Reasoning for Humans: Clear Thinking in an Uncertain World

PHIL 171

Eric Pacuit

Department of Philosophy
University of Maryland
pacuit.org

Bayes Theorem

Posterior

Prior of H

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

Likelihood

Prior of observing E

Suppose that P_1, P_2 and P_3 are three propositions where:

1. $Pr(P_1) > 0$, $Pr(P_2) > 0$ and $Pr(P_3) > 0$
2. $Pr(P_1 \vee P_2 \vee P_3) = 1$

$$Pr(X) = Pr(P_1)Pr(X \mid P_1) + Pr(P_2)Pr(X \mid P_2) + Pr(P_3)Pr(X \mid P_3)$$

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

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

Monty Hall Dilemma

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall (1)

H_1: The car is behind door 1

H_2: The car is behind door 2

H_3: The car is behind door 3

Monty Hall (2)

Reasoning 1: E: The car is not behind door 3 ($\neg H_3 \leftrightarrow H_1 \vee H_2$)

Monty Hall (2)

Reasoning 1: E: The car is not behind door 3 ($\neg H_3 \leftrightarrow H_1 \vee H_2$)

$$Pr(H_1 \mid E) \quad = \quad Pr(E \mid H_1)\frac{Pr(H_1)}{Pr(E)}$$

Monty Hall (2)

Reasoning 1: E: The car is not behind door 3 $(\neg H_3 \leftrightarrow H_1 \vee H_2)$

$$\begin{aligned} Pr(H_1 \mid E) &= Pr(E \mid H_1)\frac{Pr(H_1)}{Pr(E)} \\ &= \frac{Pr(E \mid H_1)Pr(H_1)}{Pr(E \mid H_1)Pr(H_1)+Pr(E \mid H_2)Pr(H_2)+Pr(E \mid H_3)Pr(H_3)} \end{aligned}$$

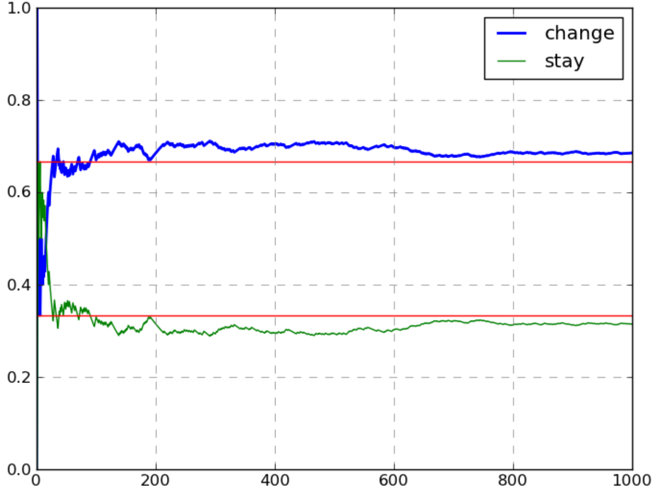
Monty Hall (2)

Reasoning 1: E: The car is not behind door 3 ($\neg H_3 \leftrightarrow H_1 \vee H_2$)

$$
\begin{aligned}
Pr(H_1 \mid E) &= Pr(E \mid H_1)\frac{Pr(H_1)}{Pr(E)} \\
&= \frac{Pr(E \mid H_1)Pr(H_1)}{Pr(E \mid H_1)Pr(H_1) + Pr(E \mid H_2)Pr(H_2) + Pr(E \mid H_3)Pr(H_3)} \\
&= 1 \cdot \frac{\frac{1}{3}}{1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} \\
&= 1 \cdot \frac{\frac{1}{3}}{\frac{2}{3}} \\
&= \frac{1}{2}
\end{aligned}
$$

Similarly for $Pr(H_2 \mid E)$, so do not switch.

Monty Hall: Reasoning 1 vs. Reasoning 2



Monty Hall (3)

Reasoning 2: F: Monty opened door number 3

$$
\begin{aligned}
Pr(H_2 \mid F) &= Pr(F \mid H_2)\frac{Pr(H_2)}{Pr(F)} \\
&= \frac{Pr(F \mid H_2)Pr(H_2)}{Pr(F \mid H_1)Pr(H_1)+Pr(F \mid H_2)Pr(H_2)+Pr(F \mid H_3)Pr(H_3)} \\
&= 1 \cdot \frac{\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+1\cdot\frac{1}{3}+0\cdot\frac{1}{3}} \\
&= 1 \cdot \frac{\frac{1}{3}}{\frac{1}{2}} \\
&= \frac{2}{3}
\end{aligned}
$$

So, $Pr(H_1 \mid F) = \frac{1}{3}$ and $Pr(H_2 \mid F) = \frac{2}{3}$, so you should switch

Monty Hall (3)

Reasoning 2: F: Monty opened door number 3

$$
\begin{aligned}
Pr(H_2 \mid F) &= Pr(F \mid H_2)\frac{Pr(H_2)}{Pr(F)} \\
&= \frac{Pr(F \mid H_2)Pr(H_2)}{Pr(F \mid H_1)Pr(H_1)+Pr(F \mid H_2)Pr(H_2)+Pr(F \mid H_3)Pr(H_3)} \\
&= 1 \cdot \frac{\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+1\cdot\frac{1}{3}+0\cdot\frac{1}{3}} \\
&= 1 \cdot \frac{\frac{1}{3}}{\frac{1}{2}} \\
&= \frac{2}{3}
\end{aligned}
$$

So, $Pr(H_1 \mid F) = \frac{1}{3}$ and $Pr(H_2 \mid F) = \frac{2}{3}$, so you should switch

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.

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

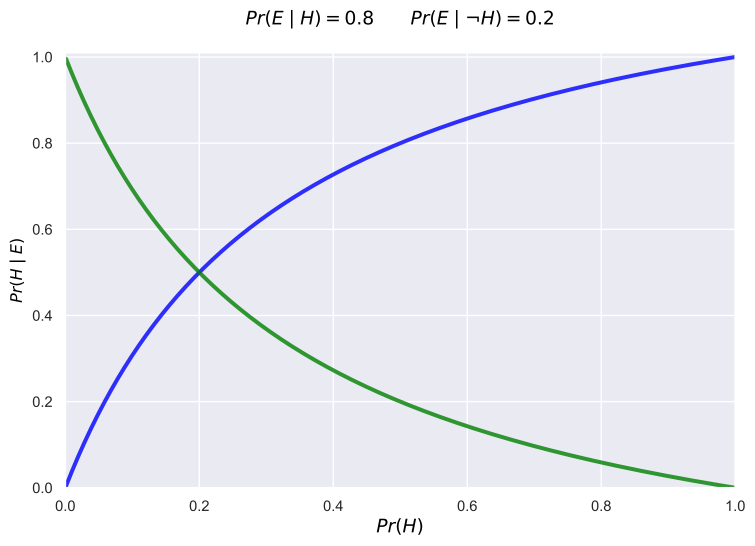
E	The witness identified the cab as blue.
H	A cab from the blue cab company was in the accident.
$\neg H$	A cab from the green cab company was in the accident.

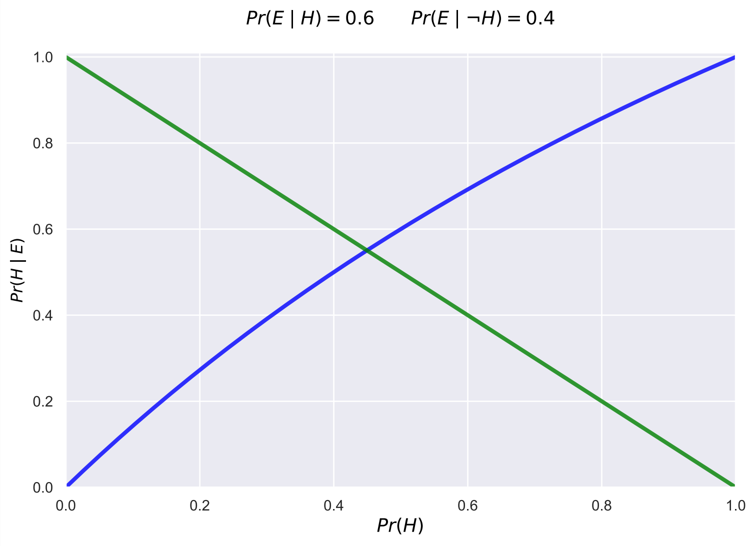
$$Pr(H) = 0.15 \qquad Pr(\neg H) = 0.85$$

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

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

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


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


$Pr(E \mid H) = 0.6$ $Pr(E \mid \neg H) = 0.4$
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.