topic 11: combinatorics & probability (a crash course)

stuff to learn today:

1. what is probability and why
2. sets
3. permutations & factorials
4. combinations
5. conditional probability
6. law of total probability

probability: what is it, and why?

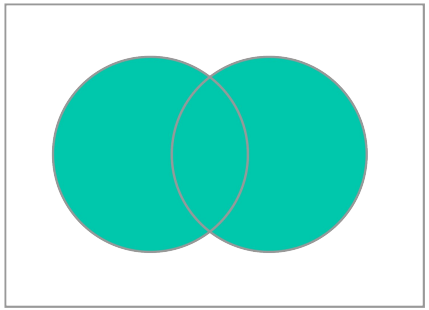
- the probability of something occurring is its proportion of all possible outcomes
- {event space} / {sample space}
- probability of flipping heads, rolling a 5, etc.

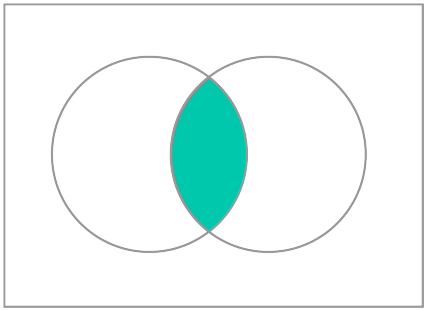
why?

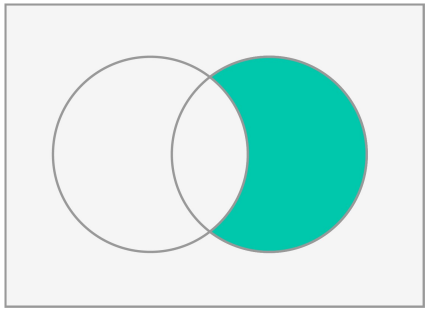
- statistics
- data science interviews!

set theory (terminology)

union, intersection, complement







terminology

- sample space: all possible outcomes
- event space: a subset of the sample space, based on the probabilities we want to calculate

the **law of relative frequency**

- the probability of an event is the ratio of positive trials to the total number of trials as we do an infinite number of trials

probability axioms

1. **positivity**: a probability of an event is always $0 \leq P(E) \leq 1$

2. **a certain event**: an event that is equivalent to the sample space, and P(Certain Event) = 1

3. **additivity**: P(the union of two exclusive events) == P(Event 1) + P(Event 2) **only if** P(intersection of 1 and 2) == 0

addition law of probability: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

factorials!

- a factorial is denoted with **!**
- the product of all positive integers leading up to the number
- n! == 1 x 2 x 3 x ... x (n-1) x n

probability use case:

- # of ways to arrange **n** objects
- 3! ways to arrange A, B, C: ABC, ACB, BAC, BCA, CAB, CBA

permutations

permutations and combinations are **counting methods** that help us find the size of our probability spaces.

permutation: the number of ways to arrange **k** out of **n** items

- formula: $n! \div (n-k)!$
- i.e. a band has a catalog of 8 songs but has to play a set of 3 songs. how many different 3-song sets are there? (a different order of the same songs counts as a different set)

- $8! \div 5! = 336 = 8 \times 7 \times 6$

combinations

combination: the number of ways to choose **k** out of **n** items, and the “order” of your items doesn’t matter

- formula: n! ÷ ((n-k)! x (k!))
- i.e. if I have a collection of 8 seashells and want to randomly pick 3 to give away, how many combinations of 3 shells are there?

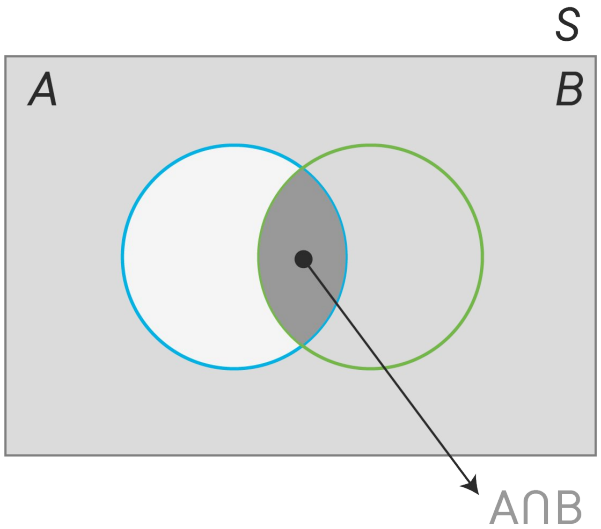
- 8! ÷ (5! x 3!) = 56

conditional probability

conditional probability is the probability of something occurring **given that something has already happened** - an important concept for Bayesian stats

P(A|B) = P(A and B)/P(B)

we come back to this in Section 17!



law of total probability

the probability of B can be calculated if we know the individual conditional probabilities: P(B|A1), P(B|A2) and P(B|A3)

the law of total probability says that P(B) is the sum of the individual conditional probabilities, given that each A segment is disjoint and there is no part of B that isn't part of an A segment