Introduction to Probability

Alessia Mondolo

July 19, 2019

Academy AI

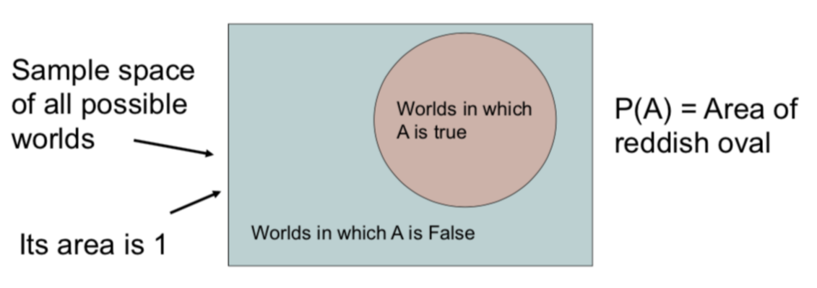
Basic concepts

Some definitions

- Sample space S: the set of possible wolds (e.g. set of students in our class, set of cards in a deck)
- Random variable A: a function defined over S
  (e.g. Gender: S $\rightarrow$ {m, f})
- Event: a subset of S (e.g. the subset of S for which Gender=f, the subset of S for which (Gender=m) AND (eyeColor=blue))
- Probability P(A): the likelyhood or chance of an event occurring, informally "the fraction of possible worlds in which A is true"

We are often interested in probabilities of specific events and of specific events conditioned on other specific events.

Visualizing A

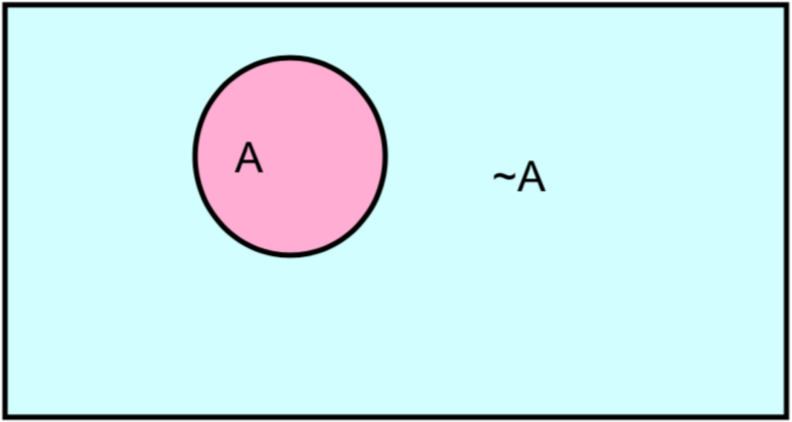


Axioms of probability

- $0 \leq P(A) \leq 1$
- $P(True) = 1$
- $P(False) = 0$
- $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

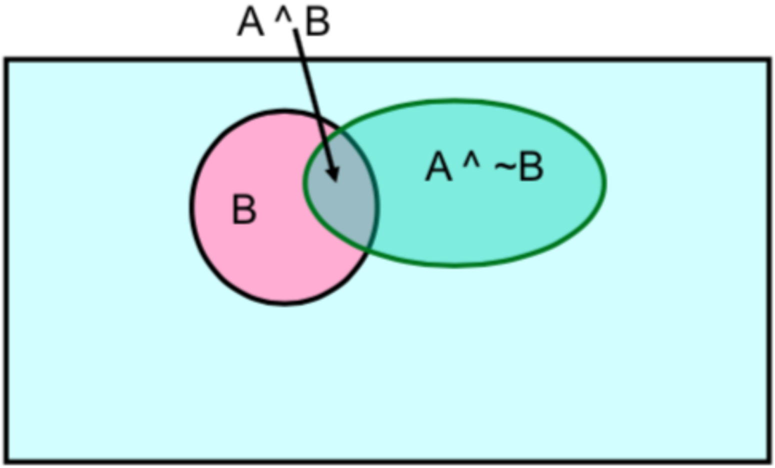
Elementary probability in pictures

$$P(\neg A) + P(A) = 1$$



Elementary probability in pictures

$$P(A) = P(A \wedge B) + P(A \wedge \neg B)$$

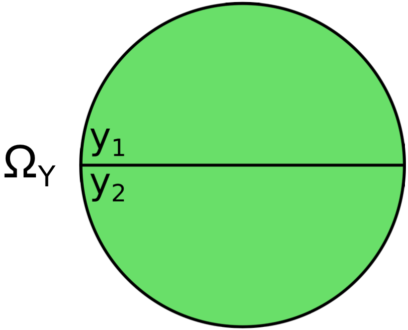


Marginal probability

The marginal probability is the probability of an event occurring (p(A)), it may be thought of as an unconditional probability. It is not conditioned on another event.

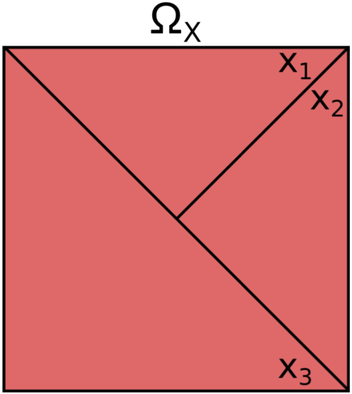
Example: the probability that a card drawn is red (p(red) = 0.5).

Another example: the probability that a card drawn is a 4 (p(four)=1/13).



$$p(Y = y_1) = 1/2$$
$$p(Y = y_2) = 1/2$$

Marginal probability



$$p(X = x_1) = 1/4$$
$$p(X = x_2) = 1/4$$
$$p(X = x_3) = 1/2$$

$$\left(\sum_{x \in \Omega_X} p(X = x) \right) = 1$$

Joint probability

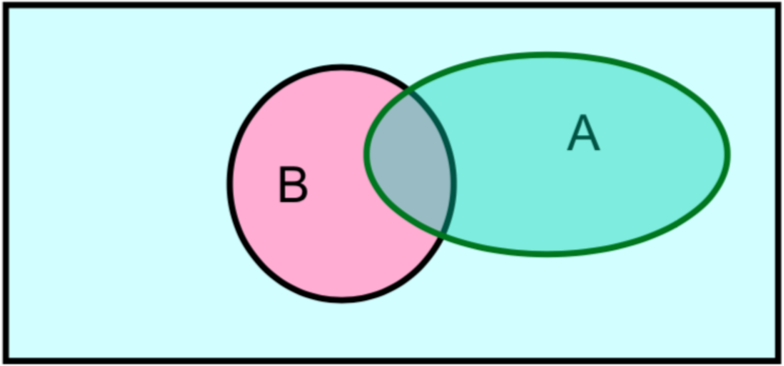
The joint probability p(A and B) is the probability of event A and event B occurring. It is the probability of the intersection of two or more events. The probability of the intersection of A and B may be written p(A $\cap$ B).
Example: the probability that a card is a four and red $= p(fourandred) = 2/52 = 1/26$.
(There are two red fours in a deck of 52, the 4 of hearts and the 4 of diamonds).

Conditional probability

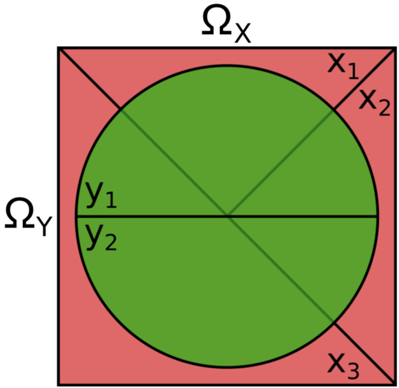
The conditional probability $p(A \mid B)$ is the probability of event A occurring, given that event B occurs.
Example: given that you drew a red card, what's the probability that it's a four ($p(\mathit{four} \mid \mathit{red}) = 2/26 = 1/13$). So out of the 26 red cards (given a red card), there are two fours so $2/26 = 1/13$.

Conditional probability

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

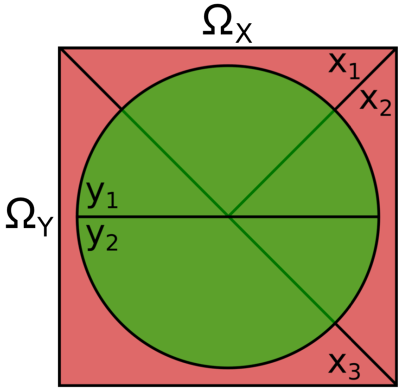


Conditional probability



$$p(Y = y_1 | X = x_1) =$$
$$p(Y = y_2 | X = x_1) =$$
$$p(Y = y_1 | X = x_2) =$$
$$p(Y = y_2 | X = x_2) =$$
$$p(Y = y_1 | X = x_3) =$$
$$p(Y = y_2 | X = x_3) =$$

Conditional probability



$$p(Y = y_1|X = x_1) = 1$$
$$p(Y = y_2|X = x_1) = 0$$
$$p(Y = y_1|X = x_2) = 1/2$$
$$p(Y = y_2|X = x_2) = 1/2$$
$$p(Y = y_1|X = x_3) = 1/4$$
$$p(Y = y_2|X = x_3) = 3/4$$

Chain rule

The chain rule for two random events A and B says:

$$P(A \wedge B) = P(A \mid B) \cdot P(B)$$

Chain rule

For all $\boldsymbol{x}$ we have that

$$p(\boldsymbol{x}) = p(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} p(x_i | x_1, ..., x_{i-1})$$

** it holds for any ordering of $X_1, ..., X_n$ **

Example: $p(x_1, x_2, x_3, x_4) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1, x_2)p(x_4 \mid x_1, x_2, x_3)$

Exercise

Chain rule

Let be $\boldsymbol{X} = (X_1, X_2, X_3, X_4)$, prove that for all $\boldsymbol{x}$ we have that $p(x_1, x_2, x_3, x_4) = p(x_1)p(x_2|x_1)p(x_3|x_1, x_2)p(x_4|x_1, x_2, x_3)$

**Remember that $p(x_i \mid x_1, ..., x_{i-1}) = \frac{p(x_1, \ldots, x_i)}{p(x_1, \ldots, x_{i-1})}$

Bayes rule

Bayes theorem describes the **probability of an event based on prior knowledge** of conditions that might be related to the event. For example, if cancer is related to age, then, using Bayes' theorem, a person's age can be used to more accurately assess the probability that they have cancer, compared to the assessment of the probability of cancer made without knowledge of the person's age.

Bayes rule:

$$P(A \mid B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

We call $P(A)$ the "prior" and $P(A \mid B)$ the "posterior".

Bayes rule

This is also Bayes rule:

$$P(A \mid B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A)}$$

$$P(A \mid B \wedge C) = \frac{P(B|A \wedge C) \cdot P(A \wedge C)}{P(B \wedge C)}$$

Exercise

$$P(A \mid B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A)}$$

A: you have a flu, B: you just coughed

Assume:

- $P(A) = 0.05$
- $P(B) = 0.8$
- $P(B \mid \neg A) = 0.2$

What is the probability that you have a flu given that you just coughed?

Independent events

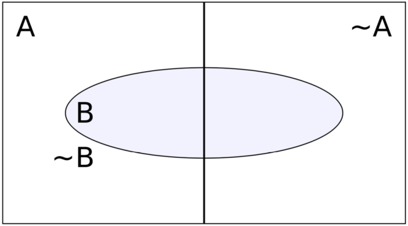
Definition

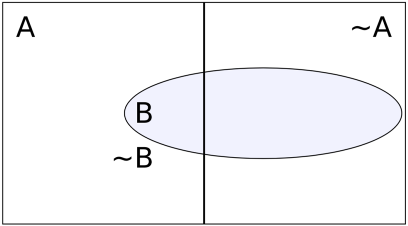
Two events A and B are independent if

$$P(A \wedge B) = P(A) * P(B)$$

** Intuition **

Knowing A tells us nothing about the value of B (and vice versa)





Permutations and combinations

Permutations

Permutations are the number of ways a subset of a specified size can be arranged from a given set, generally without replacement.
An example of this would be a 4 digit PIN with no repeated digits. The probability of having no repeated digits can be calculated by executing the following calculation:

10 x 9 x 8 x 7

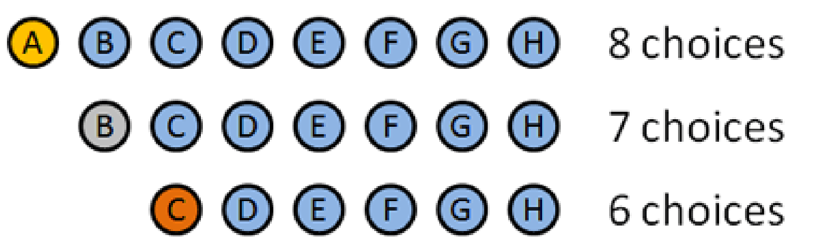
Permutations

Let's say we have 8 people:

1. Alice
2. Bob
3. Charlie
4. David
5. Eve
6. Frank
7. George
8. Horatio

Permutations

How many ways can we award a 1st, 2nd and 3rd place prize among eight contestants? (Gold / Silver / Bronze)



We're going to use permutations since **the order we hand out these medals matters.**

Permutations

Here's how it breaks down:

- Gold medal: 8 choices: A B C D E F G H. Let's say A wins the Gold.
- Silver medal: 7 choices: B C D E F G H. Let's say B wins the silver.
- Bronze medal: 6 choices: C D E F G H. Let's say... C wins the bronze.

We picked certain people to win, but the details don't matter: we had 8 choices at first, then 7, then 6.
The total number of options was $8 \cdot 7 \cdot 6 = 336$.

Permutations

Let's look at the details. We had to order 3 people out of 8. To do this, we started with all options (8) then took them away one at a time (7, then 6) until we ran out of medals.

We know the factorial is: $8! = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

Unfortunately, that does too much! We only want $8 \cdot 7 \cdot 6$. How can we "stop" the factorial at 5?

Notice how we want to get rid of $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$, which is 5 factorial!

So, if we do $8!/5!$ we get:
$\frac{8!}{5!} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 8 \cdot 7 \cdot 6$

Permutations

Why did we use the number 5? Because it was left over after we picked 3 medals from 8. So, a better way to write this would be:

$$\frac{8!}{(8-3)!}$$

where 8!/(8-3)! is just a fancy way of saying "Use the first 3 numbers of 8!". If we have n items total and want to pick k in a certain order, we get:

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

So, if You have **n** items and want to find the number of ways **k** items can be ordered:

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

Combinations

In combinations **order doesn't matter.**

Let's say I can't afford separate Gold, Silver and Bronze medals. In fact, I can only afford empty tin cans.

How many ways can I give 3 tin cans to 8 people? In this case, the order we pick people doesn't matter. If I give a can to Alice, Bob and then Charlie, it's the same as giving to Charlie, Alice and then Bob.

This raises an interesting point — we've got some redundancies here. Alice Bob Charlie = Charlie Bob Alice. For a moment, let's just figure out how many ways we can rearrange 3 people.

We have 3 choices for the first person, 2 for the second, and only 1 for the last. So we have $3 \cdot 2 \cdot 1$ ways to re-arrange 3 people.

Combinations

If we have 3 tin cans to give away, there are 3! or 6 variations for every choice we pick. If we want to figure out how many combinations we have, we just c**reate all the permutations and divide by all the redundancies**. In our case, we get 336 permutations (from above), and we divide by the 6 redundancies for each permutation and get $336/6 = 56$.

The general formula is

$$C(n,k) = \frac{P(n,k)}{k!}$$

Combinations

So, the combination formula, or the number of ways to combine k items from a set of n is:

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

Sometimes C(n,k) is written as:

$$\binom{n}{k}$$

which is the the binomial coefficient.