PSCB57 - PROF. HANNO REIN

BAYES'S THEOREM

PLAN:

1. Probability
2. Conditional probability
3. Bayes's theorem
4. Bayesian statistics

PLAN:

1. Probability
2. Conditional probability
3. Bayes’s theorem
4. Bayesian statistics
5. Example

PROBABILITY

- A number between 0 and 1
- Represents a degree of belief in a fact or prediction
- 1 means outcome is certainly true
- 0 represents certainty that fact is false

CONDITIONAL PROBABILITY

- A probability based on some background knowledge

CONDITIONAL PROBABILITY

Example: What is the probability of me having a heart attack in the next year?

According to the CDC 785,000 Americans have a heart attack. There are 311 million Americans.

Thus a random person has a heart attach with a probability of 0.3%.

CONDITIONAL PROBABILITY

Obviously not completely random. Age, blood pressure, smoking, etc play a role.

CONDITIONAL PROBABILITY

Notation:

$$p(A|B)$$

Probability of A given that B is true. A is the prediction. B is a set of conditions.

CONJOINT PROBABILITY

Fancy way of saying: probability that two things are true.

$$p(A \text{ and } B)$$

CONJOINT PROBABILITY

Example 1: Tossing two coins

$$p(A \text{ and } B) = p(A)\, p(B)$$

Probability that first coin lands face up:

$$p(A) = 0.5$$

Probability that second coin lands face up:

$$p(B) = 0.5$$

Works only because coin tosses are independent!

CONJOINT PROBABILITY

Example 2:
A means it rains today. B means it rains tomorrow.

If we know it rained today, then it's more likely that it rains tomorrow.

$$p(B|A) > p(B)$$

CONJOINT PROBABILITY

In general, the probability of a conjunction:

$$p(A \text{ and } B) = p(A)\, p(B|A)$$

So if the probability that it rains on any given day is 0.5. Then the probability of it raining on two consecutive days is greater than 0.25.

CONJOINT PROBABILITY

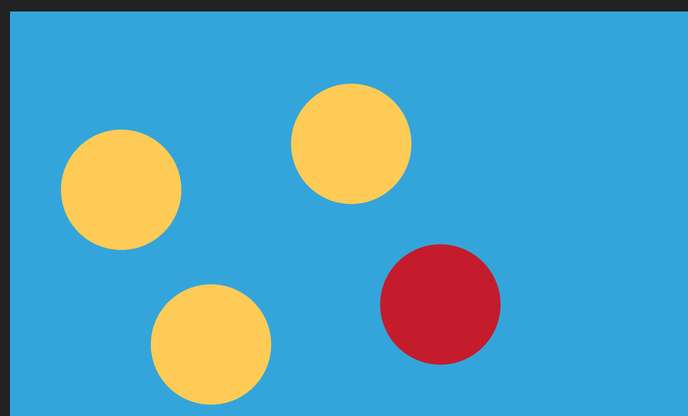
In general, the probability of a conjunction:

$$p(A \text{ and } B) = p(A)\, p(B|A)$$

So if the probability that it rains on any given day is 0.5. Then the probability of it raining on two consecutive days is greater than 0.25.

COOKIE PROBLEM

Two bowls of cookies.



30 vanilla

10 chocolate

20 vanilla

20 chocolate

Choose a bowl randomly, then choose a cookie randomly. It’s vanilla. What is the probability it came from bowl 1?

COOKIE PROBLEM

This is a hard question! Asking a different question is easy:

What is the probability of a vanilla cookie in bowl 1?

$$p(\text{vanilla}|\text{bowl 1}) = 3/4$$

$$p(\text{bowl 1}|\text{vanilla}) = ?$$

BAYES'S THEOREM

Idea: use Bayes's theorem to calculate

$$p(\text{bowl } 1|\text{vanilla}) = ?$$

DERIVATION OF BAYES'S THEOREM

Note that:

$$p(A \text{ and } B) = p(B \text{ and } A)$$

Writing down the conjoined probabilities:

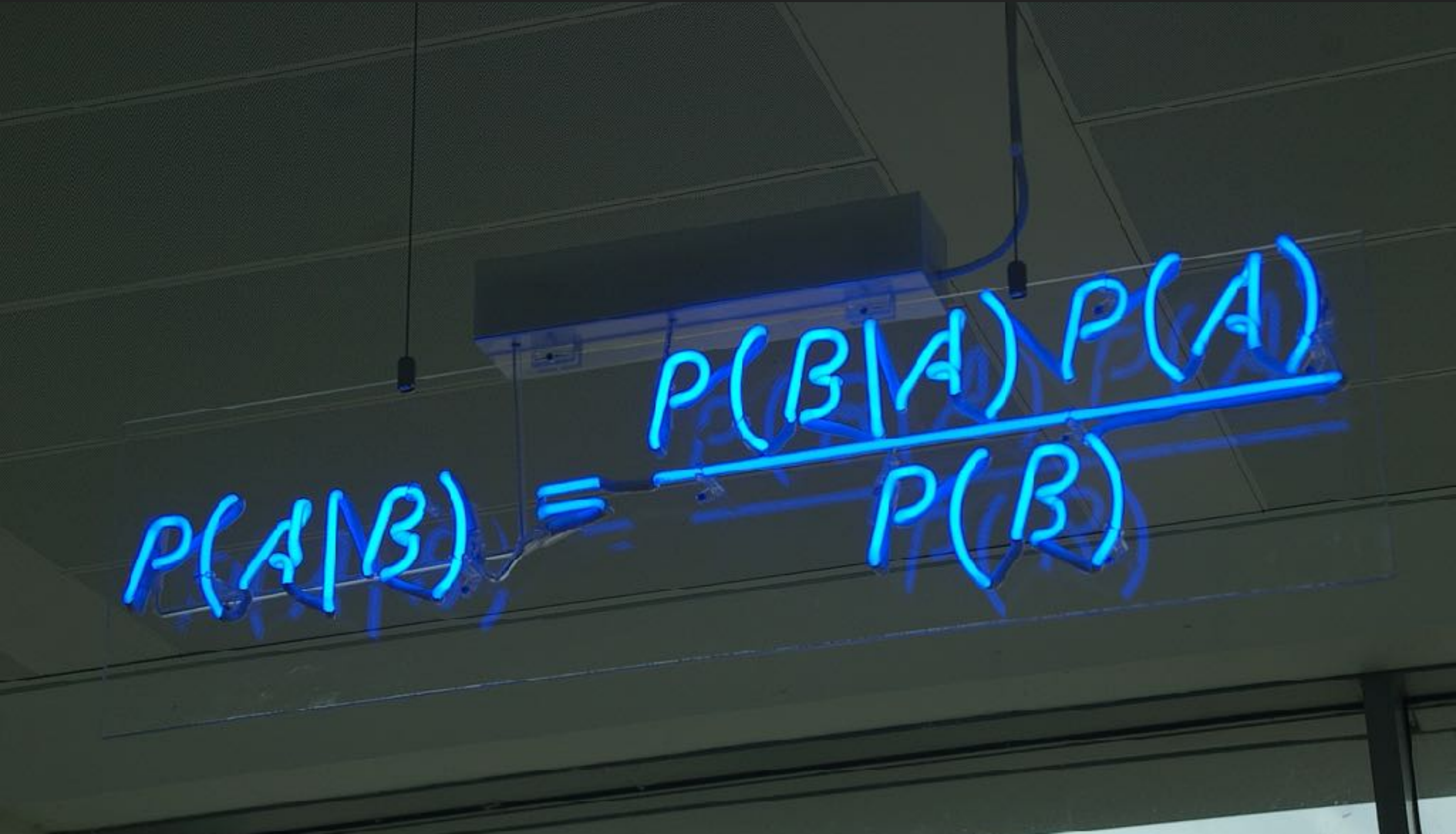
$$p(A \text{ and } B) = p(A)\, p(B|A)$$

$$p(B \text{ and } A) = p(B)\, p(A|B)$$

This gives us:

$$p(A|B) = \frac{p(A)\, p(B|A)}{p(B)}$$

BAYES'S THEOREM



COOKIE PROBLEM REVISITED

Let's call B_1 the hypothesis that the cookie came from bowl 1 and V for vanilla. Bayes's theorem gives us:

$$p(B_1|V) = \frac{p(B_1)p(V|B1)}{p(V)}$$

COOKIE PROBLEM REVISITED

$$p(B_1|V) = \frac{p(B_1)p(V|B1)}{p(V)}$$

$$p(B_1) = \frac{1}{2}$$

$$p(V|B_1) = \frac{3}{4}$$

$$p(V) = \frac{5}{8}$$

$$p(B_1|V) = \frac{\frac{1}{2}\ \frac{3}{4}}{\frac{5}{8}} = \frac{3}{5}$$

DIACHRONIC INTERPRETATION

One way of thinking about Bayes's theorem. Suppose we have a hypothesis H and some data D.

$$p(H|D) = \frac{p(H)\ p(D|H)}{p(D)}$$

These terms now have names and can be interpreted as follows.

DIACHRONIC INTERPRETATION

$$p(H|D) = \frac{p(H)\, p(D|H)}{p(D)}$$

$p(H)$	Prior
$p(H\|D)$	Posterior
$p(D\|H)$	Likelihood
$p(D)$	Normalization constant

DIACHRONIC INTERPRETATION

- Prior is subjective. But that is ok as long as you can write it down.
- Likelihood is easy to compute.
- Normalization constant is tricky. Often this is a high dimensional integral. Monte Carlo methods!

MONTY HALL PROBLEM