Bayesian Statistics

Lecture 6

Xiaofei Shi

Learning objective

- Have a basic understanding of Bayesian statistics
- How to define a prior

Bayesian statistics updates belief with data

Belief?

Update?

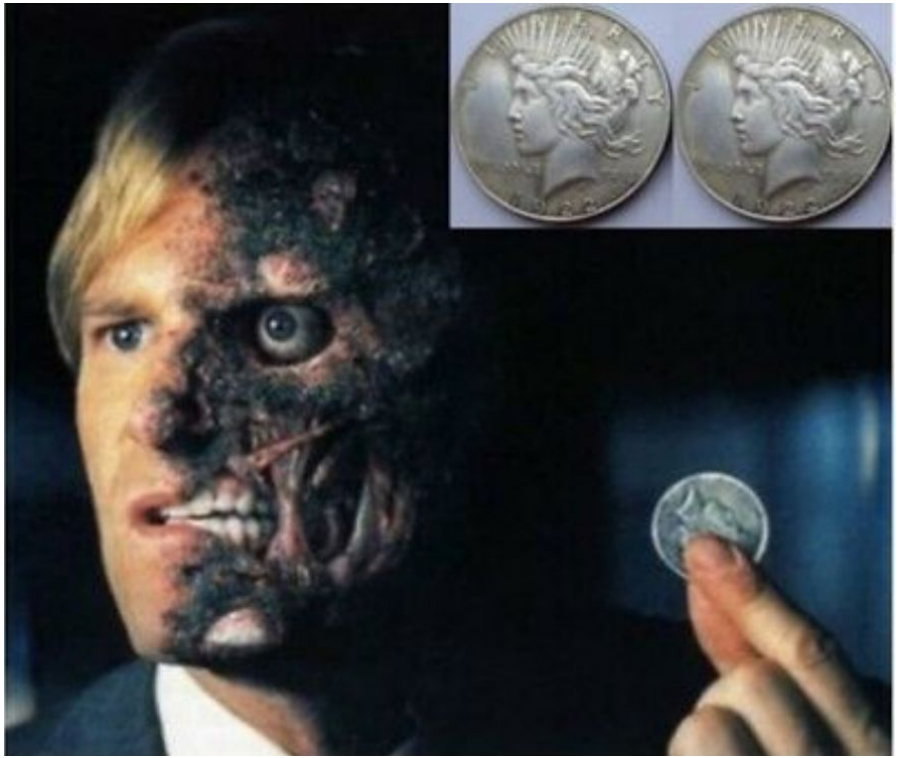
Thought experiment 0

Imagine 2 coins:

- Biased coin - both sides are Heads
- Fair coin

Someone chooses one coin and tosses it.

What's the chance it's the fair coin?



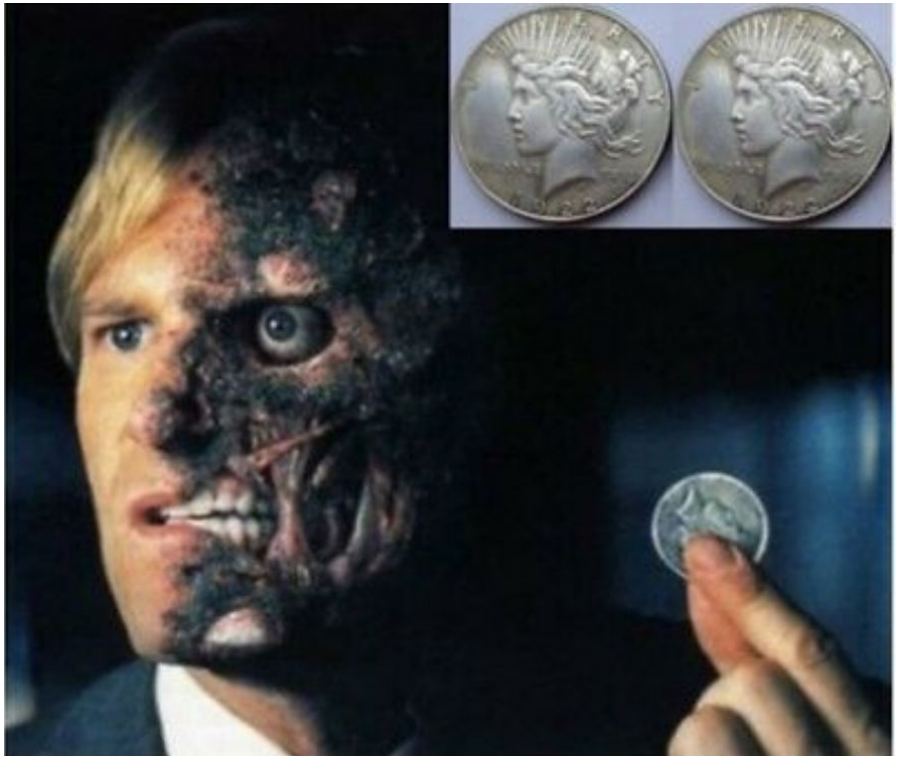
Thought experiment 1

Imagine 2 coins:

- Biased coin - both sides are Heads
- Fair coin

Someone randomly chooses one coin and tosses it.

What’s the chance it’s the fair coin?



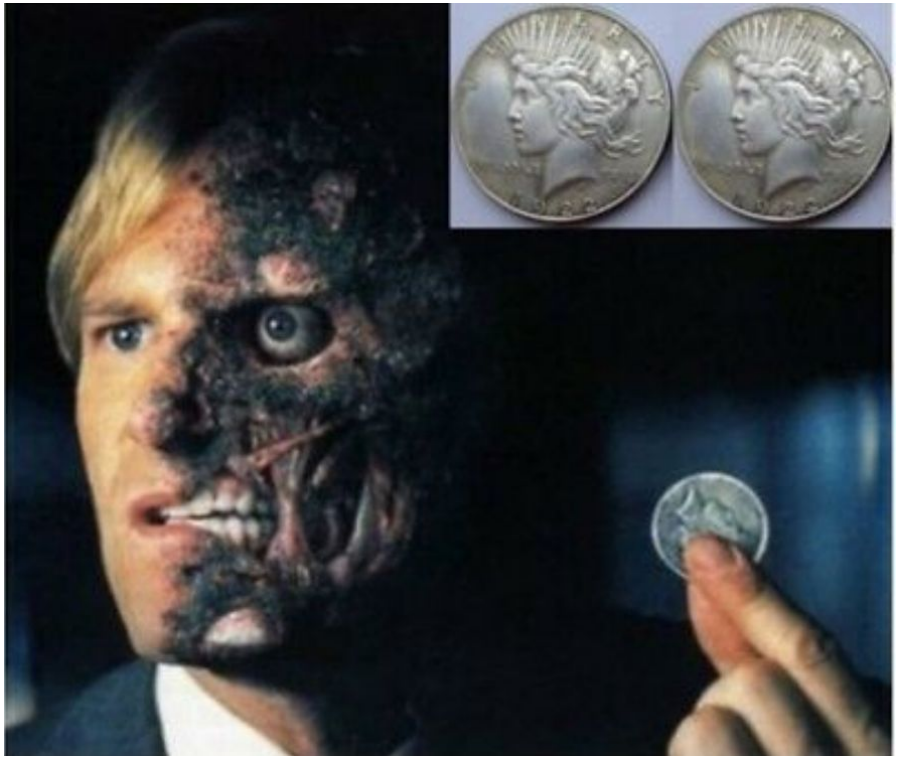
Thought experiment 2

Imagine 2 coins:

- Biased coin - both sides are Heads
- Fair coin

Someone randomly chooses one coin and tosses it.

You see a **Tail**, what's the chance it's the fair coin?



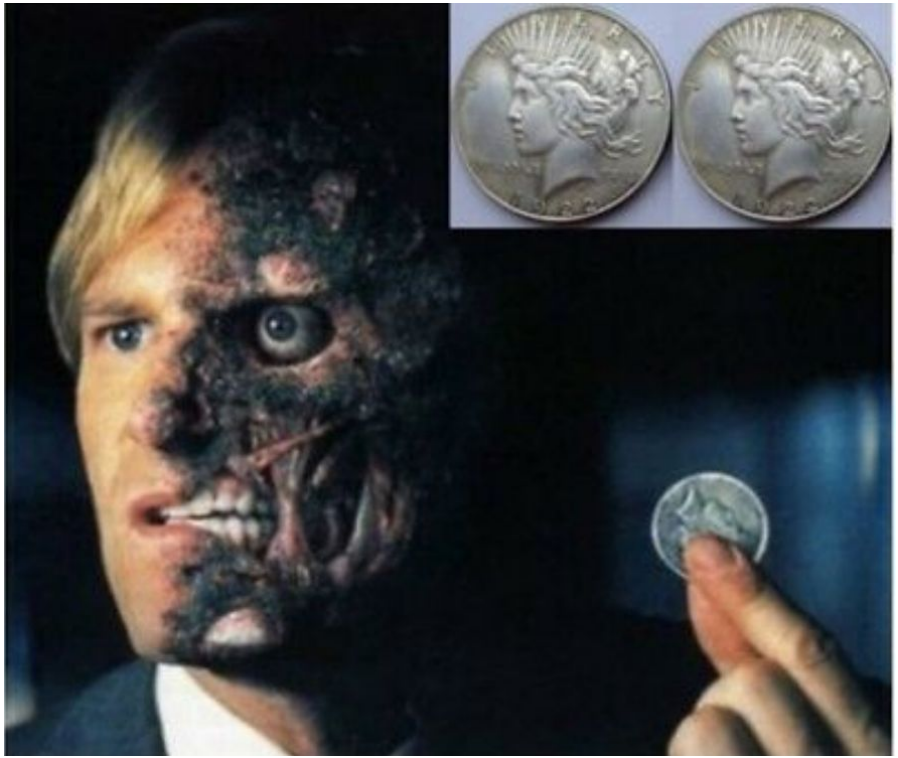
Thought experiment 3

Imagine 2 coins:

- Biased coin - both sides are Heads
- Fair coin

Someone randomly chooses one coin and tosses it.

You see a **Head**, what's the chance it's the fair coin?



Setting the problem up - what do we know

Setting the problem up - what do we know

$$P(\text{toss=Heads}|\text{coin is fair}) = 1/2$$
$$P(\text{toss=Heads}|\text{coin is biased}) = 1$$
$$P(\text{coin in biased}) = P(\text{coin in biased}) = 1/2$$

Setting the problem up - what are we solving?

$$P(\text{coin is fair}|\text{toss=Heads}) = ?$$
$$P(\text{toss=Heads}|\text{coin is fair}) = 1/2$$
$$P(\text{toss=Heads}|\text{coin is biased}) = 1$$
$$P(\text{coin in biased}) = P(\text{coin in biased}) = 1/2$$

Recalling conditional probability

$$P(\text{coin is fair}|\text{toss=Heads}) = \frac{P(\text{coin is fair AND toss=Heads})}{P(\text{toss=Heads})}$$

$$= \frac{P(\text{coin is fair})P(\text{toss=Heads}|\text{coin is fair})}{P(\text{toss=Heads})}$$

$$P(\text{toss=Heads}) = 0.5 * 0.5 + 0.5 * 1 = 0.75$$

General setup for Bayesian problem - choosing the likelihood and prior

Posterior: Given data, what's my belief about X?

- P(X | Data)

General setup for Bayesian problem - choosing the likelihood and prior

Posterior: Given data, what's my belief about X?

- P(X | Data)

Through conditional probability, this can be expressed as a function of:

Likelihood: for any specific X, what is the distribution of the data?

- P(Data | X)

General setup for Bayesian problem - choosing the likelihood and prior

Posterior: Given data, what's my belief about X?

- P(X | Data)

Through conditional probability, this can be expressed as a function of:

Likelihood: for any specific X, what is the distribution of the data?

- P(Data | X)

Prior: before seeing any data, what is my belief about X?

- P(X)

Prior is a special quantity in Bayesian Statistics

Posterior: Given data, what's my belief about X?

- P(X | Data)

Through conditional probability, this can be expressed as a function of:

Likelihood: for any specific X, what is the distribution of the data?

- P(Data | X)

Prior: before seeing any data, what is my belief about X?

- P(X)

What's the big deal?

Imagine if the 2 different coins were 2 different hypotheses?

Practice - setup the Bayesian problem

You tossed a coin 1 time, it's shows up TAILS, what is its chance to land a H?

- What is the problem?
- What is the likelihood?
- What is the prior?

See notebook on Canvas "bayesian-intro.ipynb"!

How else could we estimate this?

You tossed a coin 1 time, it’s shows up TAILS, what is its chance to land a H?

What chance would maximize the probability of seeing a TAILS from 1 toss?

Pro/Con of Bayesian Methods

Con

- Prior requires careful thought
- Can be computationally difficult
- Can have an answer with no data

Pro/Con of Bayesian Methods

Con

- Prior requires careful thought
- Can be computationally difficult
- Can have an answer with no data

Pro:

- Very flexible and intuitive (e.g. what we perceive as chance)
- Can introduce prior knowledge easily
- Can update as you see more data
- Very comprehensive framework

Notes on “Bayes” in machine learning

- Bayesian statistics is often used as a framework for motivating certain calculations
- Bayesian statistics can “regularize” some classic statistics