The psychology of conditionals
Class 4
Bayesian psychology of conditional reasoning

David Over
Psychology Department
Durham University, UK

The probability conditional

The Equation: P(if p then q) = P(q|p)

A conditional that satisfies the Equation has been called a *probability conditional* (Adams, 1998) and a *conditional event* (de Finetti, 1936). It has *P(q|p)* as its expected value.

For advanced studies this conditional, see as well Gilio (2002), Kleiter, Fugard, & Pfeifer (2018), Pfeifer & Kleiter (2009), and Sanfilippo et al. (2018).

The probability conditional: Further points

P(if p then q) = P(q|p)

P(if p then q) is the probabilty of *p & q* supposing that *if p then q* makes an assertion: P((p & q)|p) = P(q|p).

Note that the probability conditional does <u>not</u> semantically mean that P(q|p) is high. It is a <u>pragmatic</u> point that the <u>assertion</u> of a conditional suggests that P(q|p) is high.

The Ramsey test and de Finetti table

The Ramsey test and the de Finetti table, with its extension as the Jeffrey table, are the pillars that support Bayesian / probabilistic accounts of human conditional reasoning. In such accounts, indicative conditionals, conditional bets, and conditional probabilities should all be closely related to each other.

Are they closely related to each other in people's judgments?

The conditional probability hypothesis

The Equation becomes <u>the conditional probability hypothesis</u> in the psychology of reasoning, P(if p then q) = P(q|p) .

Early confirmation of this hypothesis was provided by Evans Handley, & Over (2003), Oberauer & Wilhelm (2003), and Fugard et al. (2011).

Participants were presented in these early studies with a frequency distribution, and asked to make their probability judgments on that basis.

Politzer et al. (2010)
Indicative conditional

This drawing represents chips

● ● ○

■ ■ ■ □

A chip is chosen at random. Consider the following sentence:

If the chip is square then it is black.

What are the chances the sentence is true?

Politzer et al. (2010)
Conditional bet

This drawing represents chips

● ● ○

■ ■ ■ □

A chip is chosen at random. Marie bets Pierre 1 euro that:

If the chip is square then it is black.

What are the chances that Marie wins her bet?

Politzer et al. (2010): The results

As both Ramsey and de Finetti argued, indicative conditionals and conditional bets were close in people's judgments : "not-s" cases were "irrelevant" and "void".

The majority judged the chances that *if s then b* is true, and that Marie's bet on it is won, to be P(b|s).

A small minority judged the chances to be P(s & b). Almost no one judged these chances to be P(not-s or b).

Note that Marie's bet is not fair

The expected value of a fair bet is 0. Our conditional bet would be fair if P(b|s) = 0.5, but in fact P(b|s) = 0.75. The expected value of Marie's bet for her is:

P(s & b)(1) + P(s & not-b)(-1) + P(not-s)(0)

(.43)(1) + (.14)(-1) + (.43)(0) = 29 cents

For this bet to be fair, the odds should be 3 to 1, which corresponds to P(b|s) = 0.75.

Cognitive ability

People who give P(p & q) as the answer to a question about the probability of truth of ***if p then q*****, or of winning a bet on it, are of relatively low cognitive ability.**

People who give P(q|p) as the answer to a question about the probability of truth of "if p then q", or of winning a bet on it, are of relatively high cognitive ability.

See Evans et al. (2007), and Politzer et al. (2010)

Causal conditionals

"If global warming continues, then London will be flooded."

The Equation also holds for "causal" conditionals and related counterfactuals (Over et al., 2007; Singmann et al., 2014).

A Ramsey test of such a conditional could use a causal model for the judgment, and this model could be a Bayesian network (Oaksford & Chater, 2013).

Missing-link conditionals

"If raccoons have no wings, they cannot breathe under water." (Krzyzanowska, Collins, & Hahn, 2017).

There are arguments that the Equation does not apply to such conditionals, and supporting psychological results (Douven, 2016, 2017; Douven et al, 2018; Skovgaard-Olsen et al., 2017).

But there is also evidence that people even dislike missing-link disjunctions, and this raises the question of whether the effect is semantic or pragmatic (Cruz et al., 2015).

A note on axioms

In the Bayesian approach, the relevant normative standard is axiomatic probability theory.

In the de Finetti tradition, the conditional probability of *q* given *p* is not defined to be P(p & q) / P(p), which is undefined when P(p) = 0. Rather, conditional probability is primitive, and it is presupposed that P(q|p) is fixed outside the axiom system by the Ramsey test (Pfeifer & Kleiter, 2009).

People appear to make sensible judgments about P(q|p) when P(p) = 0, e.g. when they use counterfactuals.

Dutch books:
The underlying justification

A *Dutch book* is a series of bets that the bettor can only lose.

If bettors violate the axioms of probability theory, i.e. they are incoherent, then a Dutch book can be made against them (Vineberg, 2016).

There is a debate about this formal justification of the axioms of probability theory. But suppose someone loses bet after bet after bet as a result of judging that $P(p \,\&\, q) > P(q)$?

Generalizing consistency and validity

For reasoning under uncertainty, binary consistency should be generalized to <u>coherence</u>: being consistent with probability theory.

Binary truth-preserving validity should be generalized probabilistic validity: <u>p-validity</u>.

Assumptions and belief

Traditional psychology of reasoning was assumption-based, and Bayesian approaches are belief-based.

In the tradition, participants were asked to assume that given premises were true and then to state what necessarily followed.

Bayesian approaches note that almost all inference in everyday life and science is from uncertain premises. It is usually dynamic inference from degrees of belief to degrees of belief that are lower than certainty. To assess this reasoning, we need coherence and p-validity.

Coherence and conjunction

Coherence can be seen as a generalization of the binary notion of consistency, giving us intervals for belief.

It is binary inconsistent to believe *p & q* but not *p*.

More generally, where the degree of belief in *p* is P(p) and where the degree of belief in *q* is P(q):

$$\min\{P(p), P(q)\} \geq P(p \,\&\, q) \geq \max\{0, P(p) + P(q) - 1\}$$

It is incoherent to fall outside these intervals.

Tversky & Kahneman (1983) again

Linda is single, outspoken, and intelligent. She majored in Philosophy at university, was concerned with social justice, and was anti-nuclear. Rank the following in probability:

Linda is a bank teller.

Linda is a social worker.

Linda is a feminist and a bank teller.

Linda is a farmer.

The conjunction fallacy again

Participants in experiments tend to judge P(f & t) > P(t), when they make judgments about Linda's qualities.

Tversky & Kahneman noted that judging P(f & t) > P(t) is incoherent because of the logical relation between *f & t* and *t*, but only with the coming of the new paradigm has account been taken of the much more general relation between probability and logical validity.

See Cruz et al. (2015).

Probabilistic validity defined

A single premise inference is p-valid if and only if the probability of its premise cannot be coherently greater than the probability of its conclusion.

More generally and more precisely, let the uncertainty of any premise or conclusion *s* be 1 - P(s). Then an inference is p-valid if and only if the uncertainty of its conclusion cannot be coherently greater than the sum of the uncertainties of its premises (Adams, 1998).

Supposing P(if p then q) = P(q|p), let the uncertainty of *if p then q* be 1 - P(q|p).

The coherence interval for MP

If Linda goes to the party (*p*), then she will drink too much (*q*). She will go to the party. Therefore, she will drink too much.

Let P(q|p) = .9 and P(p) = .8. By the total probability theorem:

P(q) = P(p)P(q|p) + P(not-p)P(q|not-p) = .72 + .2P(q|not-p)

P(q|not-p) is 0 at the minimum and 1 at the maximum, P(q) should fall in the interval [.72 , .92] for <u>coherence</u>. If people claim that $P(q) < .72$ or $P(q) > .92$, they are committing a fallacy like the Linda fallacy. Note that for the binary approach, confidence in the conclusion of MP cannot be too high.

Coherence intervals and p-validity

Coherence intervals have been given for a range of inferences (Cruz et al., 2015, 2016, 2017; Gilio & Over, 2012; Pfeifer & Kleiter, 2009).

There have been experimental studies of coherence intervals and p-validity as well, and people sometimes conform to these above the chance levels. See Evans, Thompson, & Over (2015) and Singmann, Klauer, & Over (2014).

“Suppressing” MP

Suppose some people judge that P(if p then q) = .9, P(p) = .8, and P(q) = .95.

Then these people are incoherent, and we can say that they have “suppressed” MP.

Note that the “suppression” of a valid inference is not the same as losing confidence in the conclusion. And that one can “suppress” a valid inference by being too confident in the conclusion, i.e. overconfident.

MP and dynamic reasoning

People conform to the p-validity of MP at above chance levels (Evans, Thompson, & Over, 2015).

Supposing the Equation, $P(\text{if } p \text{ then } q) = P(q|p)$, Bayesian conditionalization corresponds to dynamic MP.

We have a degree of belief at one time in the conditional, $P_1(\text{if } p \text{ then } q) = P_1(q|p)$. We learn p at a later time, $P_2(p) = 1$, and we infer a new, changed degree of belief in q, $P_2(q) = P_1(q|p)$. This reasoning is justified as long as invariance holds: $P_2(\text{if } p \text{ then } q) = P_1(\text{if } p \text{ then } q) = P_1(q|p)$.

Dynamic reasoning generally

Traditional psychology of reasoning did not focus on dynamic reasoning, which is not inference from assumptions, but is rather about belief change, revision, and updating (Oaksford & Chater, 2013).

There are deep questions about belief revision, including what it is to learn a conditional itself.

But an overlooked question is: "What happens to *if p then q* when we learn *not-p*?"

A murder mystery story

Suppose the police investigate the death of the master of an isolated country house in a typical "English" murder story. The evidence indicates that either the butler or the cook are responsible for the death. They conclude:

"If the butler did not do it, then the cook did."

The above inference is not p-valid. It may be well justified for the police, but it is not for the cook, who knows that she did not commit the murder. See Gilio & Over (2012).

Learning that the butler did it

Suppose the police discover that the butler was responsible for the murder. They will not use the counterfactual:

"If the butler had not done it, then the cook would have."

Instead, the police would use a counterfactual very close to the indicative conditional that the cook used from the start:

"If the butler had not done it, then there would not have been a murder."

Indicative or counterfactual?

When we have strong belief in an indicative conditional, *if p then q*, but learn *not-p*, how do we revise the conditional?

We may revise it simply as counterfactual, e.g. "If you put the litmus paper in the acid, it will turn red", but it appears that this does not always happen.

We can also revise a counterfactual as an indicative when its antecedent turns out to be true, and so MP becomes possible as an inference.

An experiment and looking ahead a final time

Students will now get the chance to take part in an experiment.

Nicole Cruz and I will jointly lead the class tomorrow. Our object will be to look more closely at counterfactuals, and to present the results of the experiment.

There will be an open discussion of the experimental results, and of what these might tell us about how people understand counterfactuals.