# Bayesian Statistics

**We'll cover the following**



## What Is Bayesian Statistics?

What we have learned so far about probability falls into the category of **Frequency Statistics**. But there is another more powerful form of statistics as well, and it called **Bayesian Statistics**, sometimes called, *Bayesian Inference*. Bayesian Statistics is a more general approach to statistics; it describes the probability of an event based on the previous knowledge of the conditions that might be related to the event. It allows us to answer questions like:

- *Has this happened before?*
- *Is it likely, based on my knowledge of the situation, that it will happen?*

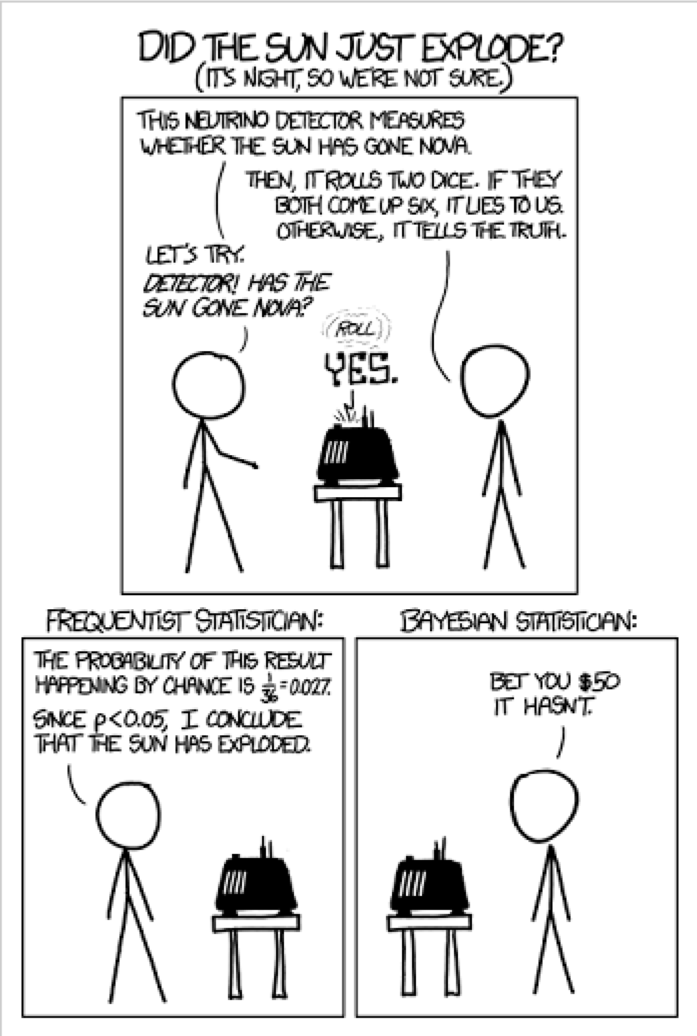


Image Credits: https://xkcd.com/

Let's look at an example. *Ever wonder how a **spam filter** could be designed?*

Say an email containing, *"You won the lottery"* gets marked as spam. The question is, how can a computer understand that emails containing certain words are likely to be spam? ***Bayesian Statistics does the magic here!***

Spam filtering based on a blacklist would be too restrictive and it would have a high false-positive rate, spam that goes undetected. Bayesian filtering can help by allowing the spam filter to learn from previous instances of spam. As we analyze the words in a message, we can compute its probability of being spam using Bayes' Theorem. And as the filter gets trained with more and more messages, it updates the probabilities that certain words lead to spam messages. Bayesian Statistics takes into account previous evidence.

Bayesian Statistics is based on the **Bayes' Theorem**: This is ***basically a way of finding a probability when we know certain other probabilities***. The magical Bayes' formula looks like this:

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

This tells us how often *A* happens given that *B* happens, written *P(A|B)*, when we have the following information:

- How often *B* happens given that *A* happens, written *P(B|A)*.
- How likely *A* is on its own, written *P(A)*.
- And how likely *B* is on its own, written *P(B)*

If we map this to the spam filter example, we get:

$$P(spam|words) = \frac{P(spam)P(words|spam)}{P(words)}$$

### Bayes' Theorem Example: #

Say we have a clinic that tests for allergies and we want to find out a patient's probability of having an allergy. We know that our test is not always right:

- For people that really do have the allergy, the test says "Positive" 80% of the time, *(Positive|Allergy)*
- The probability of the test saying "Positive" to anyone is 10%, *P(Positive)*

If 1% of the population actually has the allergy, *P(Allergy)*, and a patient's test says "Positive", what is the chance that the patient really does have the allergy, i.e., *P(Allergy|Positive)*?

Bayes' theorem tells us:

$$P(Allergy|Positive) = \frac{P(Allergy)P(Positive|Allergy)}{P(Positive)} = \frac{0.01 * 0.8}{10} = 0.08$$

This means the chance that the patient actually has the allergy is only *8%*! *Sound like a clinic we should stay away from!*