Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

> *"If we were not ignorant there would be no probability, there could only be certainty. But our ignorance cannot be absolute, for then there would be no longer any probability at all. Thus the problems of probability may be classed according to the greater or less depth of our ignorance."*
> (Poincaré)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

The state of information can be different from subject to subject

$\Rightarrow$ intrinsic subjective nature.

- No negative meaning: only an acknowledgment that several persons might have different information and, therefore, necessarily different opinions.

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

The state of information can be different from subject to subject

$\Rightarrow$ intrinsic subjective nature.

- No negative meaning: only an acknowledgment that several persons might have different information and, therefore, necessarily different opinions.
- *"Since the knowledge may be different with different persons or with the same person at different times, they may anticipate the same event with more or less confidence, and thus different numerical probabilities may be attached to the same event"* (Schrödinger)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$`P(E)' \;\longrightarrow\; P(E\,|\,I) \;\longrightarrow\; P(E\,|\,I(t))$$

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$\text{`}P(E)\text{'} \;\longrightarrow\; P(E\,|\,I) \;\longrightarrow\; P(E\,|\,I(t))$$

- *"Thus whenever we speak loosely of 'the probability of an event,' it is always to be understood: probability with regard to a certain given state of knowledge*" (Schrödinger)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$`P(E)' \;\longrightarrow\; P(E\,|\,I) \;\longrightarrow\; P(E\,|\,I(t))$$

- *“Thus whenever we speak loosely of ‘the probability of an event,’ it is always to be understood: probability with regard to a certain given state of knowledge”* (Schrödinger)
- Some examples:
  - tossing a die;
  - ’three box problems’;
  - two envelops’ paradox.

Unifying role of subjective probability

- Wide range of applicability

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain next Saturday}) = 68\%$
  - $P(\text{Inter will win Italian champion league}) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain next Saturday}) = 68\%$
  - $P(\text{Inter will win Italian champion league}) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the <u>same confidence</u> on something.

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain next Saturday}) = 68\%$
  - $P(\text{Inter will win Italian champion league}) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the <u>same confidence</u> on something.
- You might agree or disagree, but at least You know what this person has in his mind. (<u>NOT TRUE with “C.L.’s”!</u>)

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
    - $P(\text{rain next Saturday}) = 68\%$
    - $P(\text{Inter will win Italian champion league}) = 68\%$
    - $P(\text{free neutron decays before 17 s}) = 68\%$
    - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the same confidence on something.
- You might agree or disagree, but at least You know what this person has in his mind. (NOT TRUE with "C.L.'s"!)
- If a person has these beliefs and he/she has the chance to win a rich prize bound to one of these events, he/she has no rational reason to chose an event instead than the others.

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain next Saturday}) = 68\%$
  - $P(\text{Inter will win Italian champion league}) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$
- Probability not bound to a single evaluation rule.

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain next Saturday}) = 68\%$
  - $P(\text{Inter will win Italian champion league}) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

- Probability not bound to a single evaluation rule.
- In particular, combinatorial and frequency based 'definitions' are easily recovered as evaluation rules under well defined hypotheses.
- Keep separate concept from evaluation rule.

From the concept of probability to the probability theory

Ok, it looks nice, . . . but "how do we deal with 'numbers'?"

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)

From the concept of probability to the probability theory

- **Formal structure**: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- **Assess probability**: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

Coherent bet (de Finetti, Ramsey - 'Dutch book argument')

It is well understood that bet odds can express confidence†

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

  Coherent bet $\longrightarrow$ A bet acceptable in both directions:
  - You state your confidence fixing the bet odds
  - ...but somebody else chooses the direction of the bet
  - best way to honestly assess beliefs.

  $\longrightarrow$ see later for details, examples, objections, etc

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

Consistency arguments (Cox, + Good, Lucas)

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

  Consistency arguments (Cox, + Good, Lucas)
- Similar approach by Schrödinger (much less known)
- Supported by Jaynes and Maximum Entropy school

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

Consistency arguments (Cox, + Good, Lucas)

- Similar approach by Schrödinger (much less known)
- Supported by Jaynes and Maximum Entropy school

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

Lindley’s ‘calibration’ against ‘standards’

$\longrightarrow$ analogy to measures (we need to measure ’befiefs’)

From the concept of probability to the probability theory

- **Formal structure**: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- **Assess probability**: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

Lindley’s ‘calibration’ against ‘standards’

$\longrightarrow$ analogy to measures (we need to measure ’befiefs’)

$\Rightarrow$ reference probabilities provided by simple cases in which equiprobability applies (coins, dice, turning wheels,. . .).

- **Example**: You are offered to options to receive a price: a) if E happens, b) if a coin will show head. Etc. . . .

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

  Lindley's 'calibration' against 'standards'

$\longrightarrow$ Rational under everedays expressions like "there are 90 possibilities in 100" to state beliefs in situations in which the real possibilities are indeed only 2 (e.g. dead or alive)

- Example: a question to a student that has to pass an exam: a) normal test; b) pass it is a uniform random x will be ≤ 0.8.

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box, in which we have to insert some numbers.
  - Is there a very general rule?

  Lindley’s ‘calibration’ against ‘standards’
- Also based on coherence, but it avoids the ‘repulsion’ of several person when they are asked to think directly in terms of bet (it is proved that many persons have reluctance to bet money).

Basic rules of probability

They all lead to

$$
\begin{array}{ll}
1. & 0 \leq P(A) \leq 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,] \\
 & \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

where

- Ω stands for 'tautology' (a proposition that is certainly true $\rightarrow$ referring to an event that is certainly true) and $\emptyset = \overline{\Omega}$.
- $A \cap B$ is true only when both A and B are true (logical AND)

  (shorthands 'A, B' or $A\,B$ often used $\rightarrow$ logical product)
- $A \cup B$ is true when at least one of the two propositions is true (logical OR)

Basic rules of probability

Remember that probability is always conditional probability!

1. $0 \leq P(A \,|\, I) \leq 1$
2. $P(\Omega \,|\, I) = 1$
3. $P(A \cup B \,|\, I) = P(A \,|\, I) + P(B \,|\, I)$ $\quad$ [if $P(A \cap B \,|\, I) = \emptyset$]
4. $P(A \cap B \,|\, I) = P(A \,|\, B, I) \cdot P(B \,|\, I) = P(B \,|\, A, I) \cdot P(A \,|\, I)$

I is the background condition (related to information I)

$\longrightarrow$ usually implicit (we only care on ‘re-conditioning’)

Subjective $\neq$ arbitrary

Crucial role of the coherent bet

- You claim that <u>this</u> coin has 70% to give head?
  No problem with me: you place 70€ on head, I 30€ on tail and who wins take 100€.
  $\Rightarrow$ If OK with you, let's start.

Subjective $\neq$ arbitrary

Crucial role of the coherent bet

- You claim that <u>this</u> coin has 70% to give head?
  No problem with me: you place 70€ on head, I 30€ on tail and who wins take 100€.
  $\Rightarrow$ If OK with you, let's start.
- You claim that <u>this</u> coin has 30% to give head?
  $\Rightarrow$ Just reverse the bet

  (Like sharing goods, e.g. a cake with a child)

Subjective $\neq$ arbitrary

Crucial role of the coherent bet

- You claim that this coin has 70% to give head?
  No problem with me: you place 70€ on head, I 30€ on tail and who wins take 100€.
  $\Rightarrow$ If OK with you, let's start.
- You claim that this coin has 30% to give head?
  $\Rightarrow$ Just reverse the bet

  (Like sharing goods, e.g. a cake with a child)

$\Rightarrow$ Take into account all available information *in the most 'objective way'*
(Even that someone has a different opinion!)

$\Rightarrow$ It might seem paradoxically, but the 'subjectivist' is much more 'objective' than those who blindly use so-called objective methods.

Summary on probabilistic approach

- Probability means how much we believe something
- Probability values obey the following basic rules

$$
\begin{array}{ll}
1. & 0 \le P(A) \le 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \quad [\,\text{if } P(A \cap B) = \emptyset\,] \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

- All the rest by logic

$\longrightarrow$ And, please, be coherent!

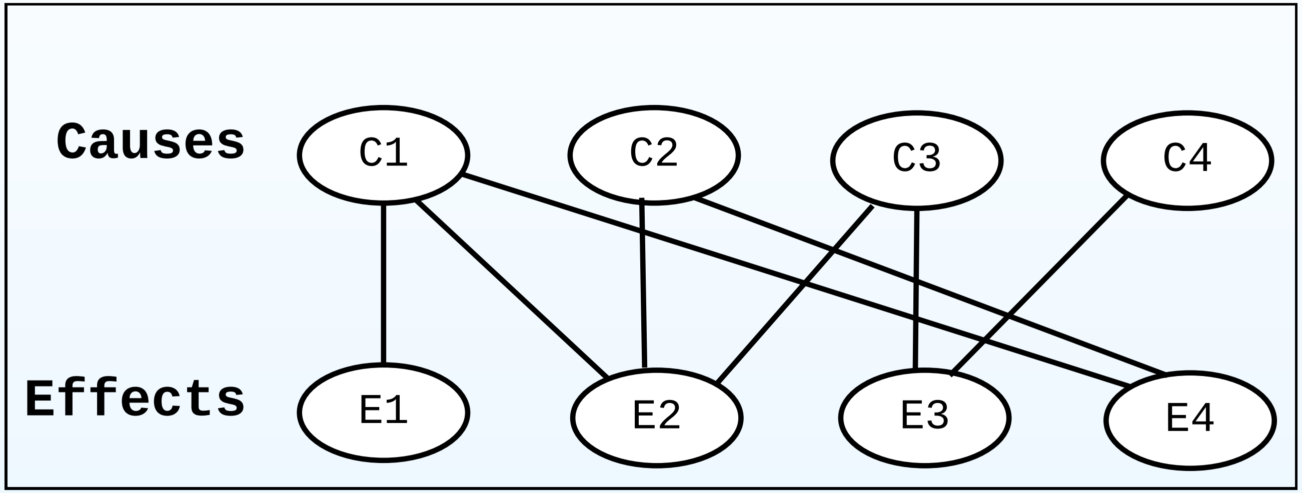
Inference

Inference

$\Rightarrow$ How do we learn from data in a probabilistic framework?

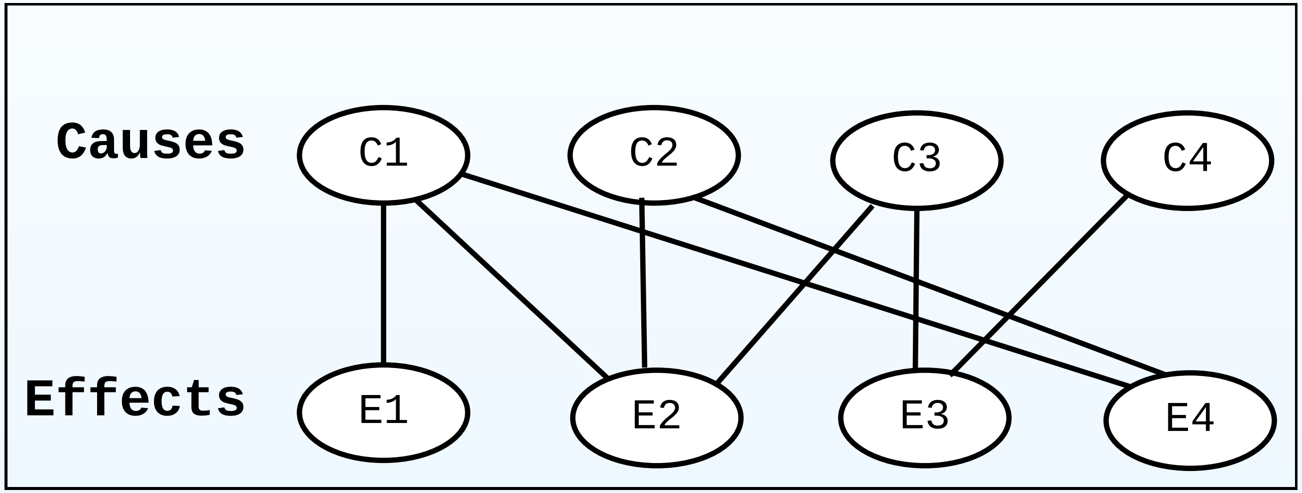
From causes to effects and back

Our original problem:



From causes to effects and back

Our original problem:

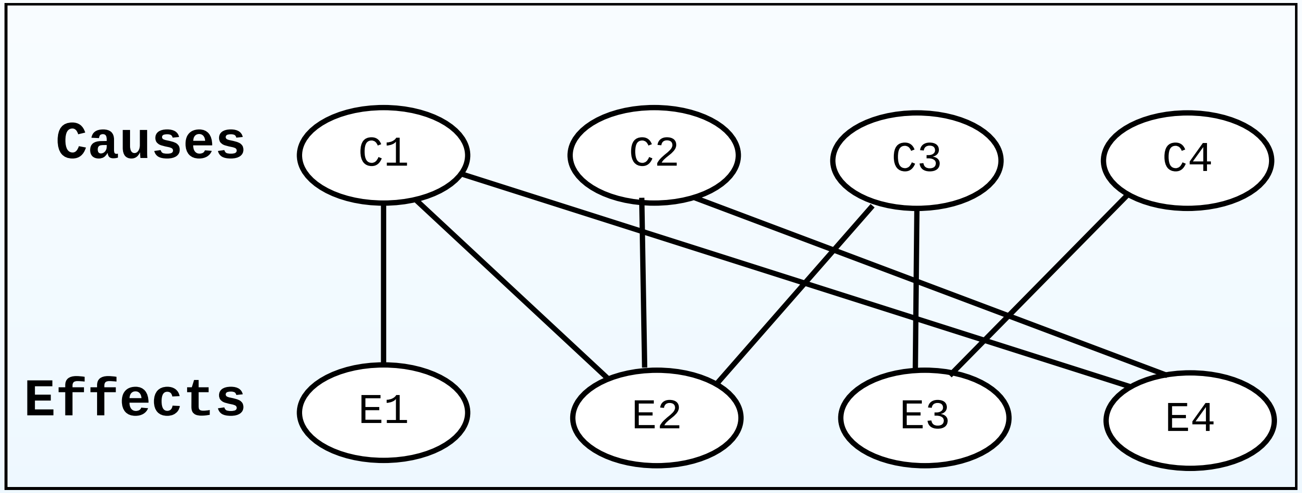


Our conditional view of probabilistic causation

$$\boxed{P(E_i \,|\, C_j)}$$

From causes to effects and back

Our original problem:



Our conditional view of probabilistic causation

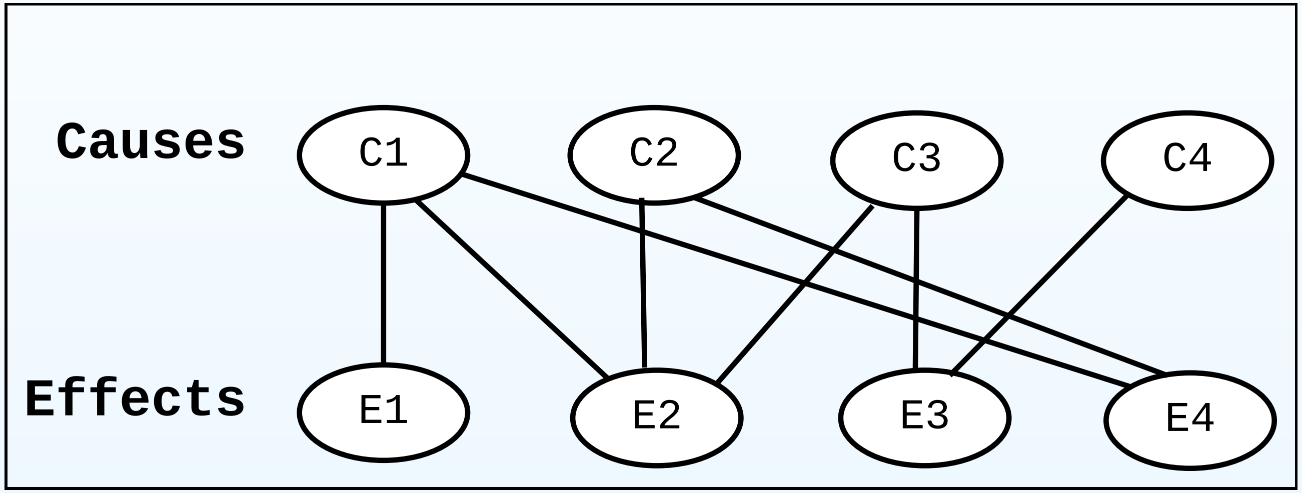
$$\boxed{P(E_i \,|\, C_j)}$$

Our conditional view of probabilistic inference

$$\boxed{P(C_j \,|\, E_i)}$$

From causes to effects and back

Our original problem:



Our conditional view of probabilistic causation

$$\boxed{P(E_i \,|\, C_j)}$$

Our conditional view of probabilistic inference

$$\boxed{P(C_j \,|\, E_i)}$$

The fourth basic rule of probability:

$$\boxed{P(C_j, E_i) = P(E_i \,|\, C_j)\, P(C_j) = P(C_j \,|\, E_i)\, P(E_i)}$$

Symmetric conditioning

Let us take basic rule 4, written in terms of hypotheses H_j and effects E_i, and rewrite it this way:

$$\frac{P(H_j \,|\, E_i)}{P(H_j)} = \frac{P(E_i \,|\, H_j)}{P(E_i)}$$

"The condition on E_i changes in percentage the probability of H_j as the probability of E_i is changed in percentage by the condition H_j."

Symmetric conditioning

Let us take basic rule 4, written in terms of hypotheses H_j and effects E_i, and rewrite it this way:

$$\frac{P(H_j \,|\, E_i)}{P(H_j)} = \frac{P(E_i \,|\, H_j)}{P(E_i)}$$

"The condition on E_i changes in percentage the probability of H_j as the probability of E_i is changed in percentage by the condition H_j."

It follows

$$P(H_j \,|\, E_i) = \frac{P(E_i \,|\, H_j)}{P(E_i)}\, P(H_j)$$

Symmetric conditioning

Let us take basic rule 4, written in terms of hypotheses H_j and effects E_i, and rewrite it this way:

$$\frac{P(H_j\,|\,E_i)}{P(H_j)} = \frac{P(E_i\,|\,H_j)}{P(E_i)}$$

"The condition on E_i changes in percentage the probability of H_j as the probability of E_i is changed in percentage by the condition H_j."

It follows

$$P(H_j\,|\,E_i) = \frac{P(E_i\,|\,H_j)}{P(E_i)}\,P(H_j)$$

Got 'after' Calculated 'before'

(where 'before' and 'after' refer to the knowledge that E_i is true.)

Symmetric conditioning

Let us take basic rule 4, written in terms of hypotheses H_j and effects E_i, and rewrite it this way:

$$\frac{P(H_j \,|\, E_i)}{P(H_j)} = \frac{P(E_i \,|\, H_j)}{P(E_i)}$$

"The condition on E_i changes in percentage the probability of H_j as the probability of E_i is changed in percentage by the condition H_j."

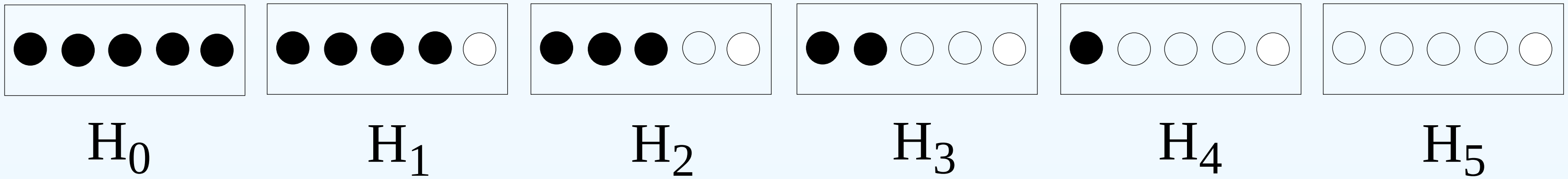
It follows

$$P(H_j \,|\, E_i) = \frac{P(E_i \,|\, H_j)}{P(E_i)}\, P(H_j)$$

"post illa observationes" *"ante illa observationes"*

(Gauss)

Application to the six box problem



Remind:

- E_1 = White
- E_2 = Black

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

Our prior belief about H_j

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)} \, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

Probability of E_i under a well defined hypothesis H_j
It corresponds to the 'response of the apparatus in measurements.
$\rightarrow$ likelihood (traditional, rather confusing name!)

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j\,|\,E_i, I) = \frac{P(E_i\,|\,H_j, I)}{P(E_i\,|\,I)}\,P(H_j\,|\,I)}$$

- $P(H_j\,|\,I) = 1/6$
- $P(E_i\,|\,I) = 1/2$
- $P(E_i\,|\,H_j,\,I)$:

$$\begin{aligned} P(E_1\,|\,H_j,\,I) &= j/5 \\ P(E_2\,|\,H_j,\,I) &= (5-j)/5 \end{aligned}$$

Probability of E_i taking account all possible H_j
$\rightarrow$ How much we are confident that E_i will occur.

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)} P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

Probability of E_i taking account all possible H_j
$\rightarrow$ How much we are confident that E_i will occur.
Easy in this case, because of the symmetry of the problem.
But already after the first extraction of a ball our opinion about the box content will change, and symmetry will break.

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

But it easy to prove that $P(E_i \,|\, I)$ is related to the other ingredients, usually easier to ‘measure’ or to assess somehow, though vaguely

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j \,|\, E_i, I) = \frac{P(E_i \,|\, H_j, I)}{P(E_i \,|\, I)}\, P(H_j \,|\, I)}$$

- $P(H_j \,|\, I) = 1/6$
- $P(E_i \,|\, I) = 1/2$
- $P(E_i \,|\, H_j,\, I)$:

$$\begin{aligned} P(E_1 \,|\, H_j,\, I) &= j/5 \\ P(E_2 \,|\, H_j,\, I) &= (5-j)/5 \end{aligned}$$

But it easy to prove that $P(E_i \,|\, I)$ is related to the other ingredients, usually easier to 'measure' or to assess somehow, though vaguely

'decomposition law': $P(E_i \,|\, I) = \sum_j P(E_i \,|\, H_j,\, I) \cdot P(H_j \,|\, I)$

($\rightarrow$ Easy to check that it gives $P(E_i \,|\, I) = 1/2$ in our case).

Collecting the pieces of information we need

Our tool:

$$\boxed{P(H_j\,|\,E_i,I) = \frac{P(E_i\,|\,H_j,I)\cdot P(H_j\,|\,I)}{\sum_j P(E_i\,|\,H_j,I)\cdot P(H_j\,|\,I)}}$$

- $P(H_j\,|\,I) = 1/6$
- $P(E_i\,|\,I) = \sum_j P(E_i\,|\,H_j,\,I)\cdot P(H_j\,|\,I)$
- $P(E_i\,|\,H_j,\,I)$:

$$\begin{aligned} P(E_1\,|\,H_j,\,I) &= j/5 \\ P(E_2\,|\,H_j,\,I) &= (5-j)/5 \end{aligned}$$

We are ready!

$\longrightarrow$ R program

First extraction

After first extraction (and reintroduction) of the ball:

- $P(H_j)$ changes
- $P(E_j)$ for next extraction changes

Note: The box is exactly in the same status as before

First extraction

After first extraction (and reintroduction) of the ball:

- $P(H_j)$ changes
- $P(E_j)$ for next extraction changes

Note: The box is exactly in the same status as before

Where is probability?

$\longrightarrow$ Certainly not in the box!

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Neglecting the background state of information I:

$$\frac{P(H_j\,|\,E_i)}{P(H_j)} = \frac{P(E_i\,|\,H_j)}{P(E_i)}$$

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Neglecting the background state of information I:

$$\frac{P(H_j \,|\, E_i)}{P(H_j)} = \frac{P(E_i \,|\, H_j)}{P(E_i)}$$

$$P(H_j \,|\, E_i) = \frac{P(E_i \,|\, H_j)}{P(E_i)} \, P(H_j)$$

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Neglecting the background state of information I:

$$\begin{aligned}
\frac{P(H_j\,|\,E_i)}{P(H_j)} &= \frac{P(E_i\,|\,H_j)}{P(E_i)} \\
P(H_j\,|\,E_i) &= \frac{P(E_i\,|\,H_j)}{P(E_i)}\,P(H_j) \\
P(H_j\,|\,E_i) &= \frac{P(E_i\,|\,H_j)\cdot P(H_j)}{\sum_j P(E_i\,|\,H_j)\cdot P(H_j)}
\end{aligned}$$

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Neglecting the background state of information I:

$$
\begin{aligned}
\frac{P(H_j \,|\, E_i)}{P(H_j)} &= \frac{P(E_i \,|\, H_j)}{P(E_i)} \\
P(H_j \,|\, E_i) &= \frac{P(E_i \,|\, H_j)}{P(E_i)}\, P(H_j) \\
P(H_j \,|\, E_i) &= \frac{P(E_i \,|\, H_j) \cdot P(H_j)}{\sum_j P(E_i \,|\, H_j) \cdot P(H_j)} \\
P(H_j \,|\, E_i) &\propto P(E_i \,|\, H_j) \cdot P(H_j)
\end{aligned}
$$

Bayes theorem

The formulae used to *infer* H_i and
to *predict* $E_j^{(2)}$ are related to the name of Bayes

Neglecting the background state of information I:

$$\frac{P(H_j \,|\, E_i)}{P(H_j)} = \frac{P(E_i \,|\, H_j)}{P(E_i)}$$

$$P(H_j \,|\, E_i) = \frac{P(E_i \,|\, H_j)}{P(E_i)} P(H_j)$$

$$P(H_j \,|\, E_i) = \frac{P(E_i \,|\, H_j) \cdot P(H_j)}{\sum_j P(E_i \,|\, H_j) \cdot P(H_j)}$$

$$P(H_j \,|\, E_i) \propto P(E_i \,|\, H_j) \cdot P(H_j)$$

Different ways to write the

Bayes' Theorem

Updating the knowledge by new observations

Let us repeat the experiment:

$\boxed{\text{Sequential use of Bayes theorem}}$

Old posterior becomes new prior, and so on

Updating the knowledge by new observations

Let us repeat the experiment:

Sequential use of Bayes theorem

Old posterior becomes new prior, and so on

$$P(H_j \,|\, E^{(1)}, E^{(2)}) \;\;\propto\;\; P(E^{(2)} \,|\, H_j, E^{(1)}) \cdot P(H_j \,|\, E^{(1)})$$

Updating the knowledge by new observations

Let us repeat the experiment:

Sequential use of Bayes theorem

Old posterior becomes new prior, and so on

$$
\begin{aligned}
P(H_j \,|\, E^{(1)}, E^{(2)}) &\propto P(E^{(2)} \,|\, H_j, E^{(1)}) \cdot P(H_j \,|\, E^{(1)}) \\
&\propto P(E^{(2)} \,|\, H_j) \cdot P(H_j \,|\, E^{(1)})
\end{aligned}
$$

Updating the knowledge by new observations

Let us repeat the experiment:

Sequential use of Bayes theorem

Old posterior becomes new prior, and so on

$$\begin{aligned} P(H_j\,|\,E^{(1)},E^{(2)}) &\propto P(E^{(2)}\,|\,H_j,E^{(1)})\cdot P(H_j\,|\,E^{(1)}) \\ &\propto P(E^{(2)}\,|\,H_j)\cdot P(H_j\,|\,E^{(1)}) \\ &\propto P(E^{(2)}\,|\,H_j)\cdot P(E^{(1)}\,|\,H_j)\cdot P_0(H_j) \end{aligned}$$

Updating the knowledge by new observations

Let us repeat the experiment:

$\boxed{\text{Sequential use of Bayes theorem}}$

Old posterior becomes new prior, and so on

$$
\begin{aligned}
P(H_j \,|\, E^{(1)}, E^{(2)}) &\propto P(E^{(2)} \,|\, H_j, E^{(1)}) \cdot P(H_j \,|\, E^{(1)}) \\
&\propto P(E^{(2)} \,|\, H_j) \cdot P(H_j \,|\, E^{(1)}) \\
&\propto P(E^{(2)} \,|\, H_j) \cdot P(E^{(1)} \,|\, H_j) \cdot P_0(H_j) \\
&\propto P(E^{(1)}, E^{(1)} \,|\, H_j) \cdot P_0(H_j)
\end{aligned}
$$

Updating the knowledge by new observations

Let us repeat the experiment:

Sequential use of Bayes theorem

Old posterior becomes new prior, and so on

$$\begin{aligned}
P(H_j\,|\,E^{(1)},E^{(2)}) &\propto P(E^{(2)}\,|\,H_j,E^{(1)})\cdot P(H_j\,|\,E^{(1)}) \\
&\propto P(E^{(2)}\,|\,H_j)\cdot P(H_j\,|\,E^{(1)}) \\
&\propto P(E^{(2)}\,|\,H_j)\cdot P(E^{(1)}\,|\,H_j)\cdot P_0(H_j) \\
&\propto P(E^{(1)},E^{(1)}\,|\,H_j)\cdot P_0(H_j) \\
P(H_j\,|\,\text{data}) &\propto P(\text{data}\,|\,H_j)\cdot P_0(H_j)
\end{aligned}$$

Updating the knowledge by new observations

Let us repeat the experiment:

$\boxed{\text{Sequential use of Bayes theorem}}$

Old posterior becomes new prior, and so on

$$
\begin{aligned}
P(H_j\,|\,E^{(1)},E^{(2)}) &\propto P(E^{(2)}\,|\,H_j,E^{(1)})\cdot P(H_j\,|\,E^{(1)}) \\
&\propto P(E^{(2)}\,|\,H_j)\cdot P(H_j\,|\,E^{(1)}) \\
&\propto P(E^{(2)}\,|\,H_j)\cdot P(E^{(1)}\,|\,H_j)\cdot P_0(H_j) \\
&\propto P(E^{(1)},E^{(1)}\,|\,H_j)\cdot P_0(H_j) \\
P(H_j\,|\,\text{data}) &\propto P(\text{data}\,|\,H_j)\cdot P_0(H_j)
\end{aligned}
$$

$\boxed{\text{Bayesian inference}}$

Updating the knowledge by new observations

Let us repeat the experiment:

Sequential use of Bayes theorem

Old posterior becomes new prior, and so on

$$
\begin{aligned}
P(H_j \,|\, E^{(1)}, E^{(2)}) &\propto P(E^{(2)} \,|\, H_j, E^{(1)}) \cdot P(H_j \,|\, E^{(1)}) \\
&\propto P(E^{(2)} \,|\, H_j) \cdot P(H_j \,|\, E^{(1)}) \\
&\propto P(E^{(2)} \,|\, H_j) \cdot P(E^{(1)} \,|\, H_j) \cdot P_0(H_j) \\
&\propto P(E^{(1)}, E^{(1)} \,|\, H_j) \cdot P_0(H_j) \\
P(H_j \,|\, \text{data}) &\propto P(\text{data} \,|\, H_j) \cdot P_0(H_j)
\end{aligned}
$$

Learning from data using probability theory