CSCI 3022

intro to data science with probability & statistics

posted tonight

September 5, 2018

A. HW 1 due next Fri. 9/14/18

Suggestion: do 2 problems by the end of this week.

1. Probability concepts and definitions.
2. Quintessential diagrams: circles and boxes
3. Quintessential problem: coins and flipping them.

Department of Computer Science
UNIVERSITY OF COLORADO BOULDER

Dan Larremore

Probability: why?

- Aspects of the world seem random and unpredictable.
    - Tall or short? Mom's eyes? Dad's chin? Is the eye of the hurricane going to pass over city x? Who survives the Titanic? How long will it take to drive to the airport? How long until the bus actually shows up?

Probability: why?

- Aspects of the world seem random and unpredictable.
    - Tall or short? Mom's eyes? Dad's chin? Is the eye of the hurricane going to pass over city x? Who survives the Titanic? How long will it take to drive to the airport? How long until the bus actually shows up?
- **Probability** is a way of thinking of these phenomena as if they were each generated by some **random process**.
- Turns out, using probability means **we can describe randomness with math**.

Basic Definitions

LaTeX ← Markdown

Jupyter Github

Definition: sample space Ω is the set of all the outcomes that we care about.

\Omega \Omega \omega \omega

Example: if we're flipping a coin, what is the sample space?

$\Omega = \{H, T\}$

$\Omega = \{1, 0\}$

$|\Omega| = 2$

Example: if we're polling to learn which month a person's birthday falls into, what is the sample space?

$\Omega = \{\text{Jan}, \text{Feb}, \text{Mar}, \ldots, \text{Dec}\}$

$|\Omega| = 12$

Basic Definitions

Definition: sample space Ω is the set of all the outcomes that we care about.

Example: if we're flipping a coin, what is the sample space?

Example: if we're polling to learn which month a person's birthday falls into, what is the sample space?

Observation: these are examples of *discrete* sample spaces because there is a countable number of outcomes.

Basic Definitions

Definition: for each outcome in Ω the probability is how likely that outcome is to occur. Probabilities are numbers in [0,1].

Observation: add up all the probabilities in the sample space and you *must* get 1. *Why?*

Something must happen. (Assumed)

Basic Definitions

Definition: an event is a set of outcomes. An event could be a single outcome or it could multiple outcomes. An event is a *subset* of the samples space Ω

"subset of"

Try it on: what is an example of an event for the coin flipping sample space? H $\Omega = \{H, T\}$ $H \subset \Omega$

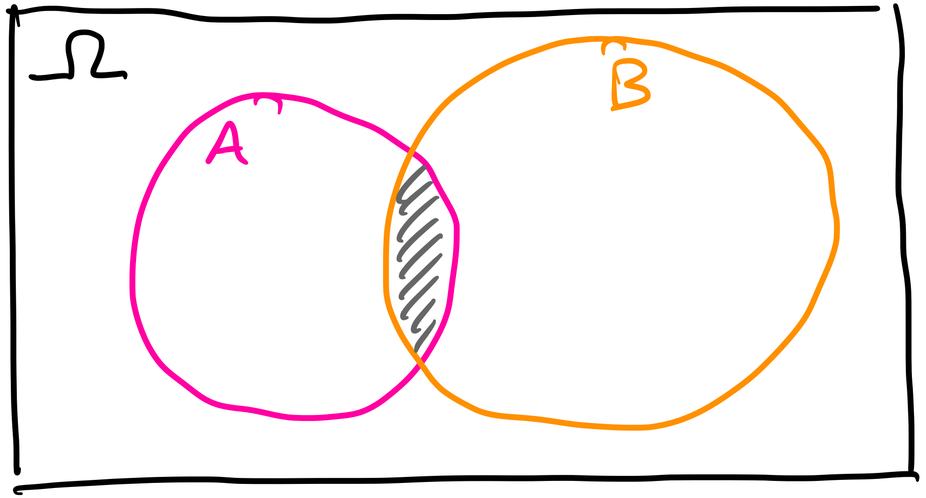
What is an example of an event for the birthday month sample space? March $\subset \Omega$ $\{Jan, June, July\} \subset \Omega$

$\{June, July, Aug\} \subset \Omega$

Definitions: set operations

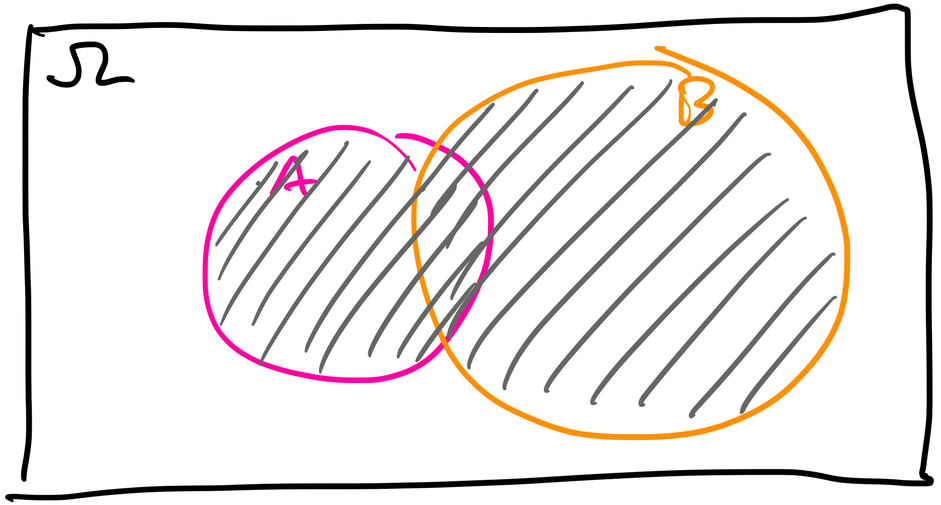
Definition: the intersection of two events is the subset of outcomes in **both** events.

intersection = “and"



Definition: the union of two events is the subset of outcomes in **one or both** events.

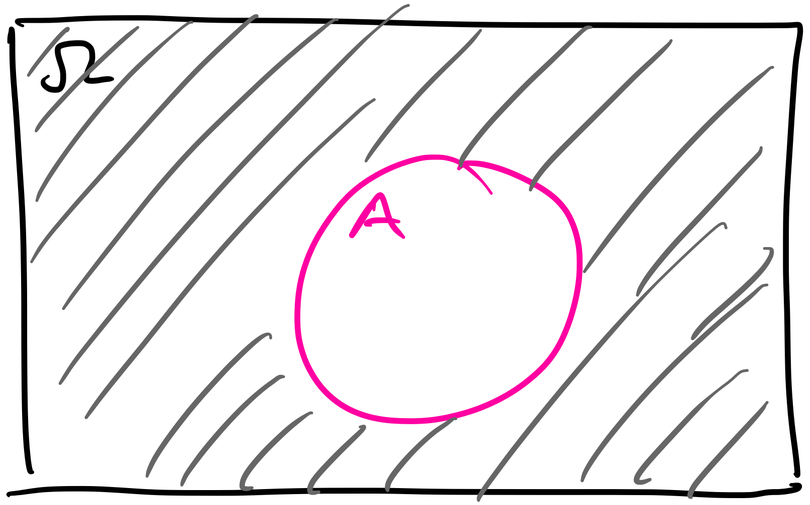
union = “or”



Definitions: set operations

not an i

Definition: the complement of an event A is the set of outcomes that are in Ω but **not in A**.



Notation time:

Complement: A^c A^C A^{C}

Intersection $A \cap B$ A \cap B

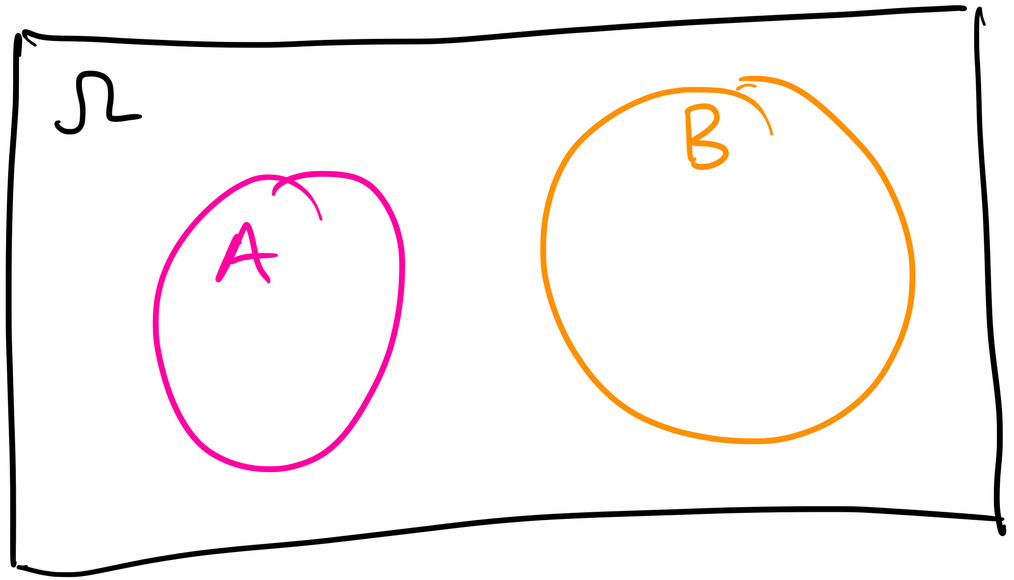
Union: $A \cup B$ A \cup B

Definitions: disjoint and null

Definition: when the intersection of two sets is empty, we call those two sets disjoint or mutually exclusive.The null set is the empty set.

$A \cap B = \emptyset$

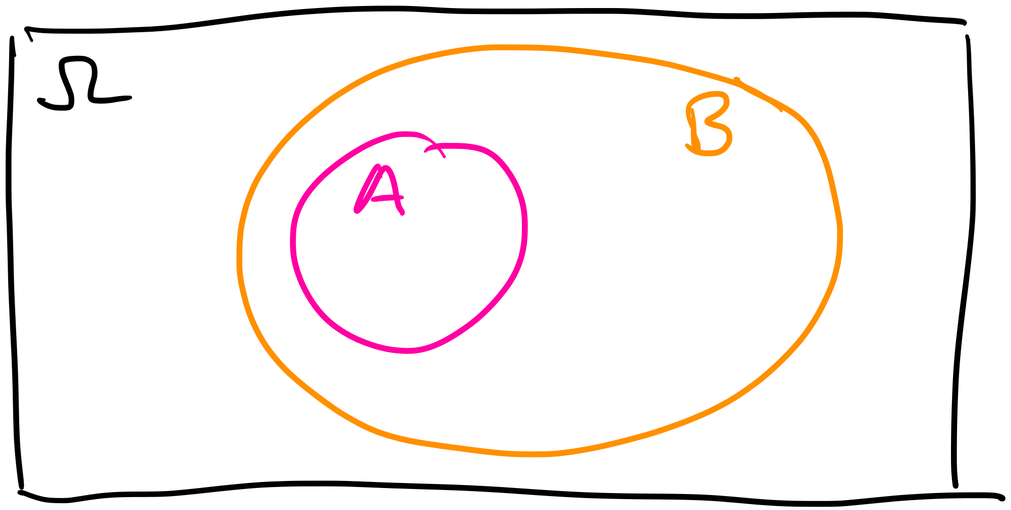
Null ?



Notation: null set $\emptyset$

Definitions: subset

Definition: if all of the outcomes in A are also outcomes in B, we say that A is a subset of B.



Notation: subset

$A \subset B$ "strict"

$A \subseteq B$ "inclusive"

$a < b$ vs. $a \leq b$

DeMorgan's Laws

$$(A \cup B)^C = A^C \cap B^C$$

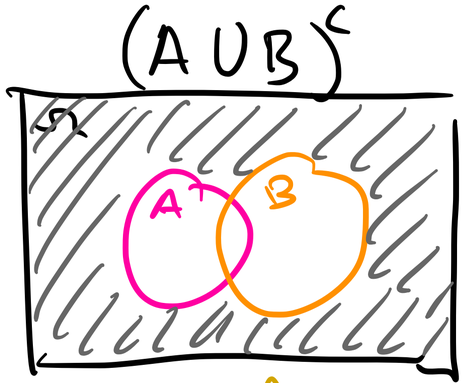
$$(A \cap B)^C = A^C \cup B^C$$

InClass: prove one of these using algebra or diagrams.

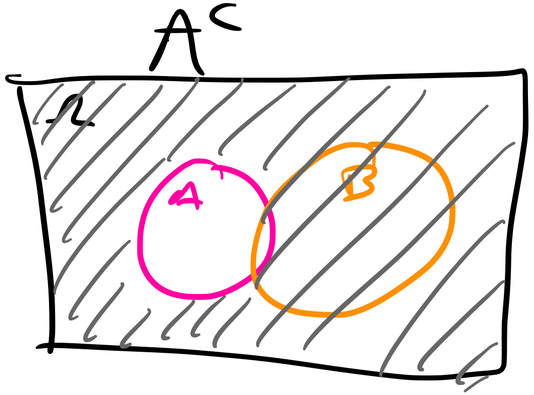
try at home w/ friends!

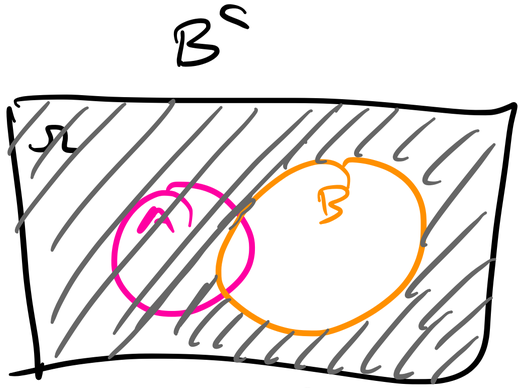
$(A \cup B)^c \neq A^c \cup B^c$

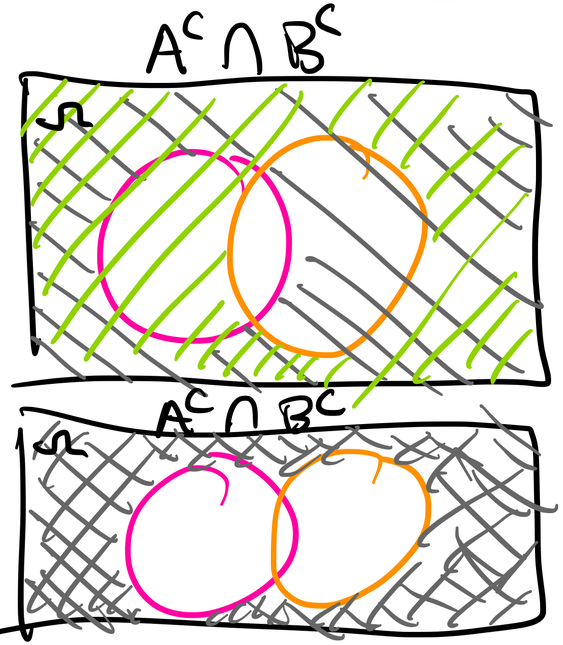
needs to flip!



=







Prob Functions: coins & birthdays

Basic Definition: a probability function associates each outcome in the sample space with a probability in [0,1].

← think, for now, "lookup table"

Try it out: what is the probability function for…flipping an unbiased coin?

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

…which month a birthday falls into?

$P(\text{Jan}) = \frac{1}{12}$ or $P(\text{Jan}) = \frac{31}{365}$

$P(E) = \frac{3}{12} = \frac{1}{4}$

Event = bday in a 3-month. $E = \{\text{Jan}, \text{June}, \text{July}\}$

Prob Functions: coins & birthdays

Basic Definition: a probability function associates each outcome in the sample space with a probability in [0,1].

Try it out: what is the probability function for…flipping an unbiased coin?

…which month a birthday falls into?

Advanced: what is the probability that a birthday falls into a month with 31 days?

← at home!

show $\frac{7 \cdot 31}{365}$

Prob Functions: the biased coin

- The biased coin is a coin with an altered probability function.
- Instead of P({H,T})={1/2, 1/2}, a biased coin's probability function is P({H,T})={p,q}. *What can we say about q?*

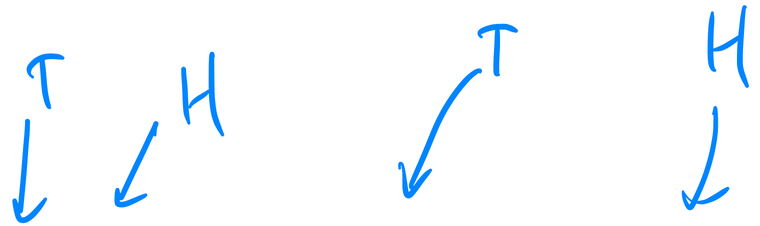
① $q = 1-p$ bc all the outcomes in Ω must add up to 1 (the probabilities of)

② "biased" = not "fair" $\Rightarrow p \neq 1-p$, i.e, $p \neq \frac{1}{2}$

true IRL, but in this class "biased" means $p \in [0,1]$ (λin)

Prob Functions: the biased coin

- The biased coin is a coin with an altered probability function.
- Instead of P({H,T})={1/2, 1/2}, a biased coin's probability function is P({H,T})={p,q}. *What can we say about q?*



- Rather than {H,T}, we can think of {0,1} or {failure,success}
- Each flip is sometimes called a *Bernoulli Trial*

sin/cos: Know your Bernoullis

We don't *really* have time for the cool history here, but…

https://en.wikipedia.org/wiki/Bernoulli_family

Brothers: Jacob, Nicolaus, **Johann**
Others: Elisabeth, Daniel, Hans…

https://en.wikipedia.org/wiki/List_of_things_named_after_members_of_the_Bernoulli_family

Probability functions

- Notice that the probability functions have two key aspects:
  - The probability of the whole sample space is 1.
  - The probability of two disjoint events is the sum of their individual probabilities.

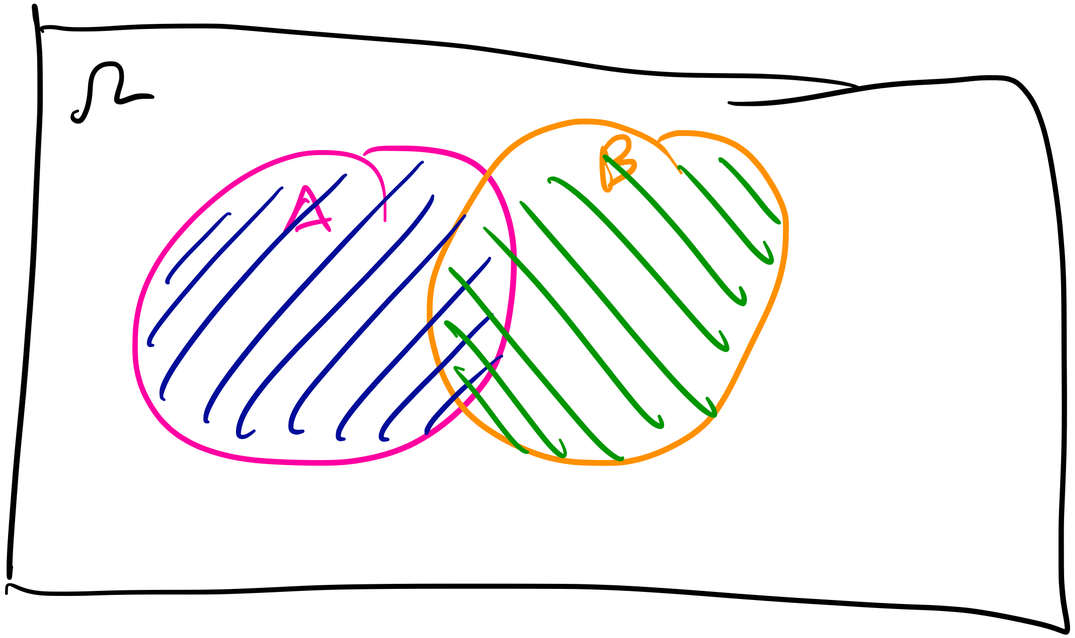
Formal Definition: a *probability function* P assigns to each event A a number P(A) in [0,1] such that:

e.g. A = Jan. B = March

1. $P(\Omega) = 1$
2. $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint.

What is the probability of two non-disjoint events?

Let's work this out:



$$P(A) + P(B) - P(A \cap B)$$

$$= P(A \cup B)$$

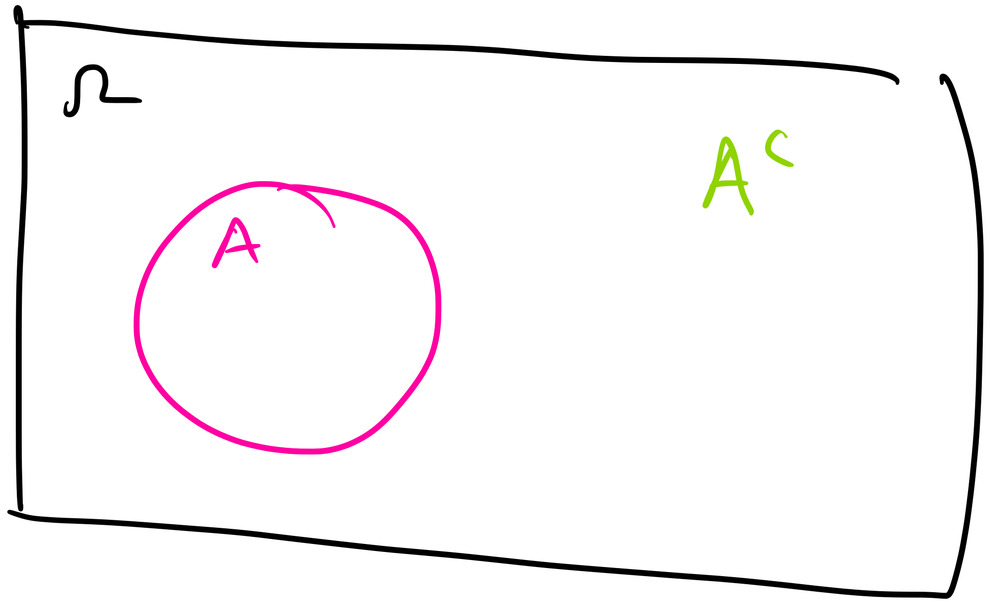
adjust for double counting

What is the probability of the complement of an event?

Let's work this out:

$$A \cup A^c = \Omega$$

$$A \cap A^c = \emptyset \text{ (disjoint)}$$



$$P(A \sim A^c) = P(A) + P(A^c)$$

$$= P(\Omega)$$

$$\Rightarrow P(A) + P(A^c) = 1$$

$$\Rightarrow \boxed{P(A^c) = 1 - P(A)}$$

Suppose I flip a biased coin over and over…

What is the probability that I flip twice and both are heads?

The sample space for a single coin flip is $\Omega = \{H,T\}$.

The sample space for two coin flips is $\Omega = \{H,T\} \times \{H,T\}$.

Suppose I flip a biased coin over and over…

What is the probability that I flip twice and both are heads?

The sample space for a single coin flip is $\Omega = \{H,T\}$.

The sample space for two coin flips is $\Omega = \{H,T\} \times \{H,T\}$.

In other words: Ω is all the possibilities of Ω_1 crossed with all the possibilities of Ω_2.

In general, if $|\Omega_1|=n$ and $|\Omega_2|=m$, then $|\Omega_1 \times \Omega_2| = nm$.

Suppose I flip a biased coin over and over…

What is the probability that I flip twice and both are heads?

The sample space for a single coin flip is $\Omega = \{H,T\}$.

The sample space for two coin flips is $\Omega = \{H,T\} \times \{H,T\}$.

In other words: Ω is all the possibilities of Ω_1 crossed with all the possibilities of Ω_2.

In general, if $|\Omega_1|=n$ and $|\Omega_2|=m$, then $|\Omega_1 \times \Omega_2| = nm$.

This is an example of a *product* of sample spaces.

Suppose I flip a biased coin over and over…

What is the probability that I flip twice and both are heads?

The sample space for two coin flips is $\Omega = \{H,T\} \times \{H,T\}$.

Intuition: does the second flip's outcome depend on the first? NO

Definition: When the outcomes of two trials do not depend on each other, we say they are *independent*.

Suppose I flip a biased coin over and over…

What is the probability that I flip twice and both are heads?

The sample space for two coin flips is $\Omega = \{H,T\} \times \{H,T\}$.

Useful fact: When the result of one trial is *independent* of the other we can **multiply their probabilities**.

For independent events, AND means multiplication:

P(H *and* H) = $P(H) \cdot P(H)$

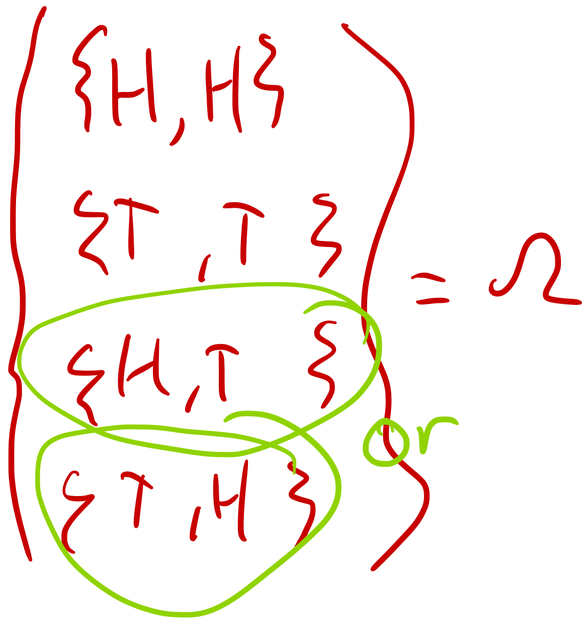
$= p \cdot p$

$= p^2$

Suppose I flip a biased coin…

What is the probability that I flip twice and end up with one heads and one tails?

Useful fact: For independent outcomes, OR means addition:



$= \Omega$

$$P(\{T,H\}) + P(\{H,T\})$$

$$(1-p)\,p \quad + \quad p\,(1-p)$$

$$\boxed{2p(1-p)}$$

Suppose I flip a biased coin over and over…

What is the probability that I flip 5 coins and get exactly one heads?

Event, E

$\Omega = \{H, T\}^5$

$E = \{$ HTTTT, THTTT, TTHTT, TTTHT, TTTTH $\}$

$p(1-p)^4$ $\quad$ $(1-p)p(1-p)^3 = p(1-p)^4$ $\quad$ $p(1-p)^4$ $\quad$ $p(1-p)^4$ $\quad$ $p(1-p)^4$

$P(E) = P(\text{I get one heads in my 5 flips}) = 5p(1-p)^4$

Let's get empirical

I have a coin, but I don't know the probabilities for heads or tails. What could I do to find out?

Flip coin n times

Count # of heads.

empirical estimate $\hat{p} = \frac{\#\text{ heads}}{n}$

- Max. Likelihood estimate.
  - unbiased
  - consistent