Informatics 1 Cognitive Science

Lecture 9: Bayesian Modeling

Frank Keller

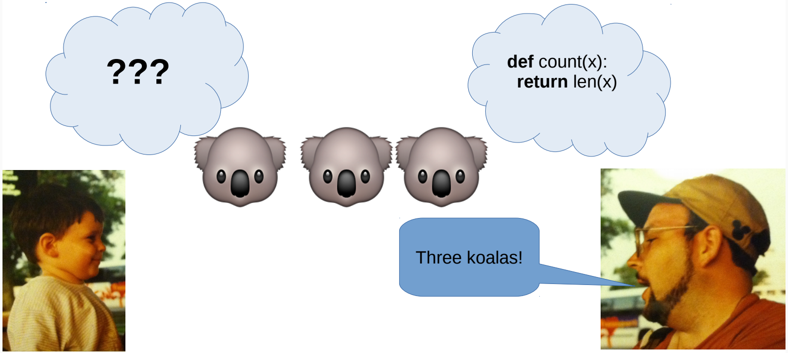
2 February 2024

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

Slide credits: Frank Mollica, Chris Lucas, Mirella Lapata

Overview

Recap



- We observe input and infer the underlying cognitive technology.
- For example: we listen to speech and infer the underlying words.
- More generally: We experiment, we observe data and we update our beliefs.
- Last time, we looked at the use of transitional probabilities $P(y|x)$.
- Today, we will generalize this: Bayesian belief update.

Priors and Posteriors

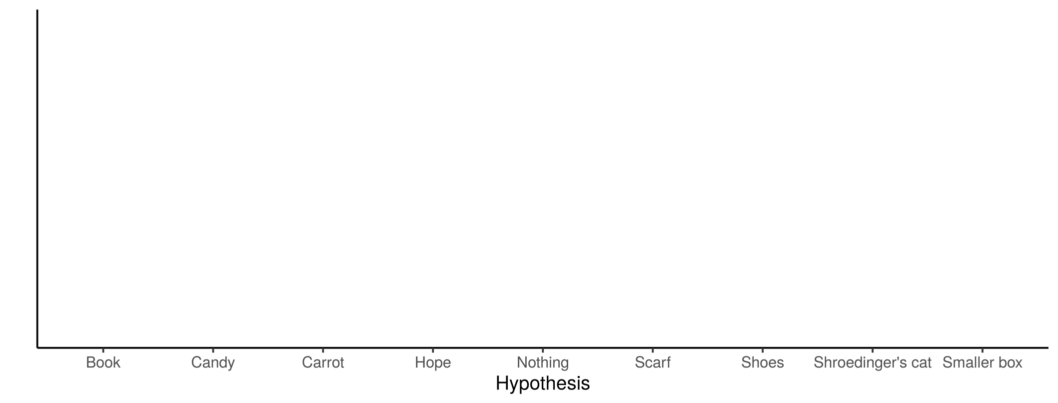
What's in the box?

What’s in the Box?

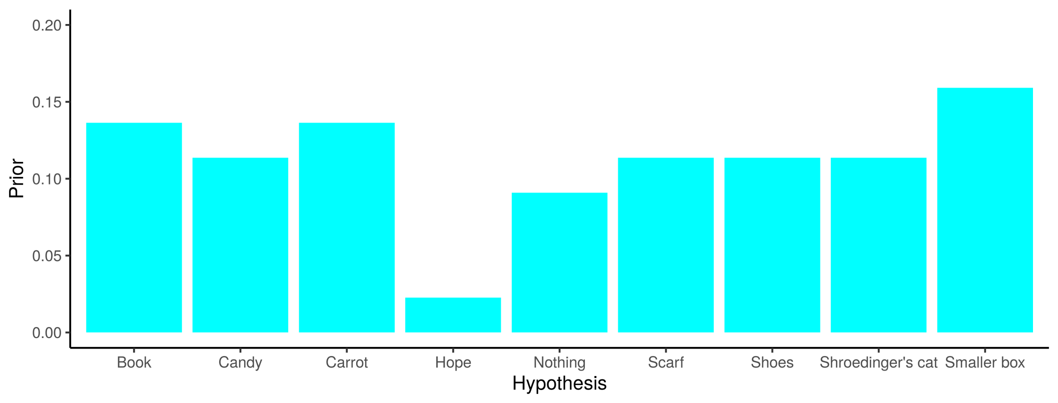
Time for a short quiz on Wooclap!

https://app.wooclap.com/YFCYTE

Hypothesis Space



Prior Beliefs



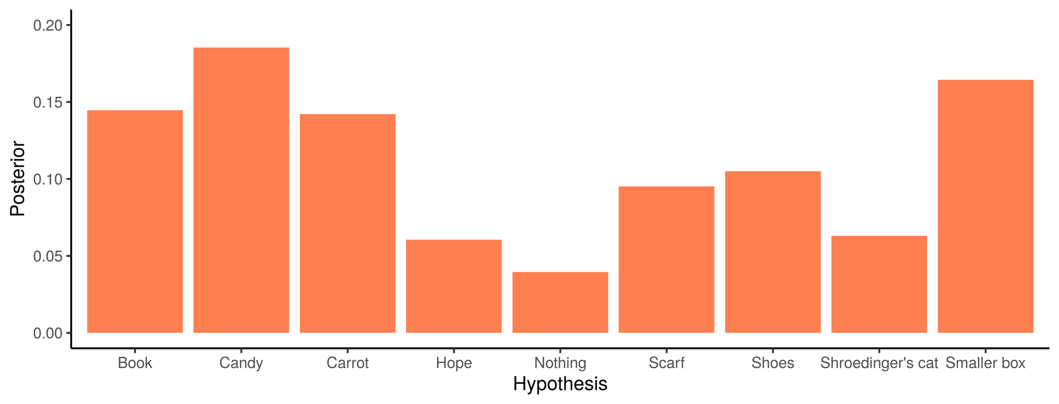
Evidence

Let's watch a short video about the box:

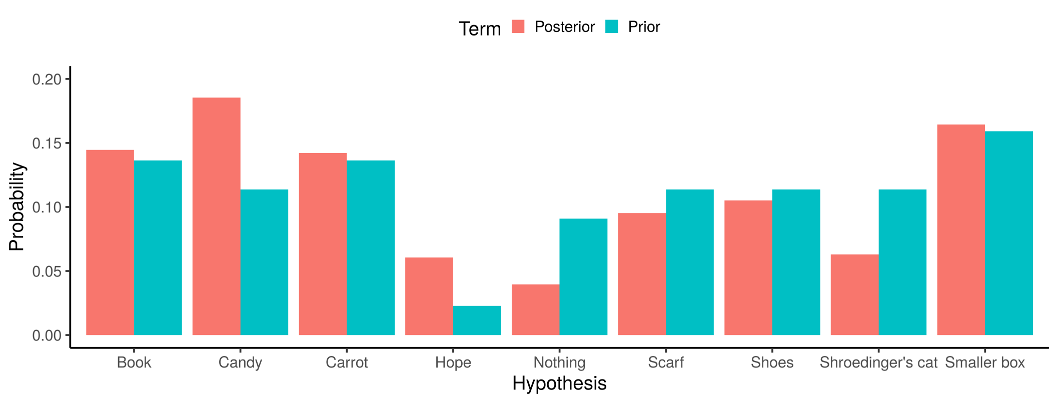
https://groups.inf.ed.ac.uk/teaching/cogsci/course/tutorials/Evidence.mp4

- The video provides additional evidence (data, observations).
- We use this to infer what's in the box.
- Specifically, for each hypothesis, if this hypothesis is true, how likely would it generate that noise?

Posterior Beliefs



Posterior Beliefs



Bayes Rule

Bayes Rule

Bayes Rule

is the calculus of belief updating:

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

- $P(\mathcal{H}|D)$ — **Posterior beliefs:** what you believe about hypotheses after seeing data
- $P(D|\mathcal{H})$ — **Likelihood:** how likely would a given hypothesis result in this data
- $P(\mathcal{H})$ — **Prior beliefs:** what you believe about hypotheses before seeing data
- $P(D)$ — **Evidence:** how likely is this data in general

Bayes Rule Example

Problem

My buddy and I are on vacation in a country where we don't know the local language. The city we are in this week has these awesome hot springs with monkeys and we want to go. Unfortunately, the hot springs do not allow clothing and my friend is really shy about bathing with women. Fortunately, the hot springs has one day a week for men and one day a week for women. Problem: This place is lo-tech and there is no signage to find out which day is for men only.

Solution?

My friend has a good idea! Every day, we will pop by the springs and see who is leaving. We can use that information to solve our problem.

Bayes Rule Example: Hypothesis Space

	Hypothesis
h_1	Today is a women only day.
h_2	Today has no restriction.
h_3	Today is a men only day.

Bayes Rule Example: Prior

What do we know know before we walk out the door?

	Hypothesis	Prior
h_1	Today is a women only day.	$\frac{1}{7}$
h_2	Today has no restriction.	$\frac{5}{7}$
h_3	Today is a men only day.	$\frac{1}{7}$

Bayes Rule Example: Likelihood

How likely would it be to see a man leave under each hypothesis?

	Hypothesis	Prior	Like$_M$
h_1	Today is a women only day.	$\frac{1}{7}$	0
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$
h_3	Today is a men only day.	$\frac{1}{7}$	1

Bayes Rule Example: Likelihood

How likely would it be to see a woman leave under each hypothesis?

	Hypothesis	Prior	Like$_M$	Like$_W$
h_1	Today is a women only day.	$\frac{1}{7}$	0	1
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{1}{2}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	0

Bayes Rule Terminology

Time for a short quiz on Wooclap!

https://app.wooclap.com/YFCYTE

Bayesian Update

Bayes Rule Example: Update

So we arrive at the hot springs on day one. We wait five minutes and see a man exit. My friend wants to be 90% sure it's men only before going in. Should we leave, go in or wait longer?

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	**Prior**	**Like$_M$**	**Post. Score**
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Bayes Rule Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Evidence $$P(D) = \sum_h P(D|h)P(h)$$

Bayes Rule Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Evidence $$P(D) = \sum_h P(D|h)P(h) = \frac{5}{14} + \frac{1}{7} = \frac{7}{14} = \frac{1}{2}$$

Bayes Rule Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score	Posterior
h_1	Women only day.	$\frac{1}{7}$	0	0	0
h_2	No restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{5}{14}$	$\frac{5}{7}$
h_3	Men only day.	$\frac{1}{7}$	1	$\frac{1}{7}$	$\frac{2}{7}$

Evidence

$$P(D) = \sum_h P(D|h)P(h) = \frac{5}{14} + \frac{1}{7} = \frac{7}{14}$$

Bayes Rule Example: Update 2

We wait another five minutes and see five more men exit. My friend wants to be 90% sure it's men only before going in. Should we leave, go in or wait longer?

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)^6(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})^6(\frac{5}{7}) = \frac{5}{448}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)^6(\frac{1}{7}) = \frac{1}{7}$

Bayes Rule Example: Update 2

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)^6(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})^6(\frac{5}{7}) = \frac{5}{448}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)^6(\frac{1}{7}) = \frac{1}{7}$

$$P(D) = \frac{5}{448} + \frac{1}{7} = \frac{69}{448}$$

Bayes Rule Example: Update 2

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score	Posterior
h_1	Women only day.	$\frac{1}{7}$	0	0	0
h_2	No restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{5}{448}$	$\frac{5}{69} = 0.072$
h_3	Men only day.	$\frac{1}{7}$	1	$\frac{64}{448}$	$\frac{64}{69} = 0.928$

$$P(D) = \frac{5}{448} + \frac{1}{7} = \frac{69}{448}$$

Rational Analysis

Rational Analysis

- Bayes Rule requires us to formalize the important components of a learning mechanism.
- If a specification perfectly captures the learning problem, Bayes Rule provides us a normative model for what the optimal behavior should look like.
- If we write down the problem incorrectly, human behavior may deviate from the model's predictions.
- We can then improve our specification and try again.
- This cyclic modeling approach using specifications using Bayes rule is called Rational Analysis.

Summary

- Bayes Rule is a way of calculating the probability of a hypothesis, given some evidence (data).
- It includes four terms: posterior $P(\mathcal{H}|D)$; likelihood $P(D|\mathcal{H})$; prior $P(\mathcal{H})$; evidence $P(D)$.
- The prior encodes how probable a hypothesis is without any evidence, the likelihood encodes how probable it is to observe this evidence if the hypothesis is true.
- The posterior is the result of combining the two.
- Bayesian update: we can re-compute the posterior as we obtain more evidence for or against the hypothesis.
- Bayes rule provides a way of analyzing learning mechanism.
- Can be used for broad range of cognitive processes: rational analysis.