Bayesian and frequentist statistics

Bayes' theorem:

$$P(M|D) = \frac{P(M) \cdot P(D|M)}{P(D)}$$

this just follows from the laws of probability!

↳ quite powerful still!

M: our model

D: the data

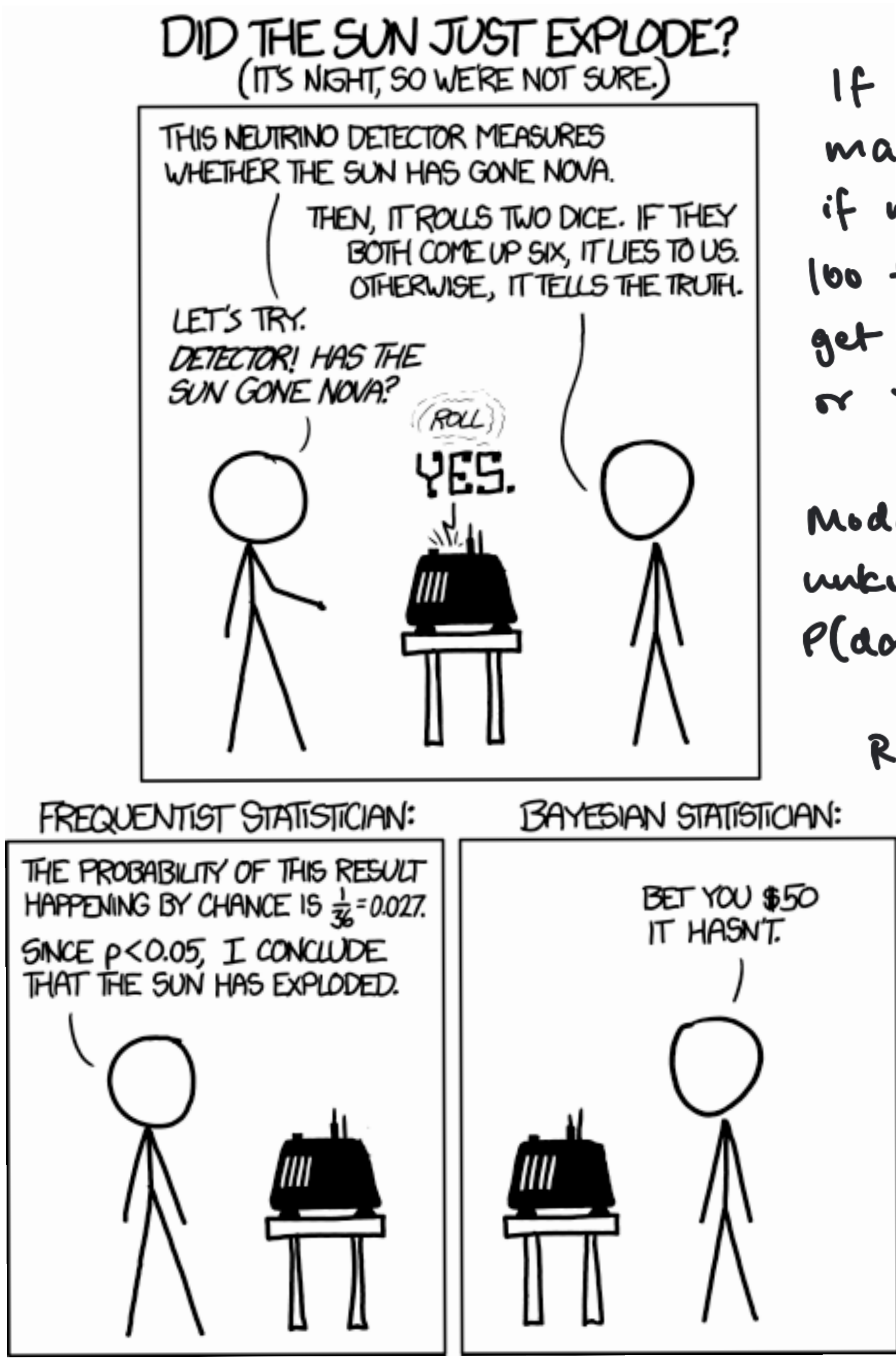
$P(M)$: the "prior" probability of M

$P(D|M)$: the likelihood

$P(D)$: a normalizing factor

↳ whether D would have happened anyway

⇒ can only calculate this if you have other hypotheses for comparison.



Frequentist

If the true value of the mass is 130 meV, then if we repeat the experiment 100 times, we would twice get a measurement <120 or > 140.

Model params. are fixed + unknown. We calculate P(data) given those params.

Random variables model outcome of data.

Often somewhat ad hoc: p-values, etc.

Bayesian

The probability of the mass being between 120 – 140 meV is 98%..

Consider data known / fixed. calculates probabilities of models / hypotheses / parameters.

Requires prior. estimation of the model's likelihood.

↓

incorporates prior knowledge.

Used increasingly bc big data / machine learning / etc ... can give POWERFUL INSIGHTS!

Coin flip

p-value for the coin. $\quad$ Prob(heads) = ?

you flip the coin 20 times and get 15 heads.

↳ frequentist approach:

a maximum likelihood calculation gives p(heads) = $3/4$.

unfair coin!

a Bayesian statistician would be a little uneasy.

↳ param. p = Prob(heads)

let's assume a uniform prior for p $\longrightarrow$ $P(m) = P(p) = 1$.

likelihood: $P(D|p) = \frac{20!}{5!\,15!} p^{15} (1-p)^5$ $\quad \leftarrow$ Binomial!

$P(D)$: probability of observing D over all hypotheses i.e all values of p

↳ $\int_0^1 dp \; P(p) \frac{20!}{5!\,15!} p^{15} (1-p)^5$

Putting it all together (uniform prior) . . .

$$P(p|D) \propto P(p) \cdot P(D|p) = \frac{20!}{5!\,15!} (1-p)^5 p^{15}$$

$\Longrightarrow$ dropped $P(D)$ because it integrates out to a constant.

what does this look like?? $\longrightarrow$ JUPYTER NOTEBOOK!

NB: this gives us <u>distributions</u> for our parameters.
no confidence intervals etc needed here.

A new prior: unfair coins are very rare.

You're 99% sure this is a normal coin.

↳ might have some spread around $p = 0.5$, let's take $\sigma = 0.01$.

1% chance of an unfair coin $\longrightarrow$ who knows what p is!

$$\Longrightarrow \quad P(p) = 0.99 \cdot \frac{1}{\sqrt{2\pi} \cdot 0.01} \exp\left(-\frac{1}{2} \cdot \frac{(p-0.5)^2}{\sigma^2}\right) + 0.01 \cdot 1$$

So this all depends on your subjective choice of prior ??

↳ Yep.

↳ CLEARLY answers: how likely is your hypothesis?

This is a good thing! It allows us to quantitatively build in things that we know / suspect.

↳ If the prior changes the answer by a LOT?

Well, then the data were not very constraining.

also: nuisance param, Occam factor.

Frequentist analysis doesn't require you to spell out assumptions so clearly.

↳ uncertainty on p would come from var(Binom) but couldn't talk about prob of a value of p!

↳ only p(data) at some chosen p. confusing...

doesn't give us a distribution for p either.