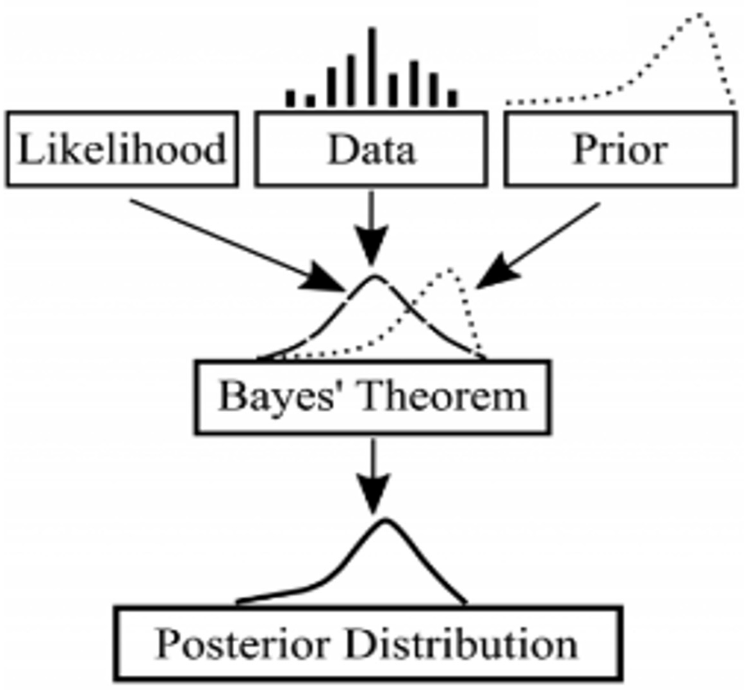
Outline

P.s. many of our slides were adapted from Cody's slides :) Go Cody!

1. What IS Bayesian statistics: a new framework of making statistical inference! (Ham)
2. An example of Bayes Magic (Alex)
3. When to use Bayesian statistics for psychological research? (Ham)



Bayesian statistics is fundamentally inference~

What is statistical inference?

Try to figure out the true quantity of interest using samples

- True quantity: 0.7 head 0.3 tail
  - We don't know this!!! But we want to
  - We know (or rather assume) the kind of probability distribution
- Sample: hhhtthhtthh
  - 6/10 heads
  - 4/10 tails
- From the sample, how can we figure out the true quantity?
  - We cannot guarantee we know the true quantity but we can justify for a best estimate

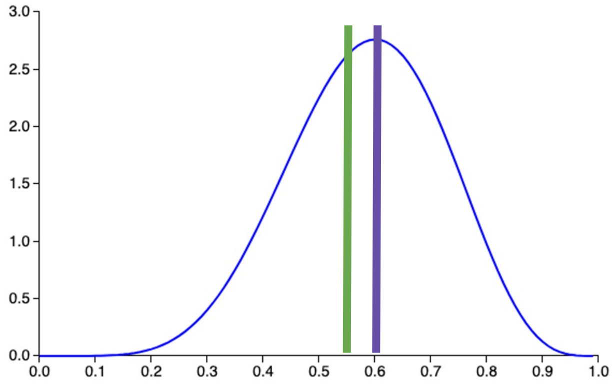
Frequentist vs Bayesian

Frequentist (MLE): **All eggs in one basket**

- Calculate P(6 heads | p_h) for all values of p_h from 0 to 1
  - P(6 heads | p_h = 0.5) = 0.20508
  - P(6 heads | p_h = 0.6) = 0.25082
  - P(6 heads | p_h = 0.8) = 0.08808
- Math tells us when p_h = 0.6, P(6 heads | p_h) has the largest value (Duh!)
- Therefore if this sample is all we got, we should infer that the true quantity is 0.6!

Bayesian: **Probability of true quantity**

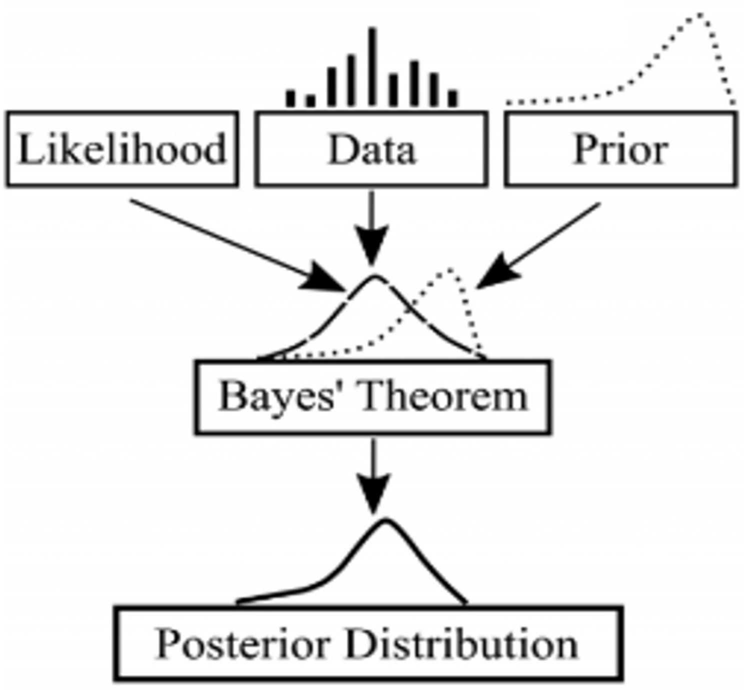
- Objective: find out P(p_h | 6 heads) for all p_h.
- P(6 heads | p_h) -> magic of Bayes rule



Outline

P.s. many of our slides were adapted from Cody's slides :) Go Cody!

1. What IS Bayesian statistics: a new framework of making statistical inference! (Ham)
2. An example of Bayes Magic (Alex)
3. When to use Bayesian statistics for psychological research? (Ham)



Bayes' rule (exciting version)

Posterior probability

Likelihood

Prior probability

$$P(h \mid d) = \frac{P(d \mid h)P(h)}{\sum_{h' \in H} P(d \mid h')P(h')}$$

Sum over space of hypotheses

h: hypothesis

d: data

Slides by Tom Griffiths

Comparing two simple hypotheses

- Contrast simple hypotheses:
  - h_1: "fair coin", $P(\mathrm{H}) = 0.5$
  - h_2: "always heads", $P(\mathrm{H}) = 1.0$
- Bayes' rule:

$$P(h \mid d) = \frac{P(d \mid h)P(h)}{\sum_{h' \in H} P(d \mid h')P(h')}$$

- With two hypotheses, use odds form

Bayes' rule in odds form

$$\frac{P(h_1 \mid d)}{P(h_2 \mid d)} = \frac{P(d \mid h_1)}{P(d \mid h_2)} \frac{P(h_1)}{P(h_2)}$$

d:	data
h_1, h_2:	models
$P(h_1 \mid d)$:	posterior probability h_1 generated d
$P(d \mid h_1)$:	likelihood of data d under model h_1
$P(h_1)$:	prior probability of h_1

Comparing two simple hypotheses

$$\frac{P(h_1 \mid d)}{P(h_2 \mid d)} = \frac{P(d \mid h_1)}{P(d \mid h_2)} \frac{P(h_1)}{P(h_2)}$$

d:	HHTHT		
h_1, h_2:	"fair coin", "always heads"		
$P(d\|h_1) =$	$1/2^5$	$P(h_1) =$	$999/1000$
$P(d\|h_2) =$	0	$P(h_2) =$	$1/1000$

$P(h_1|d) \,/\, P(h_2|d) =$ infinity

Comparing two simple hypotheses

$$\frac{P(h_1 \mid d)}{P(h_2 \mid d)} = \frac{P(d \mid h_1)}{P(d \mid h_2)} \frac{P(h_1)}{P(h_2)}$$

d:	HHHHH		
h_1, h_2:	"fair coin", "always heads"		
$P(d \mid h_1) =$	$1/2^5$	$P(h_1) =$	$999/1000$
$P(d \mid h_2) =$	1	$P(h_2) =$	$1/1000$

$$P(h_1 \mid d) \,/\, P(h_2 \mid d) \approx 30$$

Comparing two simple hypotheses

$$\frac{P(h_1 \mid d)}{P(h_2 \mid d)} = \frac{P(d \mid h_1)}{P(d \mid h_2)} \frac{P(h_1)}{P(h_2)}$$

d: HHHHHHHHHH

h_1, h_2: "fair coin", "always heads"

$P(d|h_1) =$ $1/2^{10}$ $P(h_1) =$ $999/1000$

$P(d|h_2) =$ 1 $P(h_2) =$ $1/1000$

$$P(h_1|d) \,/\, P(h_2|d) \approx 1$$

The blue cab / green cab problem

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. 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 knowing that this witness identified it as Blue?

(Tversky & Kahneman, 1982)

Comparing two simple hypotheses

$$\frac{P(h_1 \mid d)}{P(h_2 \mid d)} = \frac{P(d \mid h_1)}{P(d \mid h_2)} \frac{P(h_1)}{P(h_2)}$$

d:	Blue		
h_1, h_2:	Blue, Green		
$P(d\|h_1) =$	0.8	$P(h_1) =$	0.15
$P(d\|h_2) =$	0.2	$P(h_2) =$	0.85

$P(h_1|d) \,/\, P(h_2|d) \approx 0.70 \quad P(h_1|d) = 0.41$

Bayesian statistics

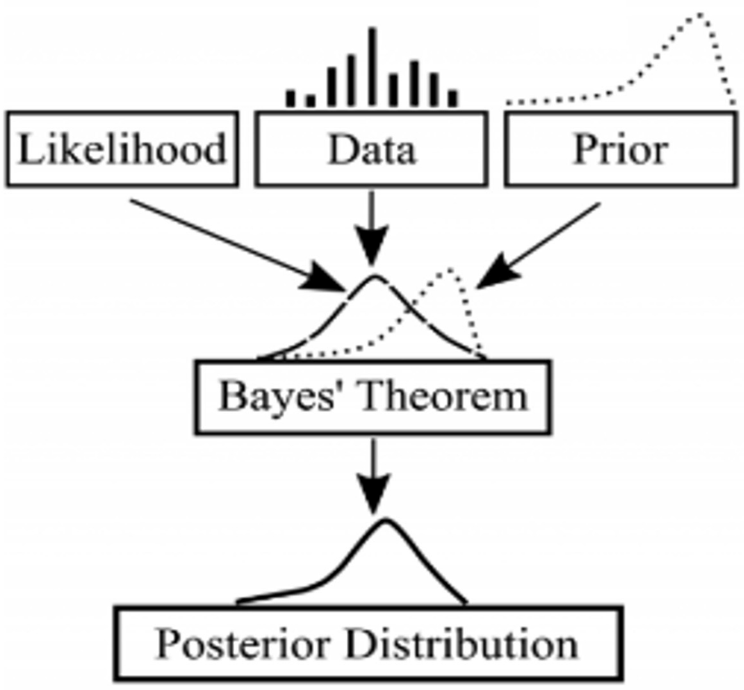
- Applying Bayes' rule with two hypotheses is fairly common in psychology
  - usually to show people don't seem to use it
- Extends to a finite number of hypotheses

- Much of Bayesian statistics involves working with infinitely many hypotheses

Outline

P.s. many of our slides were adapted from Cody's slides :) Go Cody!

1. What IS Bayesian statistics: a new framework of making statistical inference! (Ham)
2. An example of Bayes Magic (Alex)
3. When to use Bayesian statistics for psychological research? (Ham)



Pros and Cons

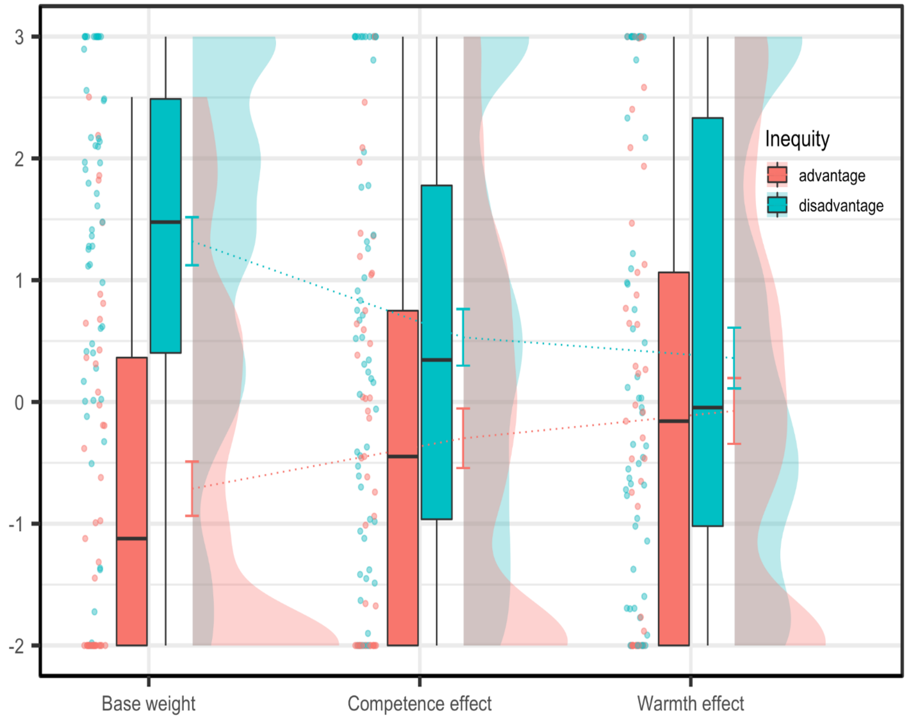
Cons: In practice, you will not be able to solve for posterior but need to approximate, meaning:

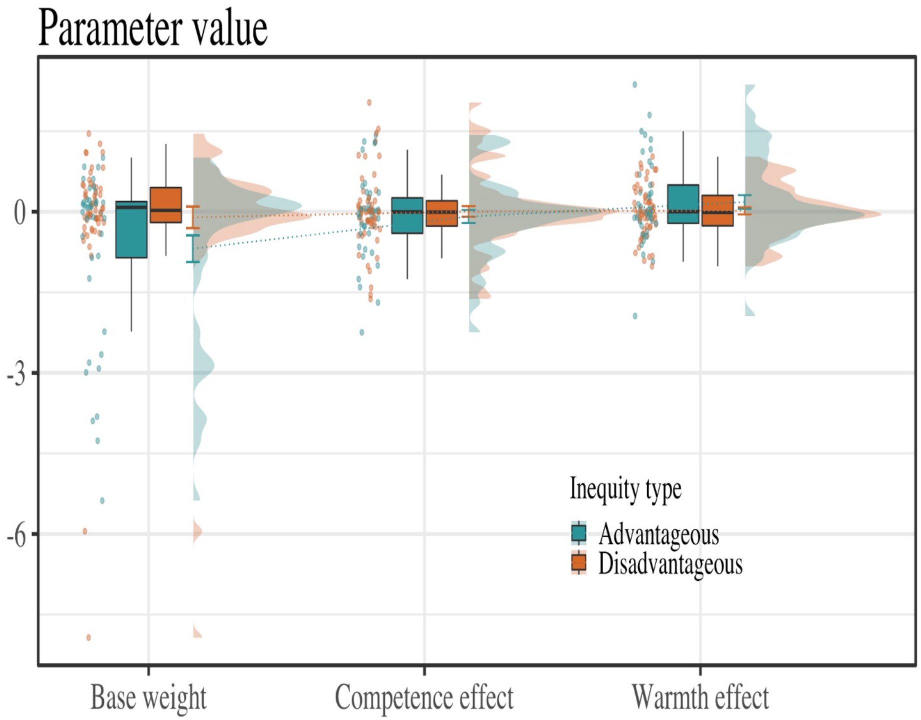
1. **Computationally intensive**!!
   a. Mostly likely you can find the equation for P(H|D) but need to approximate it
2. **Annoying Model Checks!**
   a. Many things to check to make sure your model estimation is interpretable
   b. Endless loop of tuning the prior to make sure no divergence
3. **Requires prior**:
   a. Solution: choose uninformative prior

Pros:

1. **Give you more information:**
   a. Allow direct probabilistic statements
   b. Capture uncertainty of estimates
2. **Better philosophy of statistics**:
   a. According to some people
   b. Makes prior choice explicit
   c. Doesn’t need central limit theorem
3. **Regularization**!!!!!
   a. Give you reliable stable estimates for complicated models
4. **Don’t need Null hypothesis**

Regularization is a big pro!!





They are related!

Bernstein-von Mises Theorem

- Basically (mathematicians will kill me for saying this), that **under large sample**, **Bayesian = Frequentist**
- For our purposes, Bayesian and Frequentist are **both correct**! But …

Ham's personal recommendation for when to go Bayesian

- **Don't fix if it's not broken**:
  - If whatever regression you are doing works fine with the usual R packages (linear regression, original regression, logistics regression, etc), just don't worry about going Bayesian. You could if you want and easy to do.
- **Go Bayesian if it is broken**:
  - Parameter estimate too noisy to be interpretable (need regularization)
  - Almost always use uninformative prior. If not possible, be as uninformative as possible.