Probabilistic Modeling and Rational Analysis

Frank Jäkel

Institut für Kognitionswissenschaft
Universität Osnabrück

KogWis 2014

Part 1: Bayes' Rule

Understanding Bayes' Rule

Bayes' Rule

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B \mid A) \cdot P(A) + P(B \mid \neg A) \cdot P(\neg A)}$$

Lie Detection

1. Studies show that 12.5% of people have at least once stolen something from their employer (pencil, stapler, belegtes Brötchen, etc.)
2. You are working for me and you have to undergo a lie detection test to check whether you have ever stolen something
3. The lie detection test has a 87.5% accuracy
4. What is the probability that someone with a test result that says he is a thief has actually ever stolen something?

Are You A Thief?

1. Take a coin and flip it 3 times.
2. If you get 3 heads (1/8=12.5%) you are a thief (T) otherwise you are not ($\neg T$).

Are You A Thief?

1. Now you are hooked up to a lie detector and I ask: Are you a thief?
2. Of course, you all say: No!
3. Take your coin and flip it 3 more times.
4. You are a thief:
    1. If you do *not* have 3 heads (7/8=87.5% accuracy) your lie is caught (L)
    2. If you have 3 heads you are not caught ($\neg L$)
5. You are not a thief:
    1. If you do *not* have 3 heads (7/8=87.5% accuracy) the lie detector states that you are speaking the truth ($\neg L$)
    2. If you have 3 heads the lie detector thinks you lie (L)!

Bayes' Rule Example: Medical Testing

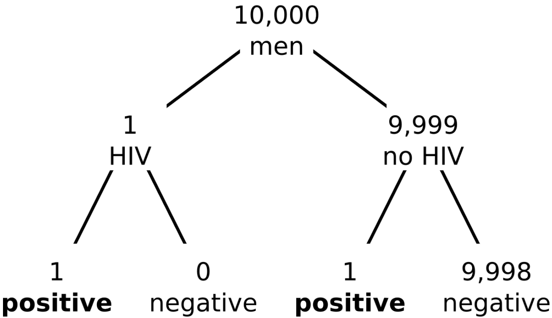
About 0.01 percent of men with no known risk behavior are infected with HIV. If such a man has the virus, there is a 99.9 percent chance that the test result will be positive. If a man is not infected, there is a 99.99 percent chance that the test result will be negative.

What is the probability that a man from the low risk group has HIV given that he gets a positive test result?

- *base rate* or *prior* probability: $P(\text{HIV}) = 0.0001$
- *hit rate* of the test: $P(\text{positive} \mid \text{HIV}) = 0.999$
- *false alarm rate* of the test: $P(\text{positive} \mid \neg\text{HIV}) = 0.0001$
- What is the *posterior* probability $P(\text{HIV} \mid \text{positive})$?

[Gigerenzer, 2003, chapter 7]

Bayes' Rule Example: Medical Testing

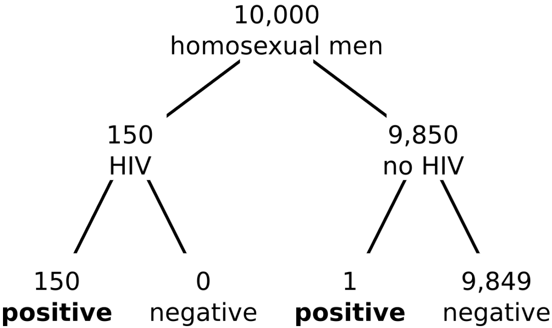


$$P(\text{HIV} \mid \text{positive}) \approx \frac{1}{1+1} = \frac{1}{2}$$

[Gigerenzer, 2003, chapter 7]

Base Rates Matter

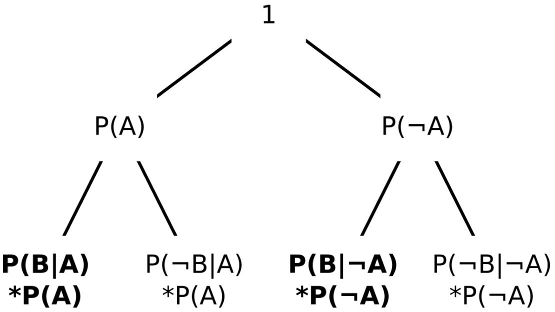
What is the reference class? The base rate for homosexual men is higher, say, 1.5 percent.



$$P(\text{HIV} \mid \text{positive}) \approx \frac{150}{1+150}$$

[Gigerenzer, 2003, chapter 7]

Bayes' Rule



$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B \mid A) \cdot P(A) + P(B \mid \neg A) \cdot P(\neg A)}$$

$$\left(\text{posterior} = \frac{\text{likelihood} \cdot \text{prior}}{\text{evidence}}\right)$$

Base Rate Fallacy

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

- 85% of the cabs in the city are Green and 15% are Blue
- a witness identified the cab as Blue.

The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time. What is the probability that the cab involved in the accident was Blue rather than Green?

The median and modal answer is typically .8: A total *neglect of the base rate*.

[Tversky and Kahneman, 1982]

Part 2: Bayesian Statistics

[Berry, 1995, Kruschke, 2011, for more]

Probability Theory and Bayesian Statistics

"It is seen in this essay that the theory of probabilities is at bottom only common sense reduced to calculus; it makes us appreciate with exactitude that which exact minds feel by a sort of instinct without being able ofttimes to give reason for it. It leaves no arbitrariness in the choice of opinions and sides to be taken; and by its use can always be determined the most advantageous choice. Thereby it supplements most happily the ignorance and the weakness of the human mind." [Laplace, 1902, p. 196]

Bayes' Rule for Statistics

Bayes' Rule

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B \mid A) \cdot P(A) + P(B \mid \neg A) \cdot P(\neg A)}$$

Bayes' Rule for Statistics

Bayes' Rule

$$P(H_1 \mid \text{data}) = \frac{P(\text{data} \mid H_1) \cdot P(H_1)}{P(\text{data} \mid H_1) \cdot P(H_1) + P(\text{data} \mid H_2) \cdot P(H_2)}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

H_2	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid \text{data}) = \frac{P(\text{data} \mid H_1) \cdot P(H_1)}{P(\text{data} \mid H_1) \cdot P(H_1) + P(\text{data} \mid H_2) \cdot P(H_2)}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid X = 5) = \frac{P(X = 5 \mid H_1) \cdot P(H_1)}{P(X = 5 \mid H_1) \cdot P(H_1) + P(X = 5 \mid H_2) \cdot P(H_2)}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid X = 5) = \frac{\frac{3}{36} \cdot P(H_1)}{\frac{3}{36} \cdot P(H_1) + \frac{9}{36} \cdot P(H_2)}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid X = 5) = \frac{1 \cdot P(H_1)}{1 \cdot P(H_1) + 3 \cdot P(H_2)}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid X = 5) = \frac{1 \cdot P(H_1)}{1 \cdot P(H_1) + 3 \cdot P(H_2)} \text{ with } P(H_1) = P(H_2) = \frac{1}{2}$$

Example

Problem

I roll two dice behind your back. I either tell you the minimum (H_1) or the maximum (H_2) of the two numbers that came up. I tell you you I got a 5. What is your probability that H_1 is true?

H_1	1	2	3	4	5	6	H_2	1	2	3	4	5	6
	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$$P(H_1 \mid X = 5) = \frac{1}{1+3} = \frac{1}{4}$$

Frequentist vs Subjective Probabilities (Caricature)

Frequentist (aka Orthodox)

Probabilities are limits of relative frequencies. They are measurable by counting events in the world.

Subjectivist (aka Bayesian)

Probabilities are degrees of belief. They are measurable by asking people for their opinion.

Bayes' Rule ...

... applies to both! But subjectivists can apply it more broadly.

[Hacking, 1975, for more]

Frequentist vs Subjective Probabilities (Caricature)

Frequentist (aka Orthodox)

Probabilities are limits of relative frequencies. They are measurable by counting events in the world.

Subjectivist (aka Bayesian)

Probabilities are degrees of belief. They are measurable by asking people for their opinion.

Bayes' Rule ...

... applies to both! But subjectivists can apply it more broadly.

[Hacking, 1975, for more]

An Aside: p-values

- It is clear from Bayes' rule that $P(A \mid B) \neq P(B \mid A)$.
- p-value is not $P(H \mid \text{data})$.

[Cohen, 1994]

Part 3: Rational Analysis

[Green and Swets, 1988, Anderson, 1991a, Anderson, 1991b, Chater and Oaksford, 1999, Oaksford and Chater, 2001, Kersten and Schrater, 2002, Kersten and Yuille, 2003, Chater et al., 2006, Tenenbaum et al., 2006, Tenenbaum et al., 2011, for more]

Categorization Example: Task

Task

White bars of varying length are presented on an otherwise black screen. The bars come from two two equally probable categories: The short and the long bars, named category 1 and category 2 respectively.

1. Stimuli from category 1 are drawn from a normal distributions with mean 10 cm and standard deviation 1 cm.
2. Stimuli from category 2 are drawn from a normal distributions with mean 12 cm and standard deviation 1 cm.

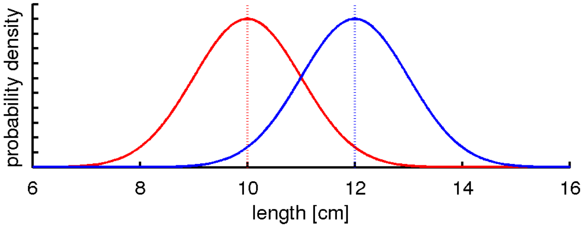
Categorization Example: Task

Task

White bars of varying length are presented on an otherwise black screen. The bars come from two equally probable categories: The short and the long bars, named category 1 and category 2 respectively. With $\sigma = 1$ and $\mu_1 = 10$ and $\mu_2 = 12$:

1. $p(X = x \mid C = 1) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right)$
2. $p(X = x \mid C = 2) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_2)^2\right)$
3. $p(C = 1) = p(C = 2) = \frac{1}{2}$

Categorization Example: Task



Categorization Example: Ideal Observer

Task

1. $p(X = x \mid C = 1) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right)$
2. $p(X = x \mid C = 2) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_2)^2\right)$
3. $p(C = 1) = p(C = 2) = \frac{1}{2}$

Question

How would an ideal subject solve the task? What is the best strategy?

Categorization Example: Posterior

Task

1. $p(X = x \mid C = 1) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(x - \mu_1\right)^2\right)$
2. $p(X = x \mid C = 2) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(x - \mu_2\right)^2\right)$
3. $p(C = 1) = p(C = 2) = \frac{1}{2}$

$$p(C = 1 \mid X = x) = \frac{p(X = x \mid C = 1)p(C = 1)}{p(X = x \mid C = 1)p(C = 1) + p(X = x \mid C = 2)p(C = 2)}$$

Categorization Example: Posterior

Task

1. $p(X = x \mid C = 1) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right)$
2. $p(X = x \mid C = 2) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_2)^2\right)$
3. $p(C = 1) = p(C = 2) = \frac{1}{2}$

$$p(C = 1 \mid X = x) = \frac{p(X = x \mid C = 1)}{p(X = x \mid C = 1) + p(X = x \mid C = 2)}$$

Categorization Example: Posterior

Task

1. $p(X = x \mid C = 1) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right)$
2. $p(X = x \mid C = 2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu_2)^2\right)$
3. $p(C = 1) = p(C = 2) = \frac{1}{2}$

$$p(C = 1 \mid X = x) = \frac{\exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}(x - \mu_1)^2\right) + \exp\left(-\frac{1}{2\sigma^2}(x - \mu_2)^2\right)}$$

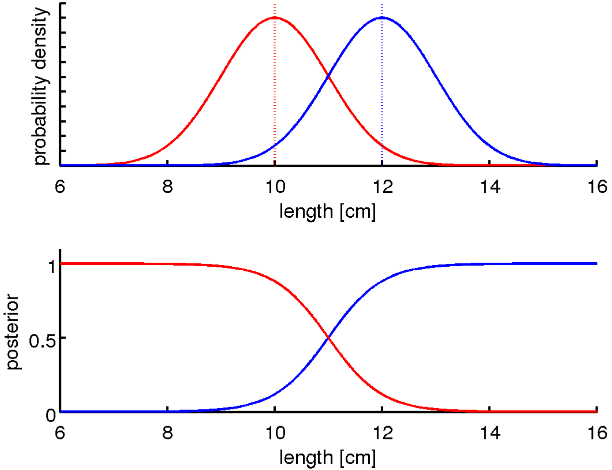
Categorization Example: Posterior

$$p(C=1 \mid X=x) = \frac{\exp\left(-\frac{1}{2\sigma^2}(x-\mu_1)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}(x-\mu_1)^2\right) + \exp\left(-\frac{1}{2\sigma^2}(x-\mu_2)^2\right)}$$

$$= \frac{1}{1+\exp\left(-\frac{1}{2\sigma^2}\left((x-\mu_2)^2-(x-\mu_1)^2\right)\right)}$$

$$p(C=2 \mid X=x) = 1 - p(C=1 \mid X=x)$$

Categorization Example: Posterior



Categorization Example: Decision Rule

Respond "1" if and only if

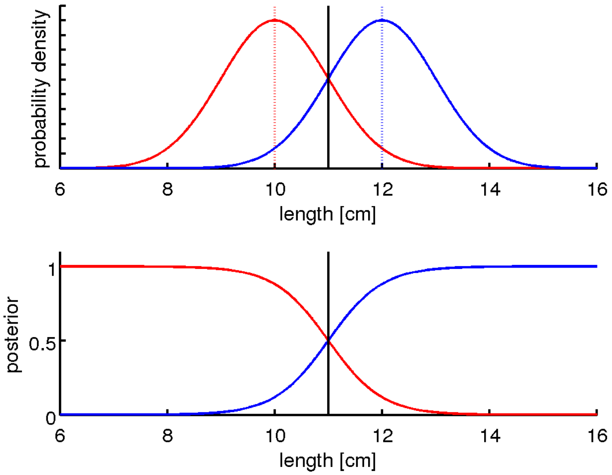
$$
\begin{aligned}
p(C=1 \mid X=x) &\geq \frac{1}{2} \\
\frac{1}{1+\exp\left(-\frac{1}{2\sigma^2}\left((x-\mu_2)^2-(x-\mu_1)^2\right)\right)} &\geq \frac{1}{2} \\
\exp\left(-\frac{1}{2\sigma^2}\left((x-\mu_2)^2-(x-\mu_1)^2\right)\right) &\leq 1 \\
\left((x-\mu_1)^2-(x-\mu_2)^2\right) &\leq 0 \\
(x-\mu_1)^2 &\leq (x-\mu_2)^2
\end{aligned}
$$

Categorization Example: Decision Rule

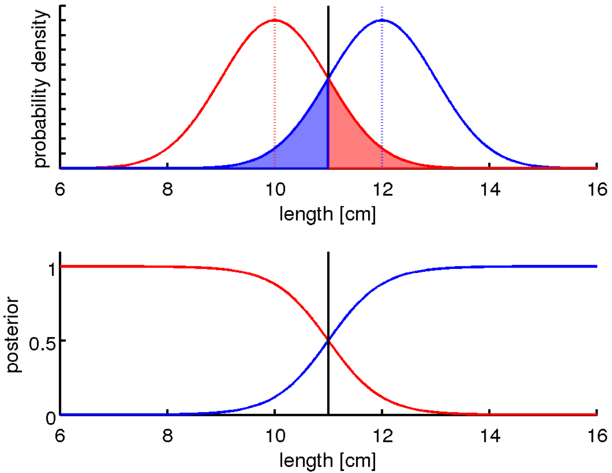
Respond "1" if and only if

$$
\begin{aligned}
(x-\mu_1)^2 &\le (x-\mu_2)^2 \\
x^2 - 2x\mu_1 + \mu_1^2 &\le x^2 - 2x\mu_2 + \mu_2^2 \\
2x(\mu_2 - \mu_1) &\le \mu_2^2 - \mu_1^2 \\
2x(\mu_2 - \mu_1) &\le (\mu_2 + \mu_1)(\mu_2 - \mu_1) \\
x &\le \frac{1}{2}(\mu_2 + \mu_1) \\
x &\le c
\end{aligned}
$$

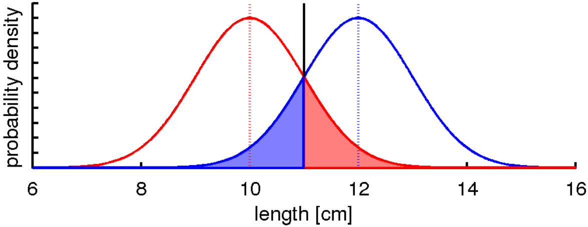
Categorization Example: Decision Rule



Categorization Example: Minimum Errors



Categorization Example: Minimum Errors

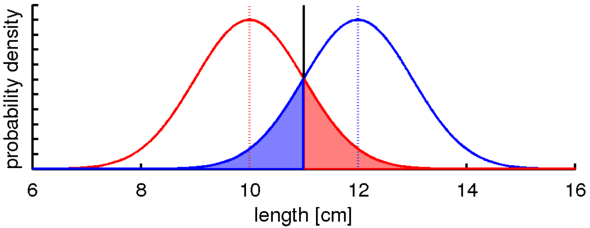


$$p(R = 1 \mid C = 2) = \int_{-\infty}^{c} \phi(x; \mu_2, \sigma) dx = \Phi(c; \mu_2, \sigma)$$

$$p(R = 2 \mid C = 1) = \int_{c}^{+\infty} \phi(x; \mu_1, \sigma) dx = 1 - \Phi(c; \mu_1, \sigma)$$

$$p(\text{Error}) = \frac{1}{2}(1 - \Phi(c; \mu_1, \sigma)) + \frac{1}{2}\Phi(c; \mu_2, \sigma)$$

Categorization Example: Minimum Errors



$$p(\text{Error}) = \frac{1}{2}(1 - \Phi(c; \mu_1, \sigma)) + \frac{1}{2}\Phi(c; \mu_2, \sigma)$$

$$0 = -\frac{1}{2}\phi(c; \mu_1, \sigma) + \frac{1}{2}\phi(c; \mu_2, \sigma)$$

$$0 < \phi'(c_{\text{opt}}; \mu_2, \sigma)$$

Categorization Example: Ideal Observer

Ideal Observer for Categorization Task

Respond "1" if and only if bar length $x \leq \frac{1}{2}(\mu_2 + \mu_1) = 11$ cm. This strategy will minimize the expected number of errors.

Note

The ideal observer for this categorization task does not compute probabilities, does not use Bayes' rule and it does not optimize. It only checks the stimulus against a decision criterion. Probability theory was merely used as a tool to derive the optimal strategy.

Categorization Example: Ideal Observer

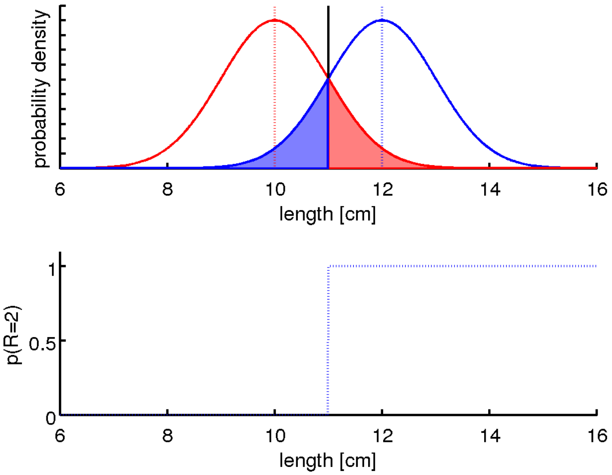
Ideal Observer for Categorization Task

Respond "1" if and only if bar length $x \leq \frac{1}{2}(\mu_2 + \mu_1) = 11$ cm. This strategy will minimize the expected number of errors.

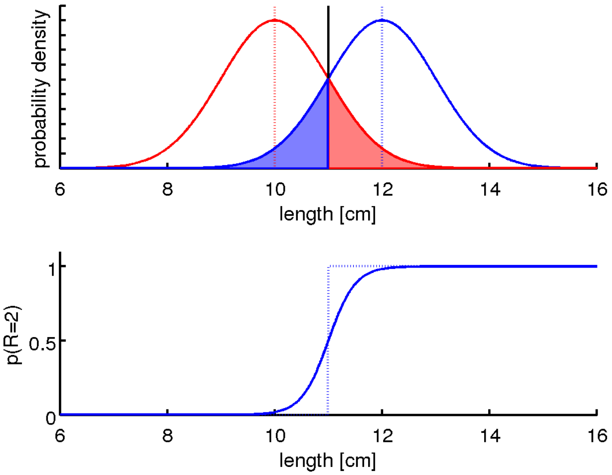
Note

The ideal observer for this categorization task does not compute probabilities, does not use Bayes' rule and it does not optimize. It only checks the stimulus against a decision criterion. Probability theory was merely used as a tool to derive the optimal strategy.

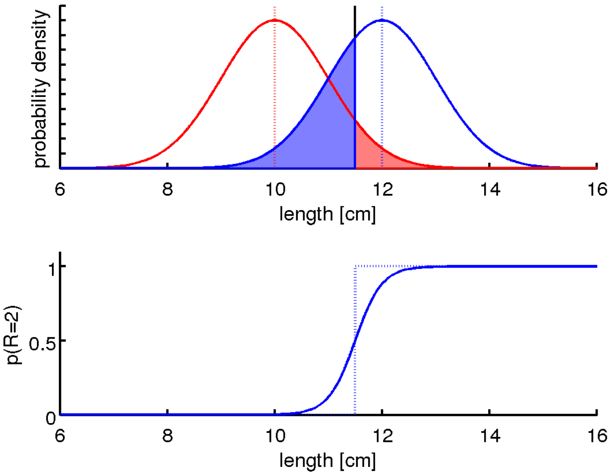
Categorization Example: Ideal Observer Responses



Categorization Example: Actual Responses



Categorization Example: Actual Responses



Ideal Observer Analysis: A Heuristic Research Strategy

“[...] In any study, it is desirable to specify rigorously the factors pertinent to the study. [...] The ideal performance, in other words, constitutes a convenient base from which to explore the complex operation of a real organism. [...] The problem then becomes one of changing the ideal model in some particular so that it is slightly less than ideal. [...] This method of attack has been found to generate useful hypotheses for further studies. Thus, whereas it is not expected that the human observer and the ideal detection device will behave identically, the emphasis in early studies is on similarities. If the differences are small, one may rule out entire classes of alternative models, and regard the model in question as a useful tool in further studies. Proceeding on this assumption, one may then in later studies emphasize the differences, the form and extent of the differences suggesting how the ideal model may be modified in the direction of reality.“
[Swets et al., 1961, p. 311]

Category Learning as Ideal Learning

Task

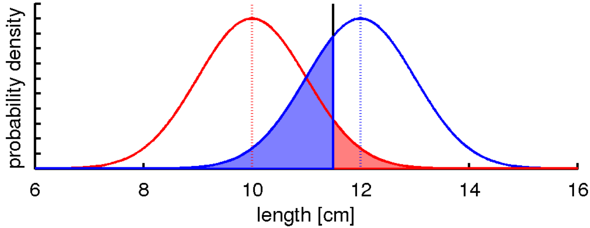
White bars of varying length ($x_1 \ldots x_n$) are presented on an otherwise black screen. The lengths of the bars are drawn i.i.d. from two equally probable categories: The short and the long bars, named category 1 and category 2, respectively. With $\sigma = 1$ and $\mu_1 = 10$ and $\mu_2 = 12$:

1. $p(X_k = x_k \mid C_k = 1, \mu_1) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(x_k - \mu_1\right)^2\right)$
2. $p(X_k = x_k \mid C_k = 2, \mu_2) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(x_k - \mu_2\right)^2\right)$
3. $p(C_k = 1) = p(C_k = 2) = \frac{1}{2}$

Question

Before we assumed that the ideal observer knows μ_1, μ_2 and σ. How would an ideal observer solve the task without this knowledge?

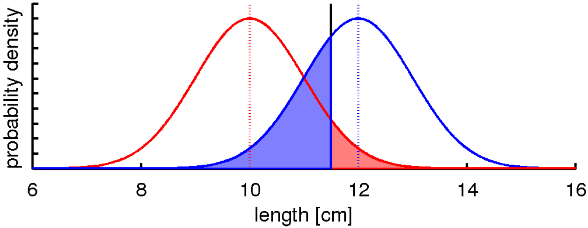
Category Learning as Error Learning



Question

As we know from the ideal observer analysis that the optimal strategy implements a decision criterion c, can we come up with a learning mechanism that iteratively updates the criterion and converges (in some sense) to the optimal criterion?

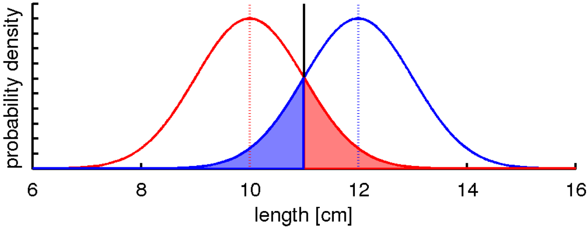
Category Learning as Error Learning



$$p(R = 1 \mid C = 2) = \Phi(c; \mu_2, \sigma)$$
$$p(R = 2 \mid C = 1) = 1 - \Phi(c; \mu_1, \sigma)$$
$$c_{t+1} = \begin{cases} c_t - \delta & \text{if } R = 1 \text{ and } C = 2 \\ c_t + \delta & \text{if } R = 2 \text{ and } C = 1 \end{cases}$$

[Kac, 1969]

Category Learning as Error Learning



$$p(R=1 \mid C=2) = p(R=2 \mid C=1)$$
$$\Phi(c;\mu_2,\sigma) = 1-\Phi(c;\mu_1,\sigma)$$
$$c_{t+1} = \begin{cases} c_t - \delta & \text{if } R=1 \text{ and } C=2 \\ c_t + \delta & \text{if } R=2 \text{ and } C=1 \end{cases}$$

[Kac, 1969]

Part 4: Response to Criticism of Rational Analysis

Criticism of Bayesian Models

- Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition [Jones and Love, 2011]
- Bayesian just-so stories in psychology and neuroscience [Bowers and Davis, 2012]
- How Robust Are Probabilistic Models of Higher-Level Cognition? [Marcus and Davis, 2013]

Criticism of Bayesian Models

- Neglect of mechanisms
- Role of optimality/rationality
- Post-hoc rationalizations

Bayesian Fundamentalism

"[...] the primary goal of much Bayesian cognitive modeling has been to demonstrate that human behavior in some task is rational with respect to a particular choice of Bayesian model. We refer to this school of thought as *Bayesian Fundamentalism*, because it strictly adheres to the tenet that human behavior can be explained through rational analysis – once the correct probabilistic interpretation of the task environment has been identified – without recourse to process, representation, resource limitations, or physiological or developmental data." [Jones and Love, 2011, p. 170]

- Explain behavior only through rational analysis of task and environment
- Mechanisms play no role
- Psychology of the empty organism

Obsession with Rationality/Optimality?

"[...] Together these studies demonstrate that people are adept at combining prior knowledge with new evidence in a manner predicted by Bayesian statistics. [...] Recent studies have found that when combining information this way, people are also similar to optimal. [...] Their performance in these cue-combination trials can be predicted using the rules of Bayesian integration, further evidencing people's ability to optimally cope with uncertain information. [...] In typical cases cues are combined by the subjects in a fashion that is close to the optimum prescribed by Bayesian statistics." [Berniker and Körding, 2011, p. 422]

Mechanisms for Probabilistic Inference

- Bayesian modelers do care about mechanisms!
- But what role do the (optimal, rational) Bayesian models play in discovering mechanisms?

Computational Level

Computational theory	*Representation and algorithm*	*Hardware implementation*
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and the output, and what is the algorithm for the transformation?	How can the representation and the algorithm be realized physically?

[Marr, 2010, p. 25]

- Bayesian models formalize computational level analysis

A Heuristic Research Strategy

"[...] In any study, it is desirable to specify rigorously the factors pertinent to the study. [...] The ideal performance, in other words, constitutes a convenient base from which to explore the complex operation of a real organism. [...] The problem then becomes one of changing the ideal model in some particular so that it is slightly less than ideal. [...] This method of attack has been found to generate useful hypotheses for further studies. Thus, whereas it is not expected that the human observer and the ideal detection device will behave identically, the emphasis in early studies is on similarities. If the differences are small, one may rule out entire classes of alternative models, and regard the model in question as a useful tool in further studies. Proceeding on this assumption, one may then in later studies emphasize the differences, the form and extent of the differences suggesting how the ideal model may be modified in the direction of reality."
[Swets et al., 1961, p. 311]

A Heuristic Research Strategy

Marr-Bayes Reverse-Engineering

Bayesian ideal observer analysis is a formal framework for computational-level modeling. The computational level is key in reverse-engineering the mind/brain. But ultimately we want explanations on all three levels: Computational, algorithmic, and implementational.

- Bayesian model is only the first step!
- Mechanisms are important and we want to know how they solve the problems that the organism faces
- But how does the computational level analysis help us in finding mechanisms?

Strong Constraints from Bayesian Models

"Recent psychophysical experiments indicate that humans perform near-optimal Bayesian inference in a wide variety of tasks, ranging from cue integration to decision making to motor control. *This implies* that neurons both represent probability distributions and combine those distributions according to a close approximation to Bayes' rule." [Ma et al., 2006, p. 1423]

Bayesian Realism

- The brain directly implements prior, likelihood, loss functions, etc.
- There are mechanisms for representing distributions, evidence integration, computing Bayes' rule, and choosing optimal decisions.
- Look for Bayes in the brain!
- But what about the decision criterion in our categorization example?

Multiple Realizability

"Using probabilistic models to provide a computational-level explanation does not require that hypothesis spaces or probability distributions be explicitly represented by the underlying psychological or neural processes, or that people learn and reason by explicitly using Bayes' rule"
[Griffiths et al., 2010, p. 362]

- Yes, but how then does it help in reverse-engineering the mind/brain?

Instrumentalist Bayesianism

- Bayesian models are just useful in summarizing and organizing the data
- Bayesian models don't constrain mechanistic explanations at all (beyond fitting the data)

Weak Constraints from Bayesian Models

"The three levels are coupled, but only loosely. The choice of an algorithm is influenced, for example, by what it has to do and by the hardware in which it must run. But there is a wide choice available at each level, and the explication of each level involves issues that are rather independent of the other two." [Marr, 2010, p. 25]

Pragmatic Bayesianism

- Bayesian models are useful in summarizing and organizing the data
- But they are also useful in guiding a heuristic search for mechanisms!
- Useful framework to generate testable hypotheses, e.g.
  - Error learning models of category learning
  - Probabilistic population codes
  - Sampling hypothesis
  - Message passing
  - ...
- Theoretically there are many possible algorithms and implementations, but *pragmatically* not so many
  - Use successful ideas from other fields (machine learning, statistics, AI)
  - Use other constraints, like cognitive limitations, known facts about the hardware, etc.

[Zednik and Jäkel, 2014]

References I

Anderson, J. R. (1991a).
The adaptive nature of human categorization.
Psychological Review, 98(3):409–429.

Anderson, J. R. (1991b).
Is human cognition adaptive?
Behavioral and Brain Sciences, 14:471–517.

Berniker, M. and Körding, K. P. (2011).
Bayesian approaches to sensory integration for motor control.
WIREs Cognitive Science, 2:419–428.

Berry, D. A. (1995).
Statistics. A Bayesian Perspective.
Duxbury Resource Center.

Bowers, J. S. and Davis, C. J. (2012).
Bayesian just-so stories in psychology and neuroscience.
Psychological Bulletin, 138(3):389–414.

Chater, N. and Oaksford, M. (1999).
Ten years of the rational analysis of cognition.
Trends in Cognitive Sciences, 3(2).

Chater, N., Tenenbaum, J. B., and Yuille, A. (2006).
Probabilistic models of cognition: Conceptual foundations.
Trends in Cognitive Sciences, 10:287–291.

References II

Cohen, J. (1994).
The earth is round (p<.05).
American Psychologist, 49(12):997–1003.

Gigerenzer, G. (2003).
Reckoning with Risk.
Penguin Books Ltd.

Green, D. M. and Swets, J. A. (1988).
Signal Detection and Psychophysics (Reprint Edition).
Peninsula Publishing.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., and Tenenbaum, J. (2010).
Probabilistic models of cognition: exploring representations and inductive biases.
Trends in Cognitive Sciences, 14(8):357–364.

Hacking, I. (1975).
The Emergence of Probability.
Cambridge University Press.

Jones, M. and Love, B. C. (2011).
Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition.
Behavioral and Brain Sciences, 34:169–231.

Kac, M. (1969).
Some mathematical models in science.
Science, 166(3906):695–699.

References III

Kersten, D. and Schrater, P. R. (2002).
Pattern inference theory: A probabilistic approach to vision.
In Mausfeld, R. and Heyer, D., editors, *Perception and the Physical World*. John Wiley & Sons, Ltd., Chichester.

Kersten, D. and Yuille, A. (2003).
Bayesian models of object perception.
Current Opinion in Neurobiology, 13:150–158.

Kruschke, J. K. (2011).
Doing Bayesian Data Analysis.
Academic Press, Burlington, MA.

Laplace, P.-S. (1902).
A Philosophical Essay on Probabilities.
Wiley, New York.

Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006).
Bayesian inference with probabilistic population codes.
Nature Neuroscience, 9(11):1432–1438.

Marcus, G. F. and Davis, E. (2013).
How robust are probabilistic models of higher-level cognition?
Psychological Science, 24(12):2351–2360.

Marr, D. (2010).
Vision: a computational investigation into the human representation and processing of visual information.
MIT Press, Cambridge, reprint of the 1982 edition.

References IV

Oaksford, M. and Chater, N. (2001).
The probabilistic approach to human reasoning.
Trends in Cognitive Sciences, 5(8).

Swets, J., Tanner, W. P., and Birdsall, T. G. (1961).
Decision processes in perception.
Psychological Review, 68:301–340.

Tenenbaum, J. B., Griffiths, T. L., and Kemp, C. (2006).
Theory-based Bayesian models of inductive learning and reasoning.
Trends in Cognitive Sciences, 10(7):309–318.

Tenenbaum, J. B., Kemp, C., L., G. T., and Goodman, N. D. (2011).
How to grow a mind: Statistics, structure, and abstraction.
Science, 331(6022):1279–1285.

Tversky, A. and Kahneman, D. (1982).
Evidential impact of base rates.
In Kahneman, D., Slovic, P., and Tversky, A., editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 153–160. Cambridge University Press.

Zednik, C. and Jäkel, F. (2014).
How does Bayesian reverse-engineering work?
In Bello, P., Guarini, M., McShane, M., and Scasselati, B., editors, *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, pages 666–671.