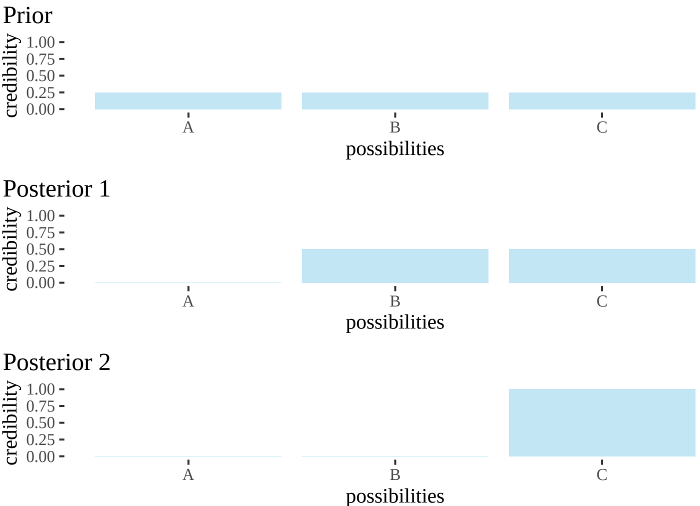
Intro to Bayesian Thinking

Rafał Urbaniak, Nikodem Lewandowski
(LoPSE research group, University of Gdansk)

Sherlock's naivete

A rather unhelpful piece of advice

"...when you have eliminated the impossible, whatever remains, however, improbable, must be the truth."



Trouble in paradise

- Data have only probabilistic relations to hypotheses

  Many people may have similar footprints

- Measurements only probabilistically narrow down the range

  We mathematically can describe the footprints up to some level of precision

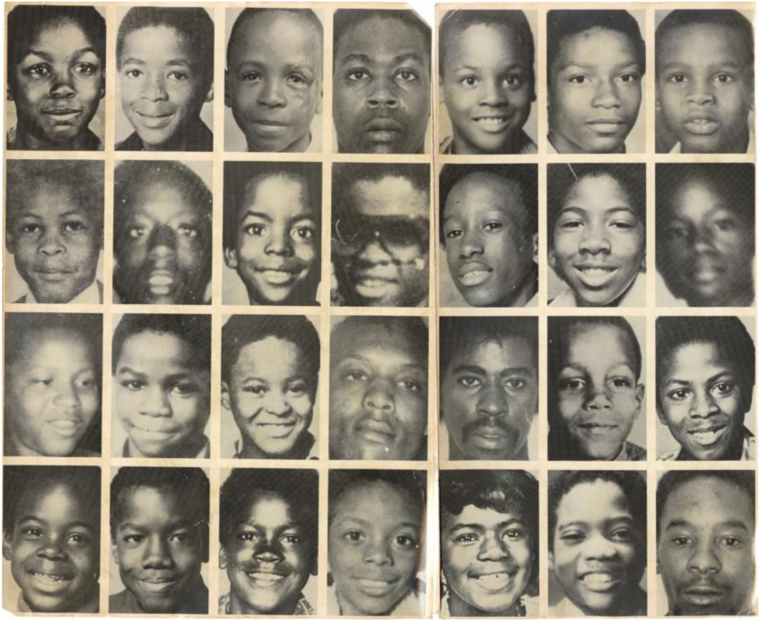
- Association does not directly translate into causation

  There may be various confounding factors explaining why people who received a given drug have lower blood pressure

- There often is natural variation

  The weight of a newborn baby may vary naturally due to genetics and environmental factors, rather than a specific cause

Wayne Williams case

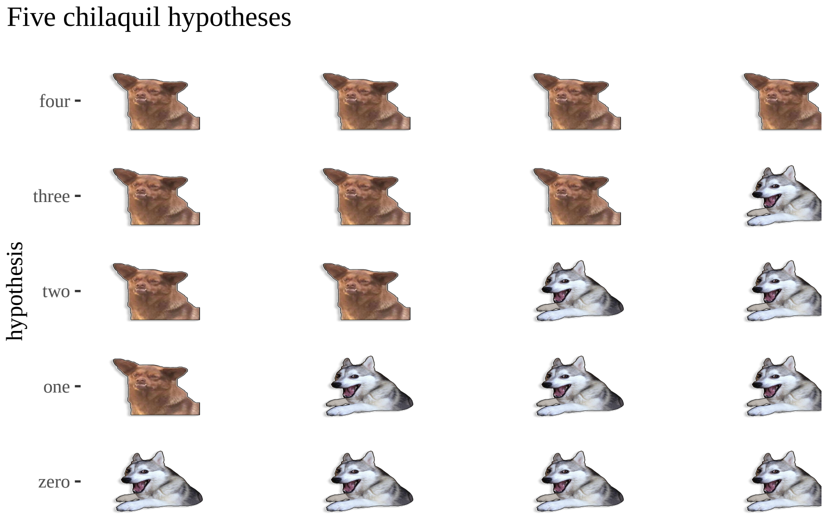


Two items of evidence

- Dog hair evidence, random match probability (RMP) is about 0.0256.
- Human hair evidence, RMP is about 0.0253

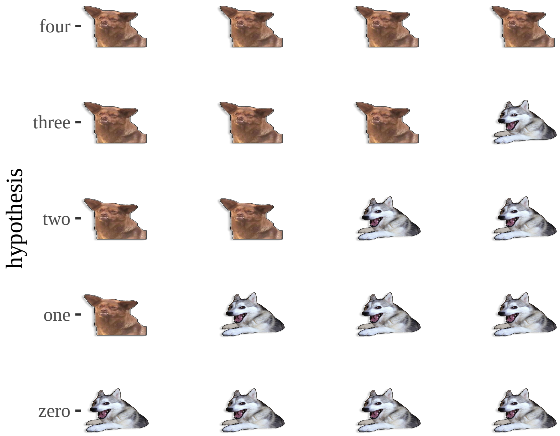
Questions that come to mind?

Let's focus on dog fur



Ways dogs could be (likelihoods)

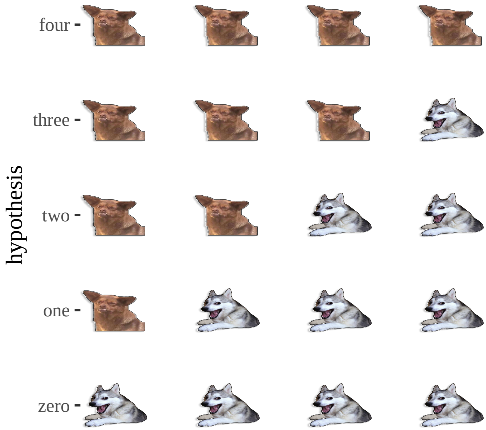
Ways to observe (h,c,h)



hypothesis	h	c	h	(h,c,h)
four	0	4	0	0
three	1	3	1	3
two	2	2	2	8
one	3	1	3	9
zero	4	0	4	0

Updating with new observations

Ways to observe (h,c,h)



	h	c	h	(h,c,h)	h	(h,c,h,h)
four	0	4	0	0	0	0
three	1	3	1	3	1	3
two	2	2	2	8	2	16
one	3	1	3	9	3	27
zero	4	0	4	0	4	0

Now with probabilities

p	ways0	ways0pr	ways1	ways1pr
0.00	0	0.00	0	0.0000000
0.25	3	0.15	3	0.0652174
0.50	8	0.40	16	0.3478261
0.75	9	0.45	27	0.5869565
1.00	0	0.00	0	0.0000000

More dogs, Bayesian style!

$$\mathrm{P}(C = c, H = h|\theta) = \frac{(c+h)!}{c!h!}\theta^c(1-\theta)^h$$

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

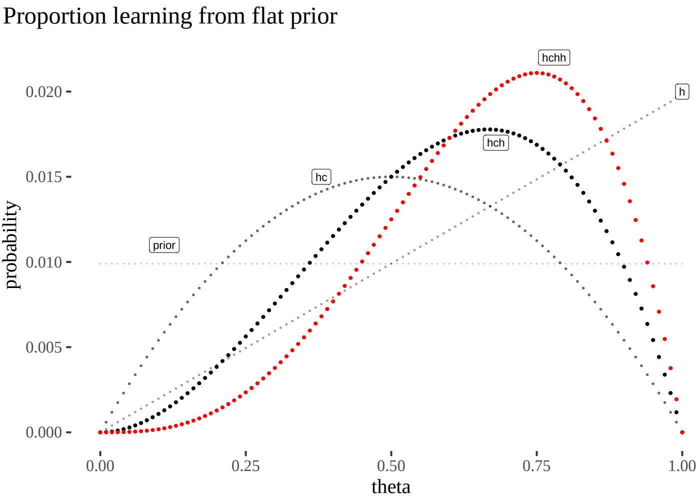
$$\begin{aligned}
H &\sim Binomial(N, \theta) \\
\theta &\sim Uniform(0, 1) \\
P(c, h, \theta) &= P(c, h|\theta)P(\theta) \\
P(c, h, \theta) &= P(\theta|c, h)P(c, h) \\
P(\theta|c, h)P(c, h) &= P(c, h|\theta)P(\theta)
\end{aligned}$$

$$\underbrace{P(\theta|c, h)}_{posterior} = \frac{\overbrace{P(c, h|\theta)}^{likelihood}\overbrace{P(\theta)}^{prior}}{\underbrace{P(c, h)}_{(average)data}}$$

The underlying mechanism

plausibility(hypothesis n|data) $\propto$

ways hypothesis n can produce data $\times$ prior plausibility of hypothesis n



Back to the fur testimony (grid approximation)

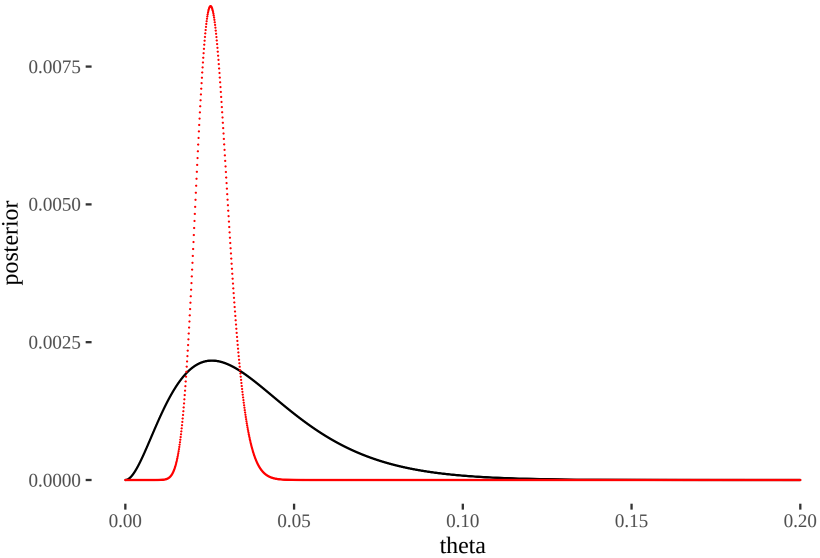
```r
theta <- seq(0,1, length.out = 10001)
prior <- rep(1/10001,10001)

likelihoodDog <- dbinom(2,78, theta)
likelihoodHuman <- dbinom(29,1148, theta)

posteriorDogUnst <-  likelihoodDog * prior
posteriorHumanUnst <-  likelihoodHuman * prior

posteriorDog <- posteriorDogUnst/sum(posteriorDogUnst)
posteriorHuman <- posteriorHumanUnst/sum(posteriorHumanUnst)
```

Back to the fur testimony (grid approximation)



Steps of Bayesian data analysis

1. Identify the data, variables, predictors
2. Define a descriptive and appropriate model
3. Specify a prior distribution (over parameters)
4. Use Bayesian inference to reallocate credibility in light of the training data
5. Test whether the posterior predictions are reasonable as compared to validation data

Build your first model!

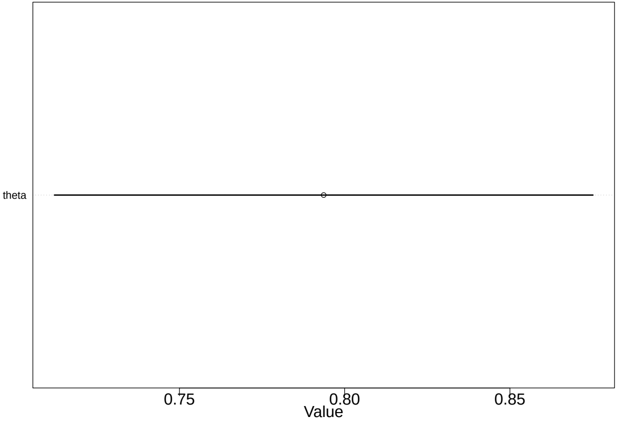
```r
dogsModel <- quap(
  alist(
    h ~ dbinom( h + c , theta),
    theta ~ dunif(0,1)
  ) ,
  data=list(h=50,c=13) )
```

Build your first model!

```
precis(dogsModel)
```

```
##              mean         sd      5.5%     94.5%
## theta 0.7936496 0.05098465 0.7121663 0.8751329
```

```
par(cex.axis=1.5, cex.lab=1.5)
plot(precis(dogsModel))
```



Liar detectors

The task

Out of 100 suspects, 10 are innocent
Polygraph sensitivity ($P(+|T)$) and specificity ($P(-|F)$) are 70%
A suspect is polygraph-positive
So what?

Liar detectors

The task

Out of 100 suspects, 10 are innocent
Polygraph sensitivity ($P(+|T)$) and specificity ($P(-|F)$) are 70%
A suspect is polygraph-positive
So what?

Population considerations

- Out of 10 000 suspects, 1000 will be guilty, 9 000 will not
- Out of 1000 guilty, 700 will be positive, out of 9 000 innocent, 2700
- So out of 2700+700 positive, 700 will be guilty. That's around 20.5%.

Liar detectors

```
pos_if_g  = .7
pos_if_ng = .3
g = .1

pos = ( pos_if_g * g + pos_if_ng * (1-g) )

g_if_pos = ( pos_if_g * g ) / pos

g_if_pos
```

```
## [1] 0.2058824
```

Signal detection and why data can't save us

Simplified structure of the goal of science

- some binary state is hidden
- we observe imperfect hints
- we use Bayes to learn

Simplified assumptions

- sensitivity is .95
- false positive rate is .05
- base rate: most hypotheses are false, with pr $=$ .01

A simplified observation

The posterior is only .16.