// Conditional Probability

Data Science Immersive

August 07, 2019

FLATIRON SCHOOL

conditional probability

- today's knowledge:

    a. Independent and Dependent Events
    b. Partitioning + Conditional Probability
    c. Bayes' Theorem

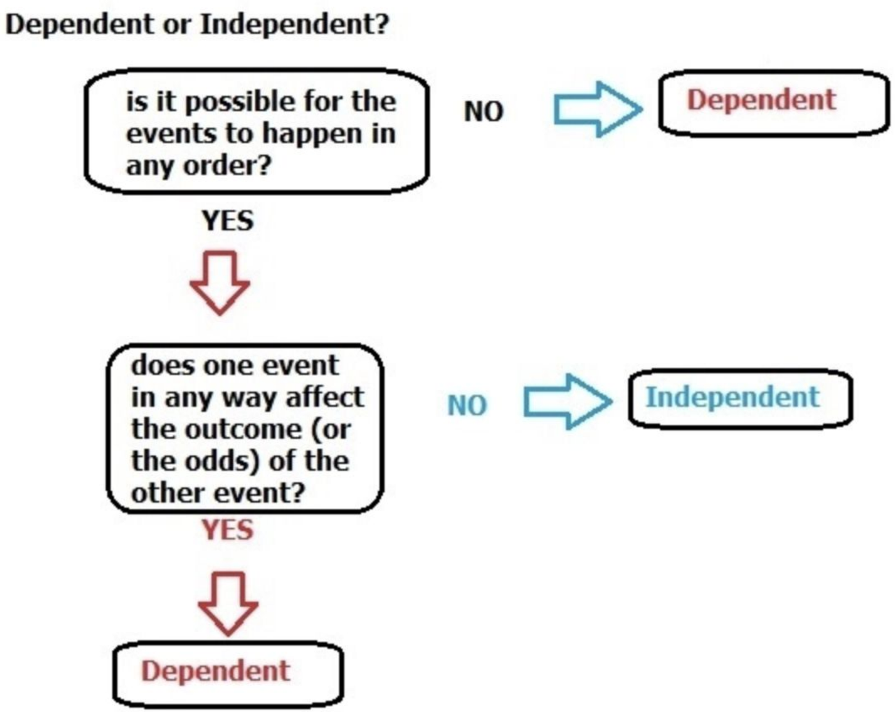
independent/dependent events

1. From a standard deck of cards:
   a. Drawing a face card
   b. Drawing a red card

2. With a standard 6-sided die:
   a. Rolling a 3
   b. Rolling a 4 or a 6

3. On any given day:
   a. Chance of rain
   b. Chance of needing coffee

4. Raining makes me tired - so rain leads to more coffee needed:
   a. Chance of rain
   b. Chance of needing coffee



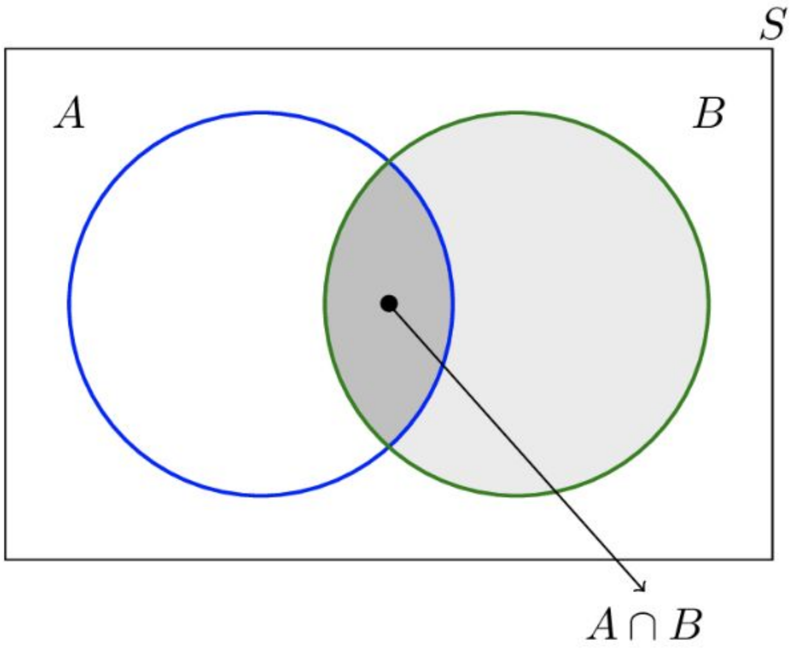
conditional probability

- conditional probability emerges in the examination of experiments **where a result of a trial may influence the results of the upcoming trials**

- the probability of drawing an Ace given that you already drew an Ace

- the probability of being in a good mood given that the weather is nice

- conceptually: you are partitioning your sample space!



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

conditional probability

You have a bag of 2 blue marbles and 3 red marbles.

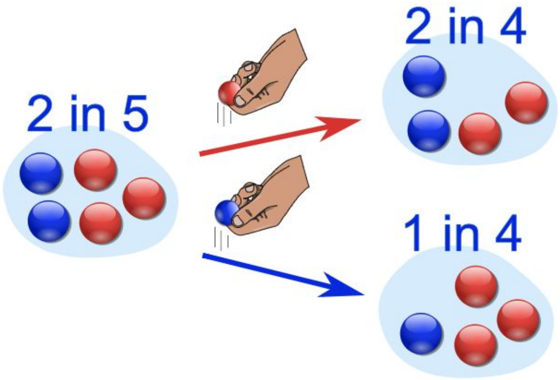
You draw two marbles without replacement.

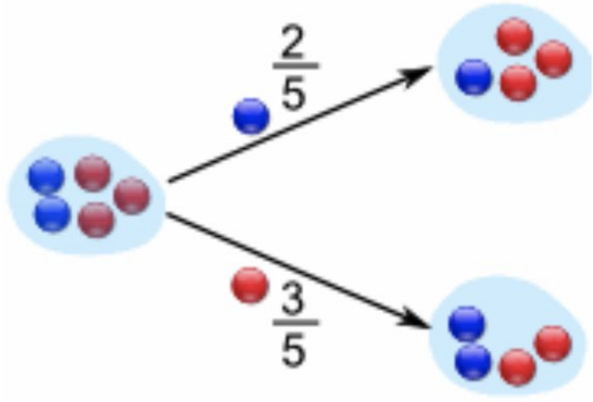
What is the probability your second marble is red, given that your first marble is blue?

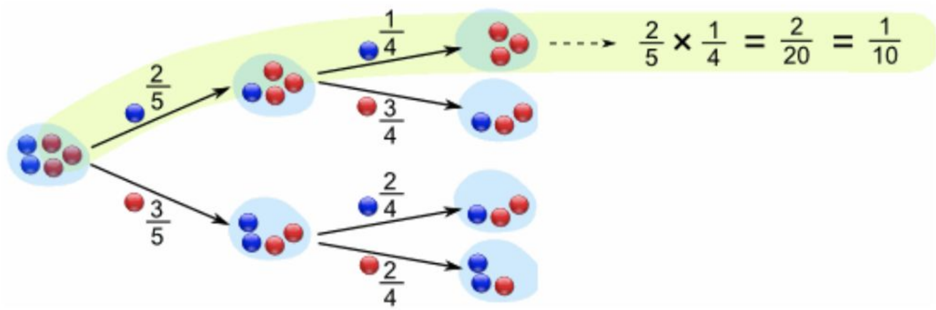
What is the probability your second marble is red, given that your first marble is red?

What is the probability your second marble is red?

conditional probability: in trees







example 1: hungry dogs

imagine a world made up of 50% cats and 50% dogs. these cats and dogs can only be hungry or not hungry.

we randomly pick one of these animals and the probability of having a hungry dog is 10% ☹️

what is the probability of picking a hungry animal given that you picked a dog?

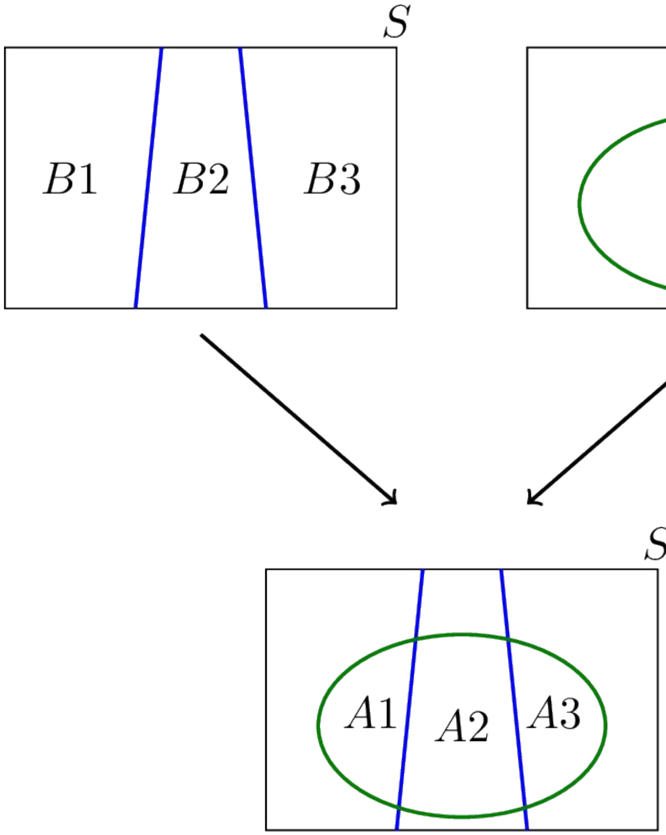
example 2: know your nieces and nephews

You are going to visit your distant cousins who recently had two children. You are told that **at least one of them is a girl**. What is the probability of both of them being girls?

example 3: seriously... know your family.

You are going to visit your distant cousins who recently had two children. You are told that the **older** one is a girl. What is the probability of both of them being girls?

Law of Total Probability



Law of
Total Probability

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)\, P(B_i)$$

Bayes' Theorem

- Bayes' Theorem (shoutout Thomas Bayes) is derived using conditional probability and is the basis of **Bayesian Inference** - an idea we'll explore way more when we talk about classification models

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

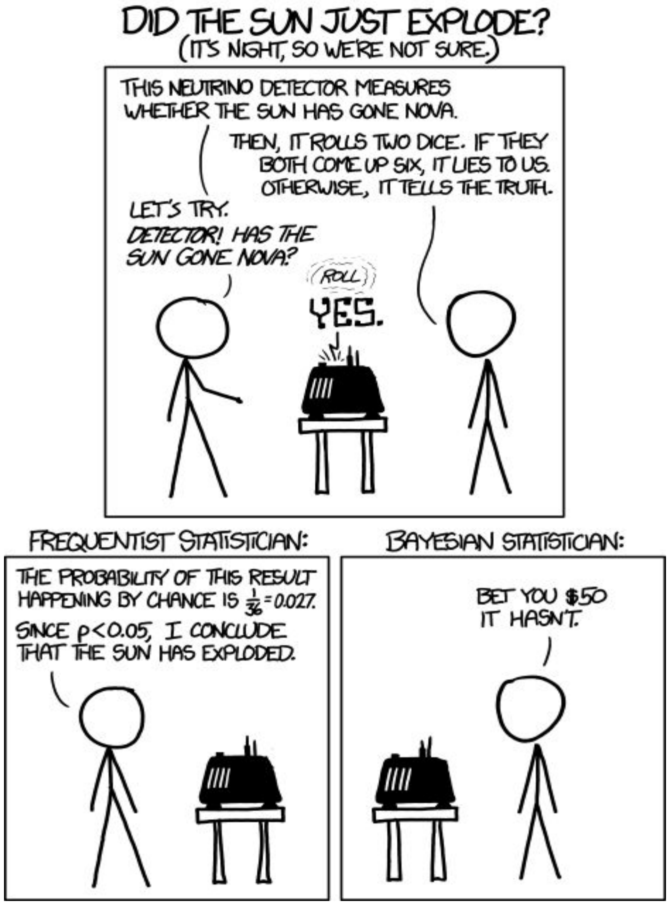
$$P(A \cap B) = P(A|B)\,P(B) = P(B|A)\,P(A)$$

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

Bayesians vs Frequentists

two main probabilistic frameworks!

- frequentists: based on **frequency** of **observed** events
- bayesians: using some **prior** beliefs/knowledge



the Bayes-ic concepts

P(A) is called the **prior**; this is the probability of our hypothesis without any additional prior information. It could also be a belief we have prior to seeing the data.

P(B) is called the **marginal likelihood**; this is the total probability of observing the evidence.

P(B|A) is called the **likelihood**; this is the probability of observing the new evidence, given our initial hypothesis.

P(A|B) is called the **posterior**; this is what we are trying to estimate.

checking for understanding

No One:

Bayesian Network:

Bayes-ically...

If $P(A) = \frac{1}{2}$ and $P(B) = \frac{1}{2}$ and $P(B|A) = 1/3$, find:

a. $P(A \text{ and } B)$

b. $P(A \text{ or } B)$

c. $P(A|B)$

totally diseased ☹️

A medical test detects a disease with P = 0.99 and fails to detect it with P = 0.01.

If there is no disease, the test indicates there is a disease with P = 0.02 and indicates there isn't with P = 0.98

The test is done on a random person from a population where the occurrence of the disease is 1/10000.

Find the probability of this person having the disease if the test comes back positive.

Now, a person has been referred by a doctor to take this test and there's a prior probability that he has P = 0.3 of having this disease ☹️