
flip 2 coins



flip 1 fair coin. If H, flip a 2nd fair coin. If T, flip a coin w/ 0.75 prob. of T.

$Pr[(T,T)] = \frac{3}{8}$

$Pr[\text{at least one } H] = 1 - \frac{3}{8}$



Def A <u>permutation</u> of a set S is a length $|S|$ sequence of elements of S with no repetitions.

ex $S = \{1, 2, 3, 4\}$

$(1, 2, 3, 4)$ ✓
$(2, 4, 3, 1)$ ✓
$(2, 2, 4, 1)$ ✗
$(1, 3, 4)$ ✗

Thm 9.8 Let S be a set w/ $|S| = n$. The number of permutations of S is $n!$.

Proof sketch #1: by product rule. $|A \times B| = |A| \cdot |B|$.

Let S_1 be S - first choice, S_2 be S_1 - 2nd choice, …, all the way to S_{n-1}.

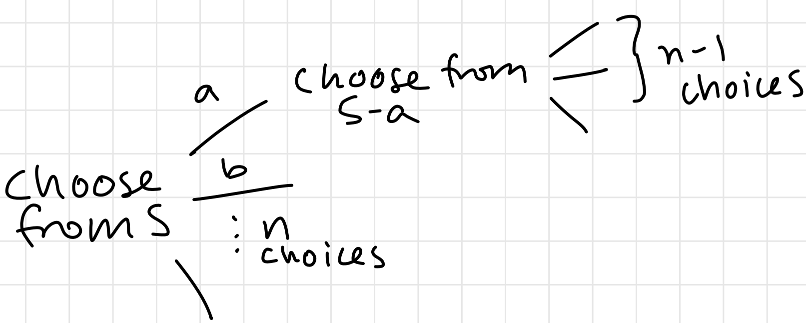
$$|\underline{S} \times \underline{S_1} \times \underline{S_2} \times \cdots \times \underline{S_{n-1}}| = |S| \cdot |S_1| \cdot |S_2| \cdots |S_{n-1}|$$

$(\ \ ,\ \ ,\ \ ,\ \ldots,\ \ )$ $= n(n-1)(n-2)\cdots(1)$

from S, from S_1, from S_2, comes from S_{n-1}

$= n!$

Proof sketch #2: w/ a tree diagram



n! leaves

<u>Def</u> Let n, k be non-negative integers.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

"n choose k"

the number of ways to choose k elements from a size n set where order does not matter.

Expected Values

Ask questions like:

how many times do we have to flip a coin to get 100 heads?

<u>def</u> A random variable X assigns a numerical value to every outcome of a sample space.

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

<u>ex</u> Suppose we flip a coin 3 times.

$S = \{H, T\}^3 = \{(HHH), \ldots \}$

$\Pr[s] = \frac{1}{8} \quad \forall s \in S$

Let X be # of heads of an outcome
Y be # of consecutive T

$X(HHH) = 3$

$Y(HHH) = 0$

Def The expectation of a random variable X, denoted $E[X]$, is the average value of X.

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

$$= \sum_{y:\, \exists s \in S:\, X(s) = y} y \cdot \Pr[X = y]$$

(10.38 in book)

ex Counting heads in 3 coin flips

$X = \#$ of heads

intuition says: expected # heads is 1.5

The following added after class...

Let's compute.

$$\text{expected \# heads} = E[X] \overset{\text{by def. above}}{=} \sum_{s \in S} X(s) \cdot \Pr[s]$$

$$= X(HHH) \cdot \Pr[HHH] + X(HHT) \Pr[HHT] +$$

$$X(HTH) \cdot \Pr[HTH] + X(HTT) \Pr[HTT] +$$

$$X(THH) \Pr[THH] + X(THT) \Pr[THT] + X(TTH) \Pr[THT]$$

$+ X(TTT)\Pr[TTT]$

$= 3 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8}$
$+ 0 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$, so the def. matches our intuition.

Now let's compute $E[X]$ using the equivalent formulation above.

$$E[X] = \sum_y y \cdot \Pr[X=y] = 0 \cdot \Pr[X=0] +$$
$$1 \cdot \Pr[X=1] + 2 \cdot \Pr[X=2] + 3 \cdot \Pr[X=3]$$
$$= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8}$$ also.

Now, let's see some examples of choosing k items from n.

ex
How many different 5-card hands are there when drawn from a 52-card deck?

we must choose 5 cards but order doesn't matter. So there are

$$\binom{52}{5} = \frac{52!}{5!(47!)} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5!}$$

(note: you don't need to evaluate something like $\binom{52}{5}$ in this class, you can leave it as-is) ex (9.41 in book)

How many different 8-bit strings are there w/ exactly 2 ones?

We can think of this as choosing two indices out of 8 to equal one.

$$\binom{8}{2}$$

Now let's see a problem about expectation where combinations come into play.

ex (10.40 in book)

What is the expected number of aces in a 13-card hand?

Let X = the number of aces in a 13-card hand. So we want to compute $E[X]$.

Recall that $E[X] = \sum_{y} y \cdot Pr[X=y]$.

Since X (a 13-card hand) can equal 0, 1, 2, 3, or 4, we need to compute $Pr[X=0]$, $Pr[X=2]$, ..., $Pr[X=4]$.

What is, for example, $Pr[X=0]$?

Since a probability is the fraction of the time an outcome occurs, it's

$$\frac{\#\text{ of ways to get 0 aces}}{\#\text{ of ways to draw 13 cards}}.$$

Let's think about # ways to draw 13 cards first. Since we are choosing 13 cards from 52, this is $\binom{52}{13}$.

Now let's think about the # ways to get 0 aces. That's just the way to choose 13 cards from the 48 non-aces in the deck, so $\binom{48}{13}$.

So overall, $\Pr[X=0]$ is $\dfrac{\binom{48}{13}}{\binom{52}{13}}$.

Now let's do $\Pr[X=1]$. Again, the denominator is $\binom{52}{13}$.

For the numerator, we know that we must choose one ace and 12 non-aces, so we have $\binom{4}{1}\binom{48}{12}$ choices — $\binom{4}{1}$ choices for the suit of the ace and $\binom{48}{12}$ choices for the remaining 12 cards.

The same logic applies for $X=2$, $X=3$, and $X=4$, so overall we have

$$E[X] = \sum_{i=0}^{4} i \cdot \Pr[X=i] =$$

$$0 \cdot \frac{\binom{4}{0}\binom{48}{13}}{\binom{52}{13}} + 1 \cdot \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}} + 2 \cdot \frac{\binom{4}{2}\binom{48}{11}}{\binom{52}{13}}$$

$$+ 3 \cdot \frac{\binom{4}{3}\binom{48}{10}}{\binom{52}{13}} + 4 \cdot \frac{\binom{4}{4}\binom{48}{10}}{\binom{52}{13}},$$

which does end up evaluating to 1.

Another example: what is the probability of drawing a full house?
A full house is 3 cards of one rank and 2 cards of another rank. For example,

2 hearts, 2 diamonds, J spades, J hearts, J clubs

is a full house.

So we need to compute $\dfrac{\text{\# ways to get full house}}{\text{\# ways to draw 5 cards}}$.

we know # ways to draw 5 cards is $\binom{52}{5}$ from above.

What is # ways to get a full house?

We can think of a full house as an element of the following set:

$$\{\text{all pairs of ranks}\} \times \{\text{all suit combos for a pair}\} \times \{\text{all triples of ranks, diff from pair}\} \times \{\text{all suit combos for triple}\}$$

By the product rule, the size of this set is the product of the sizes of these sets.

Or, equivalently, it's

$$\#\text{ways to choose 2 ranks} \cdot \#\text{ways to choose their suits} \cdot \#\text{ways to choose a diff. triple} \cdot \#\text{of ways to choose their suits.}$$

which is

$$\binom{13}{1} \cdot \binom{4}{2} \cdot \binom{12}{1} \cdot \binom{4}{3}.$$

So

$$\Pr[\text{full house}] = \frac{\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{3}}{\binom{52}{5}}$$