USAID
FROM THE AMERICAN PEOPLE

Bayesian analysis

2024-09-27

Welcome!

- Who we are
- What we do
- How we hope to help you

Benchmarks for success

By the end of this session, participants will be able to:

- Understand how to derive Bayes' Rule from the laws of probability
- Understand how to interpret Bayes' Rule in the context of a data analysis
- Understand how thinking like a Bayesian follows the scientific method and should be a normal part of our thinking about the world

Benchmarks for success

Bonus content:

- Setting Bayesian priors as qualitative research
- Naive Bayes'
- Expectation Maximization

Level Set

Deriving Bayes' Rule

The Frequentist and the Bayesian

Frequentist hypothesis test: $P(data > data_{observed}|H_0)$

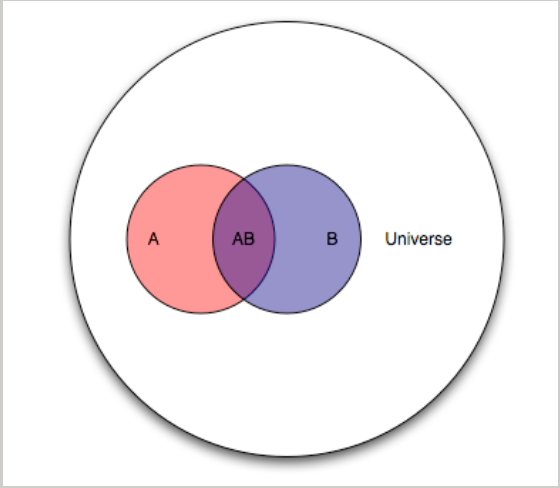
The likelihood: $P(data|model)$

The posterior probability: $P(model|data)$

Bayesian analysis requires us to update our beliefs in light of new data

Bayes' Rule

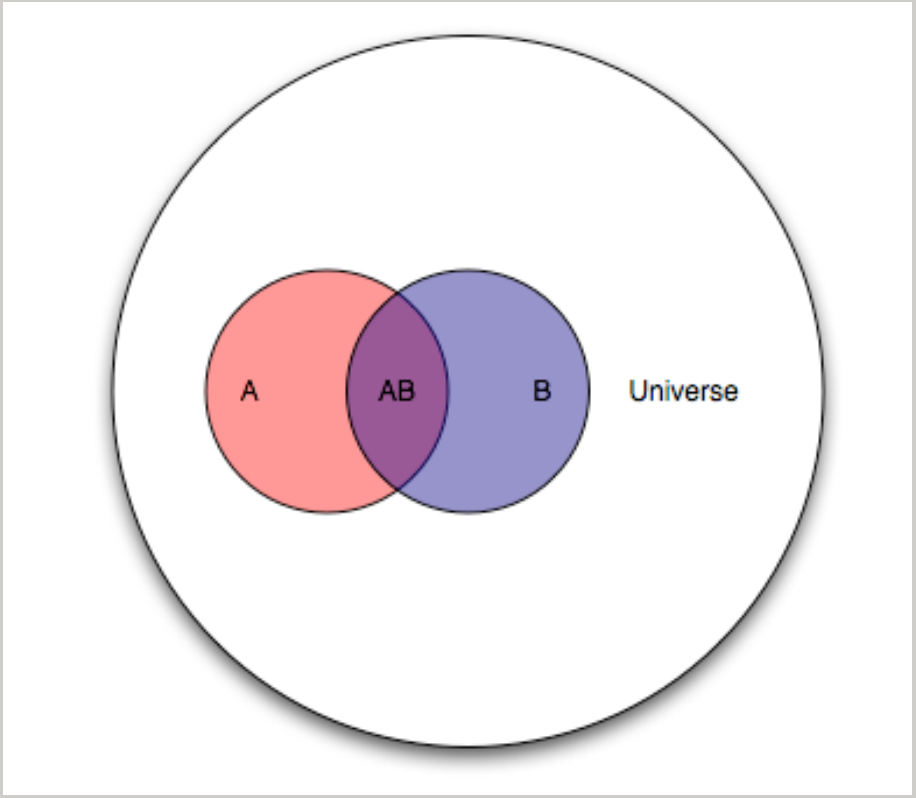
Consider two overlapping events A and B occurring within a universe U.



$$P(A) = \frac{A}{U}$$

$$P(AB) = \frac{AB}{U}$$

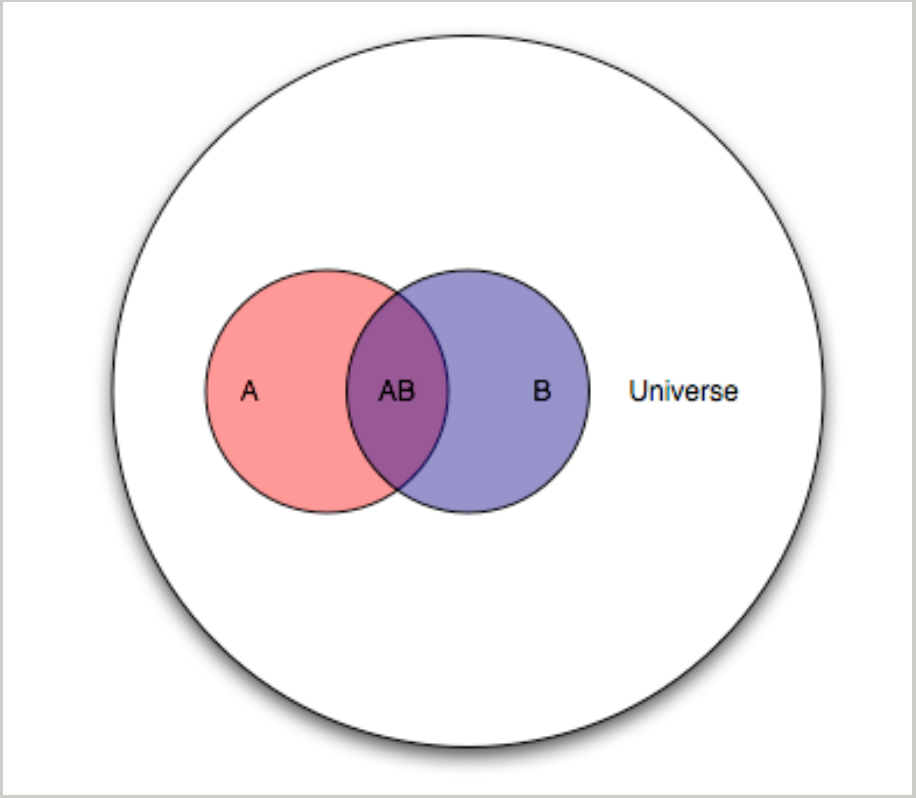
How much of A is in B?



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

$$P(AB) = P(A|B)P(B)$$

How much of B is in A?



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

$$P(AB) = P(B|A)P(A)$$

Putting the two together

We can put the two identities together and solve for $A|B$:

$$P(A|B)P(B) = P(B|A)P(A)$$

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

And that's it. That's Bayes' Rule.

Bayes' Rule as an analytical tool

We just used the laws of probability to derive Bayes' Rule:

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

Now let's use this in the context of a data analysis

$$P(model|data) = \frac{P(data|model)P(model)}{P(data)}$$

Using Bayes' Rule

$$P(model|data) = \frac{P(data|model)P(model)}{P(data)}$$

$P(model|data)$: probability of a hypothesis given data

$P(data|model)$: the likelihood of our data for each hypothesis

$P(model)$: the prior probability of the model, before data

Bayesian inference - example problem

- 40 subjects, half randomly assigned a treatment
- The treatment is expected to reduce the probability of an event
- What is the probability p that an observed event occurred within the treatment group?

Setting up our hypotheses

$H_0 : p = 50\%$ No treatment effect

$H_1 : p < 50\%$ Treatment effect

- 20 events - 4 events in the treatment group and 16 events in the control group
- How likely are these four events to have occurred within the treatment group?

Setting up the Bayesian engine

1. Set a range of plausible values (the model space)
2. Calculate the likelihood of the data for each plausible value
3. Set the prior probability of each plausible value
4. Multiply the likelihood by the prior (numerator)
5. Divide by the denominator to get the posterior probability

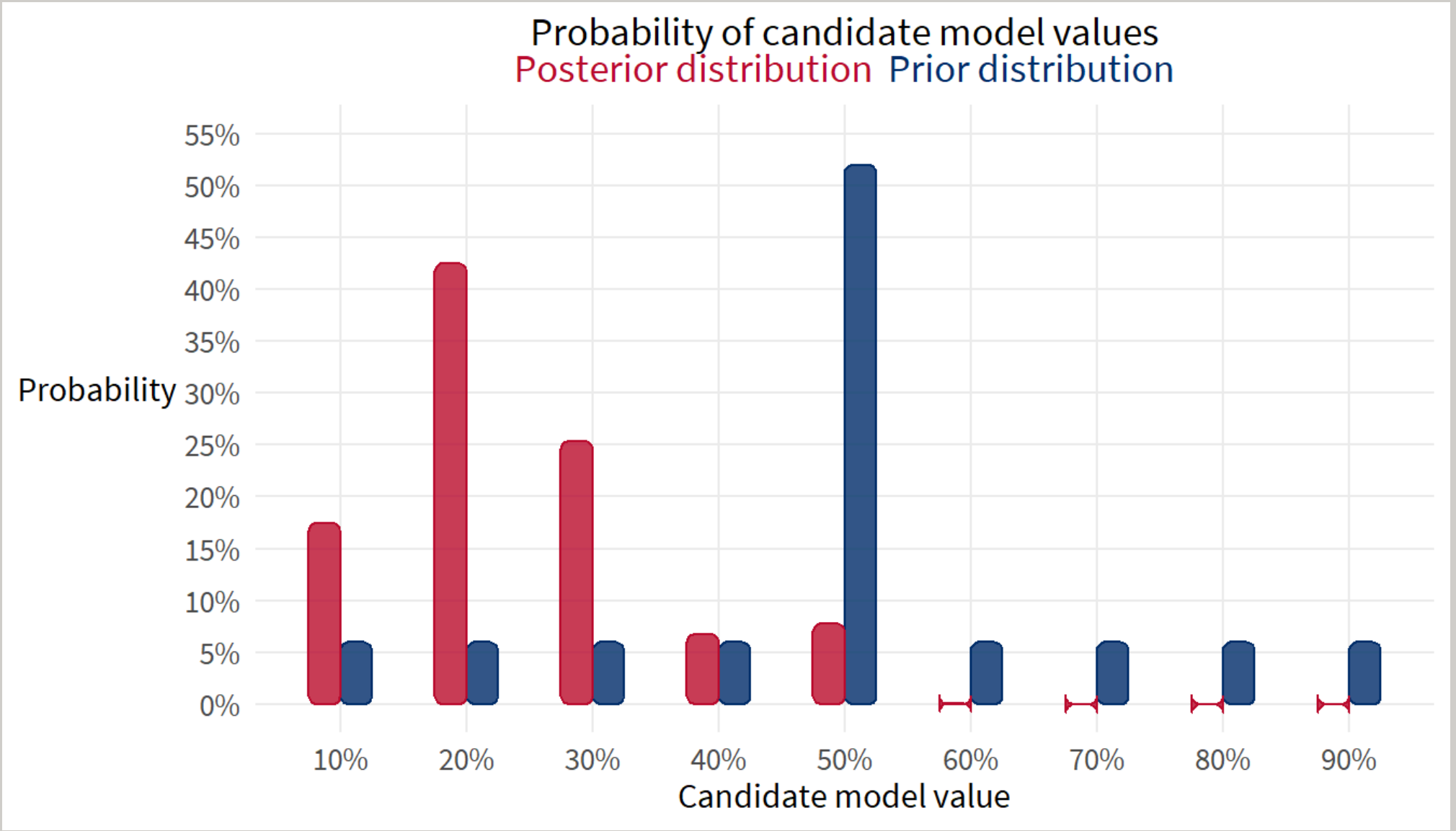
Results of the Bayesian engine

hypothesis	likelihood	prior	numerator	posterior
10%	0.090	6%	0.005	17.5%
20%	0.218	6%	0.013	42.5%
30%	0.130	6%	0.008	25.4%
40%	0.035	6%	0.002	6.8%
50%	0.005	52%	0.002	7.8%
60%	0.000	6%	0.000	0.1%
70%	0.000	6%	0.000	0.0%
80%	0.000	6%	0.000	0.0%
90%	0.000	6%	0.000	0.0%

A treatment effect of 20 percent is most likely

But notice that we get back an entire distribution, not just a point estimate

From prior to posterior



Bayesian analysis as science

- Recall what we learned as kids about the scientific method:
  - Observe a state of the world
  - Develop a hypothesis about how the world works
  - Test your hypothesis with new data
  - Update your beliefs and repeat

Think like a Bayesian

- Using Bayes' Rule to conduct inference follows the scientific method!
- Let's think like a Bayesian
- Stay tuned for Bayes' Rule used in machine learning

Thank you!

Bonus content

- Setting Bayesian priors as qualitative research
- Naive Bayes Classifier
- Expectation Maximization (EM) algorithm