Reasoning for Humans: Clear Thinking in an Uncertain World

PHIL 171

Eric Pacuit

Department of Philosophy
University of Maryland
pacuit.org

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. 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?

D. Kahneman. *Thinking, Fast and Slow*. Macmillan, 2011.

Base-Rate Fallacy

E	The witness identified the cab as blue.
H	A cab from the blue cab company was in the accident.
$\neg H$	A cab from the green cab company was in the accident.

What is $Pr(H \mid E)$?

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.

E	The witness identified the cab as blue.
H	A cab from the blue cab company was in the accident.
$\neg H$	A cab from the green cab company was in the accident.

$$Pr(H) = 0.15 \qquad Pr(\neg H) = 0.85$$

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.

E	The witness identified the cab as blue.
H	A cab from the blue cab company was in the accident.
$\neg H$	A cab from the green cab company was in the accident.

$$Pr(E \mid H) = 0.8 \qquad Pr(E \mid \neg H) = 0.2$$

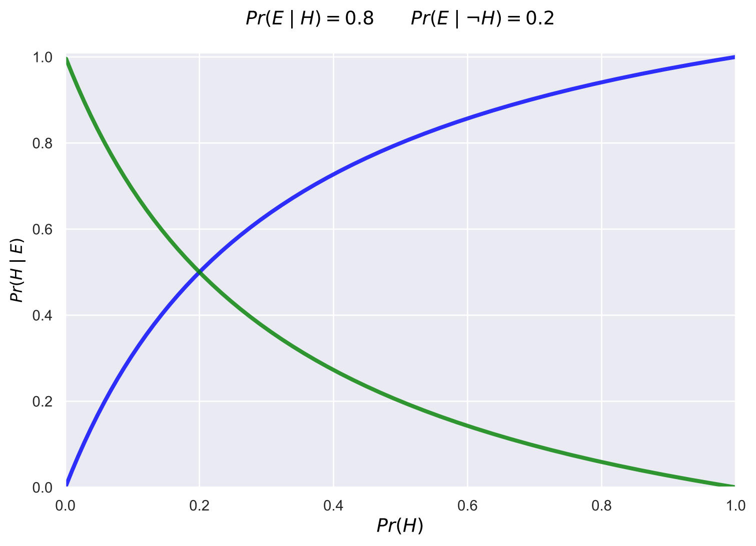
E	The witness identified the cab as blue.
H	A cab from the blue cab company was in the accident.
$\neg H$	A cab from the green cab company was in the accident.

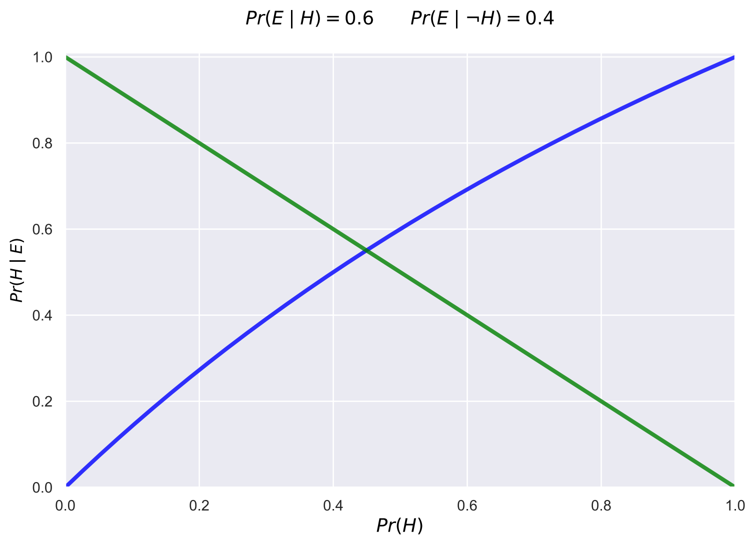
$$Pr(H) = 0.15 \qquad Pr(\neg H) = 0.85$$

$$Pr(E \mid H) = 0.8 \qquad Pr(E \mid \neg H) = 0.2$$

$$Pr(H \mid E) = Pr(E \mid H)\frac{Pr(H)}{Pr(E)} = 0.8\frac{0.15}{0.15*0.8+0.85*0.2} \approx 0.41$$

$$Pr(\neg H \mid E) = Pr(E \mid \neg H)\frac{Pr(\neg H)}{Pr(E)} = 0.2\frac{0.85}{0.15*0.8+0.85*0.2} \approx 0.59$$


$Pr(E \mid H) = 0.8$ $Pr(E \mid \neg H) = 0.2$
1.0
0.8
0.6
0.4
0.2
0.0
$Pr(H \mid E)$
0.0
0.2
0.4
0.6
0.8
1.0
$Pr(H)$


$Pr(E \mid H) = 0.6 \qquad Pr(E \mid \neg H) = 0.4$
1.0
0.8
0.6
0.4
0.2
0.0
$Pr(H \mid E)$
0.0
0.2
0.4
0.6
0.8
1.0
$Pr(H)$

$$Pr(H \mid E) = Pr(E \mid H)\frac{Pr(H)}{Pr(E)}$$

The **Base Rate Fallacy** is our tendency to give more weight to the event-specific information than we should, and sometimes even ignore base rates entirely.

$$Pr(H \mid E) = Pr(E \mid H)\frac{Pr(H)}{Pr(E)}$$

The **Base Rate Fallacy** is our tendency to give more weight to the event-specific information than we should, and sometimes even ignore base rates entirely.

Measuring Arguments

How do we *measure* the strength of an argument?

1. X evidentially supports Y
2. X is positively relevant to Y.

$Pr(Y \mid X)$ measures the evidential support of the argument.
But, how do we measure the relevance of X to Y?

How do we *measure* the strength of an argument?

1. X evidentially supports Y
2. X is positively relevant to Y.

$Pr(Y \mid X)$ measures the evidential support of the argument.
But, how do we measure the relevance of X to Y?
$d(X, Y) = Pr(Y \mid X) - Pr(Y)$

The Paradox of the Ravens

The Paradox of the Ravens (1)

(IC) A hypothesis of the form "All As are Bs" is confirmed by any positive instance, i.e., any instance that is both A and B.

- A black raven confirms that all ravens are black.
- A green emerald confirms that all emeralds are green.

The Paradox of the Ravens (2)

(EQ) If H and H' are logically equivalent, then if E confirms H, then E confirms H'.

H: All ravens are black.

H': All non-black things are non-ravens.

The Paradox of the Ravens (3)

But, then does a brown jacket confirm H?

1. (IC) implies that a brown jacket confirms that "all non-black things are non-ravens".
2. "all non-black things are non-ravens" is equivalent to "all ravens are black".
3. (EC) implies that a brown jacket confirms that "all ravens are black".

The Paradox of the Ravens (3)

But, then does a brown jacket confirm H?

1. (IC) implies that a brown jacket confirms that "all non-black things are non-ravens".
2. "all non-black things are non-ravens" is equivalent to "all ravens are black".
3. (EC) implies that a brown jacket confirms that "all ravens are black".

We can run the same argument using a blue jacket, red carpet, white chair,. . .

But, surely you can't learn something about the color of ravens by looking around your office.

Wason Selection Task and the Paradox of Confirmation

L. Humberstone. *Hempel Meets Wason*. Erkenntnis 41 (1994), 391 - 402.

B. Fitelson and J. Hawthorne. *The Wason Selection Task(s) and the Paradox of Confirmation*. Philosophical Perspectives, Volume 24, Issue 1, pages 207 - 241, 2010.