Reasoning for Humans: Clear Thinking in an Uncertain World

PHIL 171

Eric Pacuit

Department of Philosophy
University of Maryland
pacuit.org

Valid Arguments

$$X \to Y, Z \to Y \models (X \vee Z) \to Y$$

$$X \to Y, \neg X \to Y \models Y$$

If $X \models Y$ and $\neg X \models Y$, then Y is a tautology

"P evidentially supports Q given R" means the same thing as "$P \wedge R$ evidentially supports Q"

P evidentially supports Q given R	$Pr(Q \mid P \wedge R) > 0.5$
$P \wedge R$ evidentially supports Q	$Pr(Q \mid P \wedge R) > 0.5$

"P is positively relevant to Q given R" does not mean the same thing as "$P \wedge R$ is positively relevant to Q"

P is positively relevant to Q given R	$Pr(Q \mid P \wedge R) > Pr(Q \mid R)$
$P \wedge R$ is positively relevant to Q	$Pr(Q \mid P \wedge R) > Pr(Q)$

Sure-thing Principle: If

1. P is "good evidence" for Q given R, and
2. P is "good evidence" for Q given $\neg R$

then P is "good evidence" for Q

Unconditional Sure-thing Principle: If

1. $P \wedge R$ is "good evidence" for Q and
2. $P \wedge \neg R$ is "good evidence" for Q

then P is "good evidence" for Q

Sure-thing Principle: Evidential Support

In any stochastic truth table, if

1. $Pr(Q \mid P \wedge R) > 0.5$; and
2. $Pr(Q \mid P \wedge \neg R) > 0.5$

then $Pr(Q \mid P) > 0.5$

(assuming that all the conditional probabilities are defined)

Unconditional Sure-thing Principle: Positive Support

In any stochastic truth table, if

1. $Pr(Q \mid P \wedge R) > Pr(Q)$; and
2. $Pr(Q \mid P \wedge \neg R) > Pr(Q)$

then $Pr(Q \mid P) > Pr(Q)$

(assuming that all the conditional probabilities are defined)

Sure-thing Principle: Positive Support

There is a stochastic truth table such that

1. $Pr(Q \mid P \wedge R) > Pr(Q \mid R)$; and
2. $Pr(Q \mid P \wedge \neg R) > Pr(Q \mid \neg R)$

but $Pr(Q \mid P) \leq Pr(Q)$

(assuming that all the conditional probabilities are defined)

	P	Q	R
0.2	T	T	T
0.125	T	T	F
0.05	T	F	T
0.075	T	F	F
0.4	F	T	T
0	F	T	F
0.11	F	F	T
0.04	F	F	F

$$Pr(Q \mid P \wedge R) = 0.8$$
$$Pr(Q \mid R) \approx 0.789$$

$$Pr(Q \mid P \wedge \neg R) = 0.625$$
$$Pr(Q \mid \neg R) \approx 0.521$$

$$Pr(Q \mid P) \approx 0.722$$
$$Pr(Q) = 0.725$$

Simpson's Paradox

Suppose that a University is hiring in Philosophy and Mathematics.
13 men and 13 women apply for jobs.

	Men	Women	
Mathematics	1/5	2/8	*success rate better for women*
Philosophy	6/8	4/5	*success rate better for women*

Simpson's Paradox

Suppose that a University is hiring in Philosophy and Mathematics.
13 men and 13 women apply for jobs.

	Men	Women	
Mathematics	1/5	2/8	*success rate better for women*
Philosophy	6/8	4/5	*success rate better for women*
University	7/13	6/13	*success rate better for men*

How can it be that each Department favors women applicants, and yet overall men fare better than women?

G. Malinas and J. Bigelow. *Simpson's Paradox*. The Stanford Encyclopedia of Philosophy, Fall 2012 Edition, Edward N. Zalta (ed.).

Summary

- If $X \models Y$ and $\neg X \models Y$, then $\models Y$.

- Both the Sure-thing Principle and the Unconditional Sure-Thing Principle holds for evidential support:

  If $Pr(Q \mid P \wedge R) > 0.5$ and $Pr(Q \mid P \wedge \neg R) > 0.5$, then $Pr(Q \mid P) > 0.5$

- The Sure-thing Principle does not hold for positive relevance:

  There are stochastic truth tables where: $Pr(Q \mid P \wedge R) > Pr(Q \mid R)$ and $Pr(Q \mid P \wedge \neg R) > Pr(Q \mid \neg R)$, but $Pr(Q \mid P) \leq Pr(Q)$

- The Unconditional Sure-Thing Principle holds for positive relevance:

  If $Pr(Q \mid P \wedge R) > Pr(Q)$ and $Pr(Q \mid P \wedge \neg R) > Pr(Q)$, then $Pr(Q \mid P) > Pr(Q)$

Bayes Theorem

Conditional Probability

The probability of H *given* E, $Pr(H \mid E)$, is defined to be

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

provided $Pr(E) > 0$.

Bayes Theorem

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

Bayes Theorem

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

Bayes theorem is important because it expresses the quantity $Pr(H \mid E)$ (the probability of a hypothesis H given the evidence E)—which is something people often find hard to assess—in terms of quantities that can be drawn directly from experiential knowledge.

Example: Suppose you are in a casino and you hear a person at the next gambling table announce "Twelve". We want to know wether he was rolling a pair of dice or a roulette wheel.

That is, compare $Pr(Dice \mid Twelve)$ with $Pr(Roulette \mid Twelve)$.

Based on our background knowledge of gambling we have $Pr(Twelve \mid Dice) = 1/36$ and $Pr(Twelve \mid Roulette) = 1/38$.

Based on our observations about the casino, we can judge the prior probabilities $Pr(Dice)$ and $Pr(Roulette)$.

But this is now enough to *calculate* the required probabilities.

Posterior

Prior of H

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

Likelihood

Prior of observing E

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

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

Bayes Theorem

Posterior

Prior of H

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

Likelihood

Prior of observing E

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

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

Suppose you select one card from a standard deck of cards.

Q is "the card is a queen card" and F is "the card is a face card".

What is $Pr(Q \mid F)$?

	Q	F	$Q \wedge F$
$\frac{4}{52}$	T	T	T
0	T	F	F
$\frac{8}{52}$	F	T	F
$\frac{40}{52}$	F	F	F

$$Pr(Q \mid F) = \frac{Pr(Q \wedge F)}{Pr(F)} = \frac{\frac{4}{52}}{\frac{12}{52}} = \frac{4}{12} = \frac{1}{3}$$

	Q	F	$Q \wedge F$
$\frac{4}{52}$	T	T	T
0	T	F	F
$\frac{8}{52}$	F	T	F
$\frac{40}{52}$	F	F	F

$$Pr(Q \mid F) = \frac{Pr(Q \wedge F)}{Pr(F)} = \frac{\frac{4}{52}}{\frac{12}{52}} = \frac{4}{12} = \frac{1}{3}$$

Note that:

$Pr(F \mid Q) = 1$

$Pr(F) = \frac{12}{52}$

$Pr(Q) = \frac{4}{52}$

$$Pr(Q \mid F) = Pr(F \mid Q)\frac{Pr(Q)}{Pr(F)} = 1 \times \frac{\frac{4}{52}}{\frac{12}{52}} = \frac{4}{12} = \frac{1}{3}$$